Journal de l'École polytechnique
Mathématiques

Nicolas Champagnat, Sylvie Méléard, Sepideh Mirrahimi, & Viet Chi Tran

Filling the gap between individual-based evolutionary models and Hamilton-Jacobi equations

Tome 10 (2023), p. 1247-1275.

https://doi.org/10.5802/jep.244

Publié avec le soutien
du Centre National de la Recherche Scientifique

Publication membre du
Centre Mersenne pour l'édition scientifique ouverte
www.centre-mersenne.org
e-ISSN : 2270-518X

Journal de l'École polytechnique
Mathématiques
Tome 10, 2023, p. 1247–1275 DOI: 10.5802/jep.244

FILLING THE GAP BETWEEN INDIVIDUAL-BASED EVOLUTIONARY MODELS AND HAMILTON-JACOBI EQUATIONS

by Nicolas Champagnat, Sylvie Méléard, Sepideh Mirrahimi & Viet Chi Tran

Abstract. — We consider a stochastic model for the evolution of a discrete population structured by a trait with values on a finite grid of the torus, and with mutation and selection. We focus on a parameter scaling where population is large, individual mutations are small but not rare, and the grid mesh is much smaller than the size of mutation steps. When considering the evolution of the population in long time scales, the contribution of small sub-populations may strongly influence the dynamics. Our main result quantifies the asymptotic dynamics of subpopulation sizes on a logarithmic scale. We establish that under our rescaling, the stochastic discrete process converges to the viscosity solution of a Hamilton-Jacobi equation. The proof makes use of almost sure maximum principles and careful control of the martingale parts.

Résumé (Des modèles stochastiques d'évolution aux équations de Hamilton-Jacobi)
Nous considérons un modèle stochastique pour l'évolution d'une population discrète structurée en trait à valeurs dans une grille finie du tore, avec mutation et sélection. On se place dans une limite d'échelle de grande population, de petites mutations (mais pas rares), et où le maillage tend vers zéro. En temps long, la contribution de petites sous-populations peut fortement influencer la dynamique. Nous montrons que dans ce cadre, le processus stochastique discret converge sur une échelle logarithmique vers la solution de viscosité d'une équation de Hamilton-Jacobi. La preuve fait appel à un principe du maximum presque-sûr et à des estimées fines des parties martingales.

Contents

Mathematical subject classification (2020). — 92D25, 92D15, 60J80, 60F99, 35F21.
Keywords. — Stochastic birth death models, large population approximation, selection, mutation, viscosity solution, maximum principle, Hamilton-Jacobi equation.

This work is funded by the European Union (ERC, SINGER, 101054787). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. This work has also been supported by the Chair "Modélisation Mathématique et Biodiversité" of Veolia Environnement-École Polytechnique-Museum National d'Histoire Naturelle-Fondation X. S.Mirrahimi was partially supported by the ANR project DEEV ANR-20-CE40-0011-01. V.C.T. also acknowledges support from Labex Bézout (ANR-10-LABX-58).

e-ISSN: 2270-518X http://jep.centre-mersenne.org/

1. Introduction and presentation of the model

Long-term evolutionary dynamics of biological populations may be strongly influenced by small populations and local extinction in some areas of the phenotypical trait space. Survival of small populations in very large populations is crucial when evolution proceeds by selective sweeps [24] or for the evolution of antibiotic resistance for bacteria [31, 25]. For example, in bacterial populations involving horizontal transfer, it was shown in [6, 10] that the individual-based long-term dynamics is very sensitive to random survival of very small populations, which may either drive the population to evolutionary suicide or to cyclic dynamics. In such context, the bacterial population is very large making the tracking of small populations very challenging.

From a point of view of mathematical modeling, one wishes to consider large population scalings allowing for survival of much smaller populations. Two approaches emerged: a purely deterministic one, based on partial differential equations (PDE), and a stochastic one, based on birth and death processes (so-called individual-based models in biology). Both approaches describe exponentially small populations sizes and characterize the dynamics of the exponents.

An analytical approach allowing to deal with negligible but non-extinct populations was proposed in [17] and then widely developed (see for instance [32, 3, 26]) for the asymptotic study of parabolic integro-differential selection-mutation models. Let us present it in a setting close to [3]. We consider a population whose individuals are differentiated by a trait $x \in \mathbb{T}$, the torus of dimension 1, identified below with the interval $[0, 1)$. The trait can vary from an individual to the other. The evolution of the population is driven by two effects: mutation of the traits, and selection as the reproductive and survival abilities of an individual depend on its trait x. For an individual of trait $x \in \mathbb{T}$, let us denote by $b(x)$ (resp. $d(x)$ and $p(x)$) the clonal birth rate (resp. the death rate and the birth rate with mutation), and by $G(h)$ the mutation kernel. Assuming that the population density solves the PDE

$$\tag{1.1} \begin{cases} \varepsilon\partial_t u_\varepsilon(t,x) = u_\varepsilon(t,x)\left(b(x)-d(x)\right) \\ \qquad\qquad + \displaystyle\int_{\mathbb{T}} \frac{1}{\varepsilon} G\left((x-y)/\varepsilon\right) p(y) u_\varepsilon(t,y)\mathrm{d}y, & (t,x)\in\mathbb{R}_+\times\mathbb{T}, \\ u_\varepsilon(t,0) = \exp(\beta_0(x)/\varepsilon), & x\in\mathbb{T} \end{cases}$$

in the limit $\varepsilon \to 0$ of small mutations and large time, and applying the Hopf-Cole transformation

$$\tag{1.2} \beta_\varepsilon(t,x) = \varepsilon \log u_\varepsilon(t,x), \quad \text{or} \quad u_\varepsilon(t,x) = \exp(\beta_\varepsilon(t,x)/\varepsilon),$$

it is proved in [3] (in a slightly different setting, considering $x \in \mathbb{R}$ and taking into account a competition term) that β_ε converges to the unique viscosity solution β of the Hamilton-Jacobi equation

$$\tag{1.3} \begin{cases} \dfrac{\partial}{\partial t}\beta(t,x) = b(x) - d(x) + p(x)\displaystyle\int_{\mathbb{R}} G(h) e^{h\partial_x\beta(t,x)} dh, & (t,x)\in\mathbb{R}_+\times\mathbb{T}, \\ \beta(0,x) = \beta_0(x), & x\in\mathbb{T}. \end{cases}$$

Scaling limits of individual-based models on a discrete trait space with rare mutations and large population via a scaling parameter K that will tend to infinity, and allowing to deal with negligible populations and local extinction, were proposed in [18, 8, 14, 15, 7]. These references focus on population sizes of the order of K^{β}, and characterize the asymptotic dynamics of the exponent β. In particular, local extinction is possible when the exponent β hits 0. The fact that the trait space is discrete allows to describe separately the dynamics of each small sub-population. However, this makes the detailed description of the asymptotic dynamics very complicated (see [14, 15]). Note that mutations are assumed individually rare in these references, but they are more frequent than in the scaling limits of adaptive dynamics (see e.g. [29, 16, 11, 13]), where negligible populations either fixate or go extinct fast, due to the fact that the populational mutation rate vanishes here. Scaling limits with non-vanishing populational mutation rates partly solve the criticisms raised by biologists [34] concerning the too slow evolutionary speed in adaptive dynamics, particularly in microorganism populations. Biological criticisms were also raised for the analytical approach [33], because of the so-called *tail problem*: exponentially small populations, which may actually be extinct, can have a strong influence on the future evolutionary dynamics of the population. In particular, evolutionary branching is too fast. Modifications of the Hamilton-Jacobi equation were proposed in [33, 30, 20] to solve this problem, but we believe that an individual-based approach is crucial to provide a more realistic and biologically relevant solution to the tail problem.

The purpose of our work is to provide a stochastic individual-based justification of Hamilton-Jacobi equations. To our knowledge, this is the first proof of this kind in the literature. Note however that there are many examples of spatial branching processes with space-, time- or type-dependent branching rate, for which the exponential growth can be expressed using variational formulas over paths (see e.g. [5, 4, 9, 28, 27]). These can be seen as Hopf-Lax variational formulas of certain Hamilton-Jacobi equations. Note also that the Hopf-Cole transformation (1.2) is reminiscent of large deviations scalings. Our scaling is more of a law of large numbers type. As far as we know, the large deviations interpretation of the Hamilton-Jacobi equation can be done through a Feynman-Kac interpretation of the PDE (1.1) [12]. However, the stochastic process involved in the Feynman-Kac formula does not seem to be directly related to the biological population process, even though some works suggest that it may be related to the ancestral trait process of living individuals [19]. None of the references above use a direct approach from individual-based models to Hamilton-Jacobi equations.

We follow an individual-based approach, assuming a continuous trait space with a vanishing discretization step δ_K, where K is a scaling parameter such that the population is of the order of $K^{\widetilde{\beta}^K(t,x)}$, assuming frequent and small mutations. In the individual-based model, individuals with trait $x \in \mathbb{T}$ give birth to a clone at rate $b(x)$, die at rate $d(x)$ or give birth to a mutant at rate $p(x)$. Mutant traits are drawn according to a discretization of the distribution $\log(K)G(\log(K)\cdot)$. Mutation steps are of the order of $1/\log(K)$ and the discretization step δ_K is assumed much smaller than $1/\log(K)$. In this first work, we focus on the understanding of the relevant scales allowing to capture the limiting Hamilton-Jacobi dynamics. Thus, we consider a simplified model where the birth rate b is assumed larger than the death rate d, making the stochastic process super-critical, and the trait space has no boundary. Generalization is a work in progress.

The proof of our main result makes use of uniform Lipschitz bounds on the finite variation part of $\widetilde{\beta}^K$, obtained using an almost sure maximum principle and careful bounds for the martingale part. The identification of the limit is done by checking that it is almost surely a viscosity solution of (2.7). We describe the model and state our main result in Section 2. The proof is divided into two main steps—proof of tightness and identification of the limit—which are detailed in Sections 3 and 4, respectively.

Acknowledgements. — The authors thank the anonymous reviewers and also Anouar Jeddi for their comments.

2. Model and main result

2.1. The model. — We consider a super-critical stochastic birth-death-mutation model describing an asexual population of individuals characterized by a quantitative phenotypic or genetic trait $x \in \mathbb{T}$. Starting from a finite population whose initial size is parameterized by $K \in \mathbb{N}$, our goal is to recover, in the limit $K \to +\infty$, an evolutionary dynamics described by the Hamilton-Jacobi partial differential equation (1.3). For this, we consider a discretization of the trait space $\mathbb{T}$ with step $\delta_K \to 0$. For the sake of simplicity, we will consider in what follows that $1/\delta_K \in \mathbb{N}$. Then, the population is composed of individuals with traits belonging to the discrete space

$$\mathcal{X}_K := \Big\{ i\delta_K : i \in \{0, 1, \dots, \frac{1}{\delta_K} - 1\} \Big\},$$

embedded with the torus distance: $\forall x, y \in [0,1)$,

$$\begin{aligned} \rho(x,y) &= \min\{|x'-y'|,\ x' = x \bmod 1,\ y' = y \bmod 1\} \\ &= \min\big(|x-y|, 1-|x-y|\big). \end{aligned}$$

It is enough to define $\rho(x,y)$ for $x, y \in \mathbb{T}$ by considering their representative in $[0,1)$.

The number of individuals with trait $i\delta_K$ is described by the stochastic process $(N_i^K(t), t \geqslant 0)$. The total population size at time t is then given by

$$N^K(t) = \sum_{i=0}^{1/\delta_K - 1} N_i^K(t).$$

An individual with trait $x \in \mathcal{X}_K$

– gives birth to a new individual with the same trait x at rate $b(x)$;
– dies at rate $d(x)$;
– gives birth to a mutant individual with trait $y \in \mathcal{X}_K$ at rate

$$p(x)\delta_K \log(K)\, G\big(\overline{(y-x)}\, \log(K)\big), \tag{2.1}$$

where $\overline{z}$ is the unique real number of $[-1/2, 1/2)$ that is equal to z modulo 1, i.e., $\overline{z} = z - \lfloor z + 1/2 \rfloor$, where $\lfloor \cdot \rfloor$ is the integer part function.

Jumps of N^K are thus of $+1$ or -1. In the rest of the paper, the following assumptions are made:

Assumption 2.1

(1) We assume that b, d and p are nonnegative Lipschitz continuous functions defined on $\mathbb{T}$, such that for all $x \in \mathbb{T}$,

$$b(x) > d(x) \text{ and } p(x) > 0.$$

This means that the birth-death process for each trait is super-critical. In the sequel, we denote by $\overline{b}$, $\overline{p}$ and $\overline{d}$ the upper bounds of these functions on $\mathbb{T}$, by $\underline{p} > 0$ the lower bound of p, and by $\|b\|_{\mathrm{Lip}}$, $\|d\|_{\mathrm{Lip}}$ and $\|p\|_{\mathrm{Lip}}$ their Lipschitz norm.

(2) The function G defined on $\mathbb{R}$ is a nonnegative continuous density function (satisfying $\int_{\mathbb{R}} G(y)\,dy = 1$) and has finite exponential moments of any order. Moreover, we assume that there exists $R > 0$ such that G is nonincreasing on $[R, +\infty)$ and nondecreasing on $(-\infty, -R]$.

An example of function G satisfying Assumption 2 is given by the Gaussian kernel $G(h) = \frac{1}{\sqrt{2\pi}\sigma} e^{-h^2/2\sigma^2}$.

(3) There exists a constant $a_1 > 0$ such that, for all $K \in \mathbb{N}$ and all $i \in \{0, 1, \dots, 1/\delta_K - 1\}$,

$$N_i^K(0) \geqslant K^{a_1}. \tag{2.2}$$

(4) There exists $a_2 < a_1$ such that

$$K^{-a_2/4} \ll \delta_K \ll \frac{1}{\log(K)} \quad \text{as} \quad K \longrightarrow +\infty. \tag{2.3}$$

Point (4) above implies that

$$h_K := \delta_K \log(K) \ll 1. \tag{2.4}$$

Thus, the interpretation of (2.1) is that the rate at which x gives birth to a mutant individual is close to $p(x)$. Indeed, for an individual with trait $x_K = i_K \delta_K$ with $i_K = \lfloor x/\delta_K \rfloor$ and $x \in \mathbb{T}$ fixed,

$$\begin{aligned} \lim_{K\to+\infty} p(x_K) \sum_{j=0}^{1/\delta_K - 1} h_K G\left(\overline{(i_K - j)\delta_K}\, \log(K)\right) \\ = \lim_{K\to+\infty} p(x_K) \sum_{\ell=-\lfloor 1/2\delta_K \rfloor}^{1/\delta_K - 1 - \lfloor 1/2\delta_K \rfloor} h_K G\left(h_K \ell\right) = p(x) \int_{\mathbb{R}} G(y)\,dy = p(x), \end{aligned}$$

where we used Assumption 2.1.2 to control the tails of the Riemann sum.

Therefore, the individual mutation rate is order 1 and mutation steps are small: conditionally on being a mutant, the trait y of the offspring has the distribution G scaled by a factor $1/\log(K)$ representing the order of magnitude of the mutation steps. Note also that (2.4) means that the mesh size is much smaller than the mutation step.

Our goal is to study the asymptotic behavior (when K tends to infinity) of the population sizes N_i^K when $N_i^K(0)$ is of the order of K^{α_i} for some $\alpha_i > 0$. Note that in the case where $p(x) = 0$ for all $x \in \mathbb{T}$, the process $N_i^K(t)$ is a super-critical one-dimensional branching process, hence $\mathbb{E}(N_i^K(t)) = \mathbb{E}(N_i^K(0))e^{(b(i\delta_K)-d(i\delta_K))t}$. Therefore, if the initial condition is of order K^{α_i}, then

$$\mathbb{E}[N_i^K(t\log(K))] \sim K^{\alpha_i+(b(i\delta_K)-d(i\delta_K))t}.$$

This suggests to study

$$\beta_i^K(t) = \frac{\log(N_i^K(t\log(K)))}{\log(K)},$$

with the convention that $\beta_i^K(t) = 0$ if $N_i^K(t\log(K)) = 0$. So the sub-population of trait $i\delta_K$ at time $t\log(K)$ has size $N_i^K(t\log(K)) = K^{\beta_i^K(t)}$.

We make the following Lipschitz assumption on the initial condition $\beta_i^K(0)$.

Assumption 2.2. — Assume that there exists a constant $A > 0$ such that

$$\lim_{K\to+\infty} \mathbb{P}\Big(\sup_{i\neq j} \frac{|\beta_i^K(0) - \beta_j^K(0)|}{\rho(i\,\delta_K, j\,\delta_K)} > A\Big) = 0.$$

Notations

(i) We shall use the following Riemann approximation repeatedly in the proofs: for all $L > 0$,

$$\lim_{K\to+\infty} \sum_{j=0}^{1/\delta_K-1} h_K G\left(\overline{(i_K-j)\delta_K}\,\log(K)\right) e^{L\,\log(K)\,\rho(j\delta_K,i_K\delta_K)} = \lim_{K\to+\infty} \sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K G\left(h_K\ell\right) e^{L\,h_K\,|\ell|} = \int_{\mathbb{R}} e^{L|y|}G(y)\,dy,$$

and thus, there exists a constant $\overline{G}(L)$ depending only on $\alpha > 0$ such that

$$\sup_{K\geqslant 1} \sum_{j=0}^{1/\delta_K-1} h_K G\left(\overline{(i_K-j)\delta_K}\,\log(K)\right) e^{L\,\log(K)\,\rho(j\delta_K,i_K\delta_K)} =: \overline{G}(L) < +\infty. \tag{2.5}$$

(ii) In what follows and for any function f on $\{0, 1, \ldots, 1/\delta_K - 1\}$, we will use the notation

$$\Delta_K f_i = \frac{f_{i+1} - f_i}{\delta_K},$$

with the convention that $f_{1/\delta_K} = f_0$.

2.2. The main result - Sketch of the proof. — Since we are interested in the convergence of the quantities β_i^K to a continuous function defined on the trait space $\mathbb{T}$, when $K \to +\infty$ and $\delta_K \to 0$, we introduce the following affine interpolation of the β_i^K's: for all $x \in \mathbb{T}$ and $K \geqslant 1$, let $i \in \{0, 1, \ldots, 1/\delta_K - 1\}$ be such that $x \in [i\delta_K, (i+1)\delta_K)$, and define

$$\widetilde{\beta}_t^K(x) = \beta_i^K(t)\Big(1 - \frac{x}{\delta_K} + i\Big) + \beta_{i+1}^K(t)\Big(\frac{x}{\delta_K} - i\Big), \tag{2.6}$$

with the convention that $\beta^K_{1/\delta_K}(t) = \beta^K_0(t)$. For $T > 0$, the sequence of processes $(\widetilde{\beta}^K_t, t \in [0,T])_K$ belongs to $\mathbb{D}([0,T], \mathcal{C}(\mathbb{T},\mathbb{R}))$, where $\mathcal{C}(\mathbb{T},\mathbb{R})$ is endowed with the topology of uniform convergence and $\mathbb{D}([0,T], \mathcal{C}(\mathbb{T},\mathbb{R}))$ is the Skorokhod space of càdlàg paths with the associated Skorokhod topology.

Let us state our main theorem.

Theorem 2.3. — *Let $T > 0$. Under the Assumptions* 2.1 *and* 2.2*, and assuming that $\widetilde{\beta}^K_0(\cdot)$ converges in probability for the topology of uniform convergence on $\mathcal{C}(\mathbb{T},\mathbb{R})$ to a deterministic function $\beta_0(\cdot) \in \mathcal{C}(\mathbb{T},\mathbb{R})$, the sequence $(\widetilde{\beta}^K)_K$ converges in probability in $\mathbb{D}([0,T], \mathcal{C}(\mathbb{T},\mathbb{R}))$ to the unique Lipschitz viscosity solution of the Hamilton-Jacobi equation*

$$\begin{cases} \dfrac{\partial}{\partial t}\beta(t,x) = b(x) - d(x) + p(x)\displaystyle\int_{\mathbb{R}} G(h) e^{h\partial_x \beta(t,x)} dh, & (t,x) \in (0,T] \times \mathbb{T}, \\ \beta(0,x) = \beta_0(x), & x \in \mathbb{T}. \end{cases} \tag{2.7}$$

The proof of Theorem 2.3 will be classically obtained in two steps: tightness and identification of the limiting values. Therefore we will prove the two next results, respectively in Sections 3 and 4.

Theorem 2.4. — *The distributions of the processes $(\widetilde{\beta}^K_t, t \in [0,T])_K$ form a C-tight sequence in $\mathcal{P}(\mathbb{D}([0,T], \mathcal{C}(\mathbb{T},\mathbb{R})))$, the set of probability measures on $\mathbb{D}([0,T], \mathcal{C}(\mathbb{T},\mathbb{R}))$. In addition, for all $T > 0$, there exists a constant L such that, for any β distributed as a limiting value of the laws of $(\widetilde{\beta}^K_t, t \in [0,T])_K$, we have almost surely*

$$\sup_{t\in[0,T]} \sup_{x,y\in\mathbb{T} \ s.t. \ x\neq y} \frac{|\beta(t,x) - \beta(t,y)|}{\rho(x,y)} \leqslant L. \tag{2.8}$$

Theorem 2.5. — *The limiting values β of $(\widetilde{\beta}^K_t, t \in [0,T])_K$ are characterized as the unique Lipschitz viscosity solution of the Hamilton-Jacobi equation* (2.7)*.*

The proofs of these two results require to control the increments of the functions $\widetilde{\beta}^K_t(.)$. These functions can also be written as

$$\widetilde{\beta}^K_t(x) = (x - i\delta_K)\Delta_K \beta^K_i(t) + \beta^K_i(t).$$

From this expression, we observe that two technical steps are required: to estimate uniformly $\beta^K_i(t)$ and to control uniformly $\Delta_K \beta^K_i(t) = (\beta^K_{i+1}(t) - \beta^K_i(t))/\delta_K$ (the second estimate being the harder part, it is a major difficulty and constitutes the technical interest of the paper). Such estimates are also obtained in the deterministic derivation of Hamilton-Jacobi equations of type (2.7) from parabolic integro-differential equations using the maximum principle and the Bernstein method which consists in applying again the maximum principle to the equation satisfied by the increments (see [3]). Here, since we have stochastic processes we cannot apply the Bernstein method directly. Using the Doob-Meyer decomposition, the stochastic processes can be separated into a finite variation part and a martingale part. We show indeed that the martingale part remains small with our rescaling and we apply the maximum principle almost surely on the finite variation part.

Let us detail now the semi-martingale Doob-Meyer decomposition of the processes β_i^K, $i \in \{0, \dots, 1/\delta_K - 1\}$. Using standard arguments [1],

$$N_i^K(t) - N_i^K(0) - \int_0^t (b(i\delta_K) - d(i\delta_K))N_i^K(s)ds - \sum_{\ell=-\lfloor 1/2\delta_K \rfloor}^{1/\delta_K - 1 - \lfloor 1/2\delta_K \rfloor} \int_0^t h_K p((i+\ell)\delta_K) G(h_K \ell) N_{i+\ell}^K(s)ds \tag{2.9}$$

is a square integrable martingale with quadratic variation

$$\int_0^t (b(i\delta_K) + d(i\delta_K))N_i^K(s)ds + \sum_{\ell=-\lfloor 1/2\delta_K \rfloor}^{1/\delta_K - 1 - \lfloor 1/2\delta_K \rfloor} \int_0^t h_K p((i+\ell)\delta_K) G(h_K \ell) N_{i+\ell}^K(s)ds.$$

It then follows from Itô's formula for jump processes that

$$\beta_i^K(t) = M_i^K(t) + A_i^K(t) \tag{2.10}$$

with

$$A_i^K(t) = \beta_i^K(0) + \frac{1}{\log(K)} \int_0^{t\log(K)} \Big(b(i\delta_K) N_i^K(s) \log\big(1 + 1/N_i^K(s)\big) \tag{2.11}$$

$$+ d(i\delta_K) N_i^K(s) \log\big(1 - 1/N_i^K(s)\big)\Big) ds \tag{2.12}$$

$$+ \frac{1}{\log(K)} \sum_{\ell=-\lfloor 1/2\delta_K \rfloor}^{1/\delta_K - 1 - \lfloor 1/2\delta_K \rfloor} h_K\, p((i+\ell)\delta_K) G(h_K \ell) \tag{2.13}$$

$$\times \int_0^{t\log(K)} N_{i+\ell}^K(s) \log\big(1 + 1/N_i^K(s)\big)\, ds, \tag{2.14}$$

with the conventions that, when the index $j \notin \{0, \dots, 1/\delta_K - 1\}$,

$$\begin{cases} N_j^K = N_{j - \lfloor j\delta_K \rfloor/\delta_K}^K, \\ p(j\delta_K) = p((j - \lfloor j\delta_K \rfloor/\delta_K)\delta_K), \end{cases} \text{when } j \geqslant 1/\delta_K, \qquad \begin{cases} N_j^K = N_{j + \lceil |j|\delta_K \rceil/\delta_K}^K, \\ p(j\delta_K) = p((j + \lceil |j|\delta_K \rceil/\delta_K)\delta_K), \end{cases} \text{when } j < 0. \tag{2.15}$$

The process M_i^K is a local martingale with predictable quadratic variation

$$\langle M_i^K \rangle_t = \frac{1}{\log^2(K)} \int_0^{t\log(K)} \Big(b(i\delta_K) N_i^K(s) \log^2\big(1 + 1/N_i^K(s)\big) \tag{2.16}$$

$$+ d(i\delta_K) N_i^K(s) \log^2\big(1 - 1/N_i^K(s)\big)\Big) ds \tag{2.17}$$

$$+ \frac{1}{\log^2(K)} \sum_{\ell=-\lfloor 1/2\delta_K \rfloor}^{1/\delta_K - 1 - \lfloor 1/2\delta_K \rfloor} h_K p((i+\ell)\delta_K) G(h_K \ell) \tag{2.18}$$

$$\times \int_0^{t\log(K)} N_{i+\ell}^K(s) \log^2\big(1 + 1/N_i^K(s)\big)\, ds. \tag{2.19}$$

In Section 3, we will prove technical uniform estimates on the martingale part $\Delta_K M_i^K(t)$ and the finite variation part $\Delta_K A_i^K(t)$ of the processes $\Delta_K \beta_i^K(t)$. In that aim, let us introduce sequences of stopping times playing an important role in the proofs.

Let $a \in (a_2, a_1)$ be fixed during the rest of the proof, a_1 being defined in (2.2) and a_2 in (2.3). For any K, we define

$$\tau_K' = \inf\Big\{t \geqslant 0 : \exists i \in \{0, 1, \dots, 1/\delta_K - 1\}, N_i^K(t\log(K)) < K^a\Big\}. \tag{2.20}$$

Note that $\beta_i^K(t) > 0$ for all $t \leqslant \tau_K'$.

For all $L > 0$, we also define

$$\tau_K(L) = \inf\Big\{t \geqslant 0 : \exists i \in \{0, 1, \dots, 1/\delta_K - 1\}, |\beta_{i+1}^K(t) - \beta_i^K(t)| > L\delta_K\Big\}, \tag{2.21}$$

with the usual convention that $\beta_{1/\delta_K}^K = \beta_0^K$. It is easy to check that

$$\tau_K(L) = \inf\Big\{t \geqslant 0 : \exists i, j \in \{0, 1, \dots, 1/\delta_K - 1\}, |\beta_i^K(t) - \beta_j^K(t)| > L\rho(i\delta_K, j\delta_K)\Big\}$$

and also

$$\tau_K(L) = \inf\Big\{t \geqslant 0 : \exists x, y \in \mathbb{T}, |\widetilde{\beta}_t^K(x) - \widetilde{\beta}_t^K(y)| > L\ \rho(x, y)\Big\}. \tag{2.22}$$

We will study the processes until the stopping time

$$\theta_K(L) = \tau_K(L) \wedge \tau_K'. \tag{2.23}$$

Before the stopping time $\theta_K(L)$, the functions $\widetilde{\beta}_t^K$ are Lipschitz and the population size of each trait is bounded from below by K^a, by definition. For each L fixed, we will provide uniform estimates on the martingale parts $M_i^K(t)$ and $\Delta_K M_i^K(t)$ and the finite variation parts $A_i^K(t)$ and $\Delta_K A_i^K(t)$ of the processes $\beta_i^K(t)$ and $\Delta_K \beta_i^K(t)$, before the stopping time $\theta_K(L)$. This will allow us to prove that for all T, $\theta_K(L)$ is larger than T for L large enough (see Proposition 3.6).

3. Proof of the Theorem 2.4

We will use the criterion of Theorem 3.1 in Jakubowski [22]. To prove the tightness of the sequence $(\widetilde{\beta}^K)_{K\geqslant 1}$ in $\mathbb{D}([0,T], \mathcal{C}(\mathbb{T}, \mathbb{R}))$, it is sufficient to prove:

(i) For each $\varepsilon > 0$, there exists a compact set $C_\varepsilon \subset \mathcal{C}(\mathbb{T}, \mathbb{R})$ such that

$$\forall K,\ \mathbb{P}\left(\widetilde{\beta}^K \in \mathbb{D}([0,T], C_\varepsilon)\right) > 1 - \varepsilon.$$

(ii) For each $f \in \mathcal{C}(\mathbb{T}, \mathbb{R})$, the sequence of laws of real-valued processes

$$X_f^K(\cdot) = \int_{\mathbb{T}} \widetilde{\beta}^K(., x) f(x) dx \tag{3.1}$$

is tight.

Point (i) is the hard part of the proof. Using Ascoli's characterization of compact subsets of $\mathcal{C}(\mathbb{T}, \mathbb{R})$, we need to obtain estimates related to equi-boundedness and to equi-continuity for the processes $\widetilde{\beta}_t^K(.)$. The proof relies on Lipschitz estimates (in x) of the functions $\widetilde{\beta}_t^K$. In Section 3.1, we show that the martingale part of $\beta_i^K(t\wedge\theta_K(L))$

remains small with our rescaling and in Section 3.2, we apply the maximum principle almost surely on the finite variation part of $\beta_i^K(t \wedge \theta_K(L))$. This allows us to prove that $\theta_K(L)$ is large enough in Section 3.3. The proof of the tightness is ended in Section 3.4.

3.1. Control of the martingale part. — Our first estimate will be useful to prove the tightness of the laws of $\widetilde{\beta}^K$. In the sequel, C denotes a constant not depending on any parameter and that may change from line to line.

The following estimate will be used repeatedly: by (2.22), for all $t \leqslant \tau_K(L)$ and all $i, j \leqslant 1/\delta_K - 1$,

$$\frac{N_j^K(t\log(K))}{N_i^K(t\log(K))} = \exp(\log(K)(\beta_j(t) - \beta_i(t)) \leqslant e^{L\,\rho(j\delta_K, i\delta_K)\,\log(K)}. \tag{3.2}$$

Lemma 3.1. — *For all $T > 0$, $L > 0$, $K > 0$ and $a \in (a_2, a_1)$, there exists a constant C independent of K, T, L and i such that, almost surely, for all $t \leqslant T$ and all $i \in \{0, \dots, 1/\delta_K - 1\}$,*

$$\langle M_i^K \rangle_{t\wedge\theta_K(L)} \leqslant \Big(\frac{C\overline{G}(L)}{K^a \log(K)}\Big) t. \tag{3.3}$$

Further, for all $A > 0$,

$$\mathbb{P}\Big(\sup_{t\leqslant T\wedge\theta_K(L)} \sup_i |M_i^K(t)| \geqslant A\Big) \leqslant \frac{C\overline{G}(L)T}{A^2\delta_K K^a \log(K)}. \tag{3.4}$$

Proof. — It follows from (2.16) that

$$\begin{aligned}\langle M_i^K \rangle_{t\wedge\theta_K(L)} \leqslant \frac{C(\overline{b} + \overline{d})}{\log^2(K)} \int_0^{(t\wedge\theta_K(L))\log(K)} \frac{ds}{N_i^K(s)} \\ + \frac{\overline{p}}{\log^2(K)} \int_0^{(t\wedge\theta_K(L))\log(K)} \sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K - 1 - \lfloor 1/2\delta_K\rfloor} h_K G(h_K \ell) \frac{N_{i+\ell}^K(s)}{(N_i^K(s))^2}\, ds.\end{aligned}$$

Therefore, using (2.20) and (3.2), (3.3) follows. Then

$$\begin{aligned}\mathbb{P}\Big(\sup_{t\leqslant T\wedge\theta_K(L)} \sup_i |M_i^K(t)| \geqslant A\Big) &\leqslant \sum_{i=0}^{1/\delta_K - 1} \mathbb{P}\Big(\sup_{t\leqslant T\wedge\theta_K(L)} |M_i^K(t)| \geqslant A\Big) \\ &\leqslant \frac{1}{A^2} \sum_{i=0}^{1/\delta_K - 1} \mathbb{E}\Big(\sup_{t\leqslant T\wedge\theta_K(L)} |M_i^K(t)|^2\Big),\end{aligned}$$

so we obtain (3.4) using Doob's inequality and (3.3). □

The second lemma gives controls on the martingale increments.

Lemma 3.2. — *For all $T > 0$, $K > 0$ and $a \in (a_2, a_1)$, let us set $\varepsilon_K = \delta_K^{-1}(K^a \log(K))^{-1/4}$. For $L > 0$, we define the event*

$$\Omega_K(L) = \Big\{\sup_{\substack{0\leqslant i\leqslant 1/\delta_K - 1\\ t\leqslant T\wedge\theta_K(L)}} |\Delta_K M_i^K(t)| \leqslant \varepsilon_K\Big\}. \tag{3.5}$$

Then there exists a constant $C>0$ such that

$$\mathbb{P}(\Omega_K^c(L)) \leqslant \frac{C\sqrt{\overline{G}(L)T}}{\delta_K (K^a\log(K))^{1/4}}. \tag{3.6}$$

In addition,

$$\mathbb{P}\Big(\sup_{\substack{0\leqslant i,j\leqslant 1/\delta_K-1\\ t\leqslant T\wedge\theta_K(L)}} \big|M_i^K(t)-M_j^K(t)\big| > \varepsilon_K\Big) \leqslant C\sqrt{\overline{G}(L)T}\varepsilon_K. \tag{3.7}$$

Proof. — We first prove that, for all $\varepsilon>0$ and any $i\in\{0,\dots,1/\delta_K-1\}$, for some constant C independent of t, K, ε and L,

$$\mathbb{P}\Big(\sup_{s\leqslant t\wedge\theta_K(L)} |\Delta_K M_i^K(s)| > \varepsilon\Big) \leqslant \frac{C}{\varepsilon}\sqrt{\frac{\overline{G}(L)t}{\delta_K^2 K^a\log(K)}}, \tag{3.8}$$

where the constant $\overline{G}(L)$ is defined in (2.5).

By the submartingale maximal lemma, we have that

$$\mathbb{P}\Big(\sup_{s\leqslant t} |\Delta_K M_i^K(s)| > \varepsilon\Big) \leqslant \frac{1}{\varepsilon}\,\mathbb{E}(|\Delta_K M_i^K(t)|) \leqslant \frac{1}{\varepsilon}\,\mathbb{E}\big(|\Delta_K M_i^K(t)|^2\big)^{1/2}.$$

Now, Lemma 3.1 yields

$$\langle \Delta_K M_i^K\rangle_{t\wedge\theta_K} \leqslant \frac{2}{\delta_K^2}\left(\langle M_{i+1}^K\rangle_{t\wedge\theta_K} + \langle M_i^K\rangle_{t\wedge\theta_K}\right) \leqslant \frac{4C\overline{G}(L)t}{\delta_K^2 K^a\log(K)}.$$

Hence (3.8) is proved.

Now, we choose $\varepsilon=\varepsilon_K$ and obtain

$$\begin{aligned}\mathbb{P}(\Omega_K^c(L)) &\leqslant \sum_{0\leqslant i\leqslant 1/\delta_K-1} \mathbb{P}\Big(\sup_{t\leqslant T\wedge\theta_K(L)} |\Delta_K M_i^K(t)| > \varepsilon_K\Big)\\ &\leqslant \frac{C\sqrt{\overline{G}(L)T}}{\delta_K\varepsilon_K\delta_K\sqrt{K^a\log(K)}}.\end{aligned}$$

This ends the proof of (3.6).

To complete the proof of Lemma 3.2, it is sufficient to notice that, since $|j-i|\delta_K\leqslant 1$ for all $0\leqslant i,j\leqslant 1/\delta_K-1$, we have that

$$\Omega_K(L)\subset\Big\{\sup_{\substack{0\leqslant i,j\leqslant 1/\delta_K-1\\ t\leqslant T\wedge\theta_K(L)}} \big|M_i^K(t)-M_j^K(t)\big| \leqslant \varepsilon_K\Big\}. \qquad \square$$

Note that, by (2.3), $\delta_K^4 K^a\log(K)$ tends to infinity as K goes to infinity, so $\mathbb{P}(\Omega_K(L))$ tends to 1 and the probability of $\Omega_K(L)$ converges to 1. From now on, we will work on the probability subspace $\Omega_K(L)$.

3.2. Control of the finite variation part. — Let us now focus on the finite variation part A_i^K. We will prove that

Proposition 3.3. — *Let $T > 0$. Then, there exists a constant C_1 such that for K large enough, for all $t \leqslant T$ and all $i \in \{0, \dots, 1/\delta_K - 1\}$ the following inequality holds almost surely on $\Omega_K(L)$:*
$$|A_i^K(t \wedge \theta_K(L))| \leqslant \max_{0 \leqslant j \leqslant 1/\delta_K - 1} \beta_j^K(0) + C_1 t.$$

Proof. — We first provide the proof of the upper bound on $A_i^K(t \wedge \theta_K(L))$. For simplicity we will omit the dependency in L of $\theta_K(L)$ when there is no ambiguity.

Let t and s be less than T such that $s < t$. Using that $\log(1 + x) \leqslant x$ and (3.2), and neglecting the non positive death term, we have
$$\begin{aligned}
&A_i^K(t \wedge \theta_K) - A_i^K(s \wedge \theta_K)\\
&\quad = \frac{1}{\log(K)} \int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)} \Big(b(i\delta_K) N_i^K(u) \log\big(1 + 1/N_i^K(u)\big)\\
&\qquad\qquad + d(i\delta_K) N_i^K(u) \log\big(1 - 1/N_i^K(u)\big) \Big) du\\
&\quad + \frac{1}{\log(K)} \sum_{\ell = -\lfloor 1/2\delta_K \rfloor}^{1/\delta_K - 1 - \lfloor 1/2\delta_K \rfloor} h_K p((i+\ell)\delta_K) G(h_K \ell)\\
&\qquad\qquad \times \int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)} N_{i+\ell}^K(u) \log\big(1 + 1/N_i^K(u)\big)\, du\\
&\quad \leqslant C\bar{b}(t-s)\\
&\quad + \frac{1}{\log(K)} \sum_{\ell = -\lfloor 1/2\delta_K \rfloor}^{1/\delta_K - 1 - \lfloor 1/2\delta_K \rfloor} h_K p((i+\ell)\delta_K) G(h_K \ell) \int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)} \frac{N_{i+\ell}^K(u)}{N_i^K(u)} du\\
&\quad \leqslant C\bar{b}(t-s)\\
&\quad + \sum_{\ell = -\lfloor 1/2\delta_K \rfloor}^{1/\delta_K - 1 - \lfloor 1/2\delta_K \rfloor} h_K p((i+\ell)\delta_K) G(h_K \ell) \int_{s\wedge\theta_K}^{t\wedge\theta_K} \exp\big(\log(K)(\beta_{i+\ell}^K(u) - \beta_i^K(u))\big) du.
\end{aligned}$$
Recall that on $\Omega_K(L)$,
$$\beta_j^K(u) - \beta_i^K(u) = A_j^K(u) - A_i^K(u) + M_j^K(u) - M_i^K(u) \leqslant A_j^K(u) - A_i^K(u) + \varepsilon_K.$$
Thus we obtain that, on the event $\Omega_K(L)$,
$$\begin{aligned}
A_i^K(t \wedge \theta_K) - A_i^K(s \wedge \theta_K) \leqslant C\bar{b}(t-s) + \sum_{\ell = -\lfloor 1/2\delta_K \rfloor}^{1/\delta_K - 1 - \lfloor 1/2\delta_K \rfloor} h_K p((i+\ell)\delta_K) G(h_K \ell)\\
\int_{s\wedge\theta_K}^{t\wedge\theta_K} e^{\varepsilon_K \log(K)} \exp\big(\log(K)(A_{i+\ell}^K(u) - A_i^K(u))\big) du.
\end{aligned}$$
We deduce that, almost surely on $\Omega_K(L)$ and for all $t \leqslant \theta_K(L)$,
$$\begin{aligned}
\frac{dA_i^K(t)}{dt} \leqslant C\bar{b} + \sum_{\ell = -\lfloor 1/2\delta_K \rfloor}^{1/\delta_K - 1 - \lfloor 1/2\delta_K \rfloor} h_K p((i+\ell)\delta_K) G(h_K \ell)\\
\times e^{\varepsilon_K \log(K)} \exp\big(\log(K)(A_{i+\ell}^K(t) - A_i^K(t))\big).
\end{aligned}$$

Defining $\mathbf{A}_i^K(t) = A_i^K(t) - C\bar{b}t - 2\overline{p}t$, we deduce that for any $t \leqslant \theta_K(L)$,

$$\begin{aligned}(3.9)\qquad \frac{d\mathbf{A}_i^K(t)}{dt} \leqslant \overline{p}e^{\varepsilon_K \log(K)} \sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} & h_K G(h_K\ell) \\ & \times \exp(\log(K)(\mathbf{A}_{i+\ell}^K(t) - \mathbf{A}_i^K(t))) - 2\overline{p}.\end{aligned}$$

Let us introduce

$$(i_K, t_K) = (i_K(\omega), t_K(\omega)) = \operatorname{argmax}_{i\in\{0,\dots,1/\delta_K-1\}, t\in[0,\theta_K(\omega)\wedge T]} \mathbf{A}_i^K(t).$$

We can prove that $t_K = 0$. Indeed, if conversely we assume that $t_K > 0$, then the right term of (3.9), for K large enough, is negative for $i = i_K$ and $t = t_K$ and then the left term is negative, contradicting the fact that $\mathbf{A}_{i_K}^K(t)$ is maximal for $t = t_K$. Hence, we have proved that for K large enough, almost surely on the event Ω_K, for all $t \leqslant \theta_K(L) \wedge T$ and for all $i \in \{0, ..., 1/\delta_K - 1\}$,

$$\begin{aligned} A_i^K(t) = \mathbf{A}_i^K(t) + C\bar{b}t + 2\overline{p}t &\leqslant \max_{0\leqslant j\leqslant 1/\delta_K} \mathbf{A}_j^K(0) + C\bar{b}t + 2\overline{p}t \\ &= \max_{0\leqslant j\leqslant 1/\delta_K} \beta_j^K(0) + C\bar{b}t + 2\overline{p}t. \end{aligned}$$

For the lower bound, we observe that, for $t \leqslant \theta_K(L) \wedge T$, $N_i^K(t) \geqslant K^a$ and hence, for K large enough,

$$\log\big(1 + 1/N_i^K(t)\big) \geqslant 0 \quad \text{and} \quad \log\big(1 - 1/N_i^K(t)\big) \geqslant -\frac{2}{N_i^K(t)}.$$

The proof then follows from the same argument as above. ☐

The last result has a consequence that will be useful to prove the tightness of $\widetilde{\beta}^K$ in Section 3.4.

Corollary 3.4. — *For all $T > 0$, there exists $C(T)$ such that,*

$$\lim_{K\to+\infty} \mathbb{P}\Big(\sup_{0\leqslant i\leqslant 1/\delta_K-1}\ \sup_{t\in[0,T\wedge\theta_K(L)]} \beta_i^K(t) \geqslant C(T)\Big) = 0.$$

Proof. — We use the semi-martingale decomposition (2.10) of β_i^K, the result of Proposition 3.3, (3.4) with $A = 1$ and Lemma 3.2, to deduce that, for all $t \leqslant T$ and K large enough and $i \in \{0, 1, \dots, 1/\delta_K - 1\}$,

$$|\beta_i^K(t\wedge\theta_K)| \leqslant \sup_{0\leqslant j\leqslant 1/\delta_K-1} |\beta_j^K(0)| + C_1T + 1$$

with probability at least

$$1 - \frac{C\overline{G}(L)T}{\delta_K K^a \log(K)} - \frac{C\sqrt{\overline{G}(L)T}}{\delta_K(K^a\log(K))^{1/4}}.$$

Since $\widetilde{\beta}_i^K(0)$ converges in probability to β_0, $\mathbb{P}(\sup_i \beta_i^K(0) \geqslant \|\beta_0\|_\infty + 1)$ converges to 0 when K goes to $+\infty$. Hence the result follows with $C(T) = \|\beta_0\|_\infty + C_1T + 2$. ☐

3.3. Estimates on $\theta_K(L)$

Lemma 3.5. — *Under Assumption* 2.1(1) *and* 2.1(3), *for all* $T>0$,

$$\lim_{K\to+\infty}\mathbb{P}(\tau'_K\geqslant T)=1.$$

Proof of Lemma 3.5. — By (2.9), neglecting incoming mutations, it is easy to prove using standard coupling arguments that for each i, the process $(N_i^K(t))_t$ is pathwisely bounded below by a branching process $(Z_i^K(t))_t$ with birth rate $b(i\delta_K)$, death rate $d(i\delta_K)$ and initial condition $K^{a+\varepsilon}$ for $\varepsilon=a_1-a>0$. In addition, the processes $(Z_i^K(t))_t$ for $0\leqslant i\leqslant 1/\delta_K-1$ are independent. Let us define

$$\theta''_K=\inf\Big\{t\geqslant 0:\exists i\in\{0,1,\dots,1/\delta_K-1\},\,Z_i^K(t\log(K))<K^a\Big\}.$$

In order to prove that $\lim_{K\to\infty}\mathbb{P}(\tau'_K>T)=1$, it is enough to prove that

$$\lim_{K\to\infty}\mathbb{P}(\theta''_K=+\infty)=1.$$

We have

$$\begin{aligned}\mathbb{P}(\theta''_K=+\infty)&=\mathbb{P}\Big(\forall i\in\{0,1,\dots,1/\delta_K-1\},\ \forall t\geqslant 0,\ Z_i^K(t\log(K))>K^a\Big)\\&=\prod_{i=0}^{1/\delta_K-1}\mathbb{P}\Big(\inf_{t\geqslant 0}Z_i^K(t\log(K))>K^a\Big).\end{aligned}$$

Fix $i\in\{0,\dots,1/\delta_K-1\}$. It is usual to prove (by time change) that the probability $\mathbb{P}\big(\inf_{t\geqslant 0}Z_i^K(t\log(K))>K^a\big)$ is equal to the probability that a random walk

$$M\Big(\frac{b(i\delta_K)}{b(i\delta_K)+d(i\delta_K)},\frac{d(i\delta_K)}{b(i\delta_K)+d(i\delta_K)}\Big)$$

on $\mathbb{Z}_+$ (adding $+1$ with probability $b(i\delta_K)/(b(i\delta_K)+d(i\delta_K))$ and -1 with probability $d(i\delta_K)/(b(i\delta_K)+d(i\delta_K)))$ with initial value $K^{a+\varepsilon}$ never attains K^a. This quantity is well known and equal to

$$1-\Big(\frac{d(i\delta_K)}{b(i\delta_K)}\Big)^{K^{a+\varepsilon}-K^a}.$$

Since $\alpha=\max_{x\in\mathbb{T}}d(x)/b(x)<1$, it follows from (2.3) that

$$\begin{aligned}\mathbb{P}(\theta''_K=+\infty)&\geqslant\exp\Big(\frac{1}{\delta_K}\log\big(1-\alpha^{K^{a+\varepsilon}-K^a}\big)\Big)\\&\sim\exp\Big(-\frac{1}{\delta_K}\alpha^{K^{a+\varepsilon}-K^a}\Big)\gg 1-K^{a_2/4}\alpha^{K^{a+\varepsilon}-K^a},\end{aligned}$$

which tends to 1 when K tends to infinity. □

Proposition 3.6. — *Under the Assumptions* 2.1 *and* 2.2, *for all* $T>0$, *there exists* L_T *in the definition* (2.23) *of* $\theta_K(L)$ *such that*

$$\lim_{K\to+\infty}\mathbb{P}(\theta_K(L_T)>T)=1.$$

Proof of Proposition 3.6. — In view of Lemma 3.5 it is enough to prove that there exists L_T such that $\lim_{K\to+\infty} P(\tau_K(L_T)>T)=1$. For $i\in\{0,\dots,1/\delta_K-1\}$, let us consider the increments

$$\begin{aligned}\Delta_K\beta_i^K(t\wedge\theta_K)&=\frac{\beta_{i+1}^K(t\wedge\theta_K)-\beta_i^K(t\wedge\theta_K)}{\delta_K}\\&=\Delta_K M_i^K(t\wedge\theta_K)+\Delta_K A_i^K(t\wedge\theta_K).\end{aligned}$$

We also introduce

$$g_i^K(t)=\Delta_K A_i^K(t\wedge\theta_K)+\frac{\|p\|_{\mathrm{Lip}}}{\underline{p}}A_{i+1}^K(t\wedge\theta_K). \tag{3.10}$$

To prove the result we will show that we can control $\Delta_K\beta_i^K(t\wedge\theta_K)$. To this end, we will first control $g_i^K(t)$ using an almost sure maximum principle. We will then use the fact that $\Delta_K M_i^K(t)$ is small in $\Omega_K(L)$ to obtain a control on $\Delta_K\beta_i^K(t\wedge\theta_K)$. We provide the proof in several steps.

Step 1. — As a first step we prove that for all $t\leqslant\theta_K(L)\wedge T$,

$$\frac{dg_i^K(t)}{dt}\leqslant C(K,L)+\overline{p}\log(K)\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor}h_KG(h_K\ell)\left[g_{\ell+i}^K(t)-g_i^K(t)\right]^+\,e^{h_KL|\ell|}.$$

Using (2.11), we obtain for K large enough and $0\leqslant s<t\leqslant T$,

$$\begin{aligned}
g_i^K(t)-g_i^K(s)\qquad\qquad&\\
=\frac{1}{h_K}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}&\Big[b((i+1)\delta_K)N_{i+1}^K(u)\log\big(1+1/N_{i+1}^K(u)\big)\\
&\qquad-b(i\delta_K)N_i^K(u)\log\big(1+1/N_i^K(u)\big)\Big]du\\
+\frac{1}{h_K}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}&\Big[d((i+1)\delta_K)N_{i+1}^K(u)\log\big(1-1/N_{i+1}^K(u)\big)\\
&\qquad-d(i\delta_K)N_i^K(u)\log\big(1-1/N_i^K(u)\big)\Big]du\\
+\frac{1}{h_K}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}&\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor}h_KG(h_K\ell)\\
&\Big[p((i+1+\ell)\delta_K)N_{i+1+\ell}^K(u)\log\big(1+1/N_{i+1}^K(u)\big)\\
&\qquad-p((i+\ell)\delta_K)N_{i+\ell}^K(u)\log\big(1+1/N_i^K(u)\big)\Big]du\\
+\frac{\|p\|_{\mathrm{Lip}}}{\underline{p}\log(K)}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}&\big[b((i+1)\delta_K)N_{i+1}^K(u)\log\big(1+1/N_{i+1}^K(u)\big)\\
&\qquad+d((i+1)\delta_K)N_{i+1}^K(u)\log\big(1-1/N_{i+1}^K(u)\big)\big]\,du\\
+\frac{\|p\|_{\mathrm{Lip}}}{\underline{p}\log(K)}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}&\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor}h_KG(h_K\ell)\;p((i+1+\ell)\delta_K)\\
&\qquad\times N_{i+1+\ell}^K(u)\log\big(1+1/N_{i+1}^K(u)\big)\,du
\end{aligned}\tag{3.11}$$

$$
\begin{aligned}
&\leqslant \frac{C(\overline{b}+\overline{d})}{h_K}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}\Big[\frac{1}{N^K_{i+1}(u)}+\frac{1}{N^K_i(u)}\Big]du\\
&\qquad\qquad\qquad + (C_1(\overline{b}+\overline{d})+\|b\|_{\mathrm{Lip}}+\|d\|_{\mathrm{Lip}})(t\wedge\theta_K-s\wedge\theta_K)\\
&\quad+\frac{1}{h_K}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)\\
&\qquad\times\big[N^K_{\ell+i+1}(u)\log\big(1+1/N^K_{i+1}(u)\big)-N^K_{\ell+i}(u)\log\big(1+1/N^K_i(u)\big)\big]\,du\\
&\quad+\frac{3\|p\|_{\mathrm{Lip}}}{\underline{p}\log(K)}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)\\
&\qquad\qquad\qquad\qquad N^K_{\ell+i+1}(u)\log\big(1+1/N^K_{i+1}(u)\big)\,du,
\end{aligned}
$$

where we used to prove the last inequality that, for all x such that $|x|\leqslant 1/2$, we have

$$\Big|\frac{1}{x}\log(1+x)\Big|\leqslant C, \tag{3.12}$$

$$\Big|\frac{1}{x}\log(1+x)-\frac{1}{y}\log(1+y)\Big|\leqslant C(|x|+|y|), \tag{3.13}$$

and the fact that (recalling the convention (2.15) and that p is periodic)

$$
\begin{aligned}
|p((\ell+i+1)\delta_K)-p((\ell+i)\delta_K)| &= p((\ell+i)\delta_K)\frac{|p((\ell+i+1)\delta_K)-p((\ell+i)\delta_K)|}{p((\ell+i)\delta_K)}\\
&\leqslant \frac{\|p\|_{\mathrm{Lip}}\delta_K}{\underline{p}}p((\ell+i)\delta_K).
\end{aligned}
$$

Note also that, to obtain the last inequality (3.11), we have taken K large enough such that $\|p\|_{\mathrm{Lip}}\delta_K/\underline{p}\leqslant 1$, so that

$$p((\ell+i+1)\delta_K)\leqslant 2p((\ell+i)\delta_K).$$

Next, notice that for all x',y',x,y such that $|x|,|y|\leqslant 1/2$,

$$\frac{1}{y'}\log(1+y)-\frac{1}{x'}\log(1+x)\leqslant\frac{y}{y'}-\frac{x}{x'}+C\Big(\frac{y^2}{y'}+\frac{x^2}{x'}\Big).$$

Using this inequality and (3.2), we have

$$
\begin{aligned}
&\frac{1}{h_K}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)\\
&\qquad\times\big[N^K_{\ell+i+1}(u)\log\big(1+1/N^K_{i+1}(u)\big)-N^K_{\ell+i}(u)\log\big(1+1/N^K_i(u)\big)\big]\,du\\
&\leqslant\frac{1}{h_K}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)\\
&\qquad\qquad\qquad\qquad\times\left[\frac{N^K_{\ell+i+1}(u)}{N^K_{i+1}(u)}-\frac{N^K_{\ell+i}(u)}{N^K_i(u)}\right]du\\
&\quad+\frac{C\overline{p}}{h_K}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K G(h_K\ell)e^{Lh_K|\ell|}\\
&\qquad\qquad\qquad\qquad\times\left[\frac{1}{N^K_{i+1}(u)}+\frac{1}{N^K_i(u)}\right]\,du.
\end{aligned}
$$

Therefore, using (2.5) and the definition of τ'_K, we have proved that

$$g_i^K(t) \;-\; g_i^K(s) \leqslant \Big(\frac{C\overline{G}(L)\log(K)}{K^a\ h_K} + C\Big)(t-s)$$
$$+ \frac{1}{h_K}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)} \sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)$$
$$\times\Big[\frac{N^K_{\ell+i+1}(u)}{N^K_{i+1}(u)} - \frac{N^K_{\ell+i}(u)}{N^K_{i}(u)}\Big]\,du$$
$$+ \frac{3\|p\|_{\mathrm{Lip}}}{\underline{p}\log(K)}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)} \sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)$$
$$\times\frac{N^K_{\ell+i+1}(u)}{N^K_{i+1}(u)}du.$$

Using that for any real numbers λ,α, $e^\lambda\leqslant e^\alpha + e^\lambda(\lambda-\alpha)$, we have

$$\frac{1}{h_K}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)} \sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)\Big[\frac{N^K_{\ell+i+1}(u)}{N^K_{i+1}(u)} - \frac{N^K_{\ell+i}(u)}{N^K_{i}(u)}\Big]\,du$$
$$\leqslant \frac{\log(K)}{h_K}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)} \sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)$$
$$\times\Big(\beta^K_{\ell+i+1}\Big(\frac{u}{\log(K)}\Big) - \beta^K_{\ell+i}\Big(\frac{u}{\log(K)}\Big)$$
$$-\Big(\beta^K_{i+1}\Big(\frac{u}{\log(K)}\Big) - \beta^K_{i}\Big(\frac{u}{\log(K)}\Big)\Big)\Big)\frac{N^K_{\ell+i+1}(u)}{N^K_{i+1}(u)}du$$
$$\leqslant \int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)} \sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)$$
$$\times\Big(\Delta_K\beta^K_{\ell+i}\Big(\frac{u}{\log(K)}\Big) - \Delta_K\beta^K_{i}\Big(\frac{u}{\log(K)}\Big)\Big)\frac{N^K_{\ell+i+1}(u)}{N^K_{i+1}(u)}du$$
$$\leqslant \int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)} \sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)$$
$$\times\Big(\Delta_K M^K_{\ell+i}\Big(\frac{u}{\log(K)}\Big) - \Delta_K M^K_{i}\Big(\frac{u}{\log(K)}\Big)$$
$$+\Delta_K A^K_{\ell+i}\Big(\frac{u}{\log(K)}\Big) - \Delta_K A^K_{i}\Big(\frac{u}{\log(K)}\Big)\Big)\frac{N^K_{\ell+i+1}(u)}{N^K_{i+1}(u)}du,$$

using (2.10). Thus, using (2.3) and (3.10), we deduce that

$$g_i^K(t) - g_i^K(s) - \int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)} \sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)$$
$$\times\Big(g^K_{\ell+i}\Big(\frac{u}{\log(K)}\Big) - g^K_{i}\Big(\frac{u}{\log(K)}\Big)\Big)\frac{N^K_{\ell+i+1}(u)}{N^K_{i+1}(u)}du$$

$$
\begin{aligned}
&\leqslant C_0(K,L)(t-s)\\
&+\frac{3\|p\|_{\mathrm{Lip}}}{\underline{p}\log(K)}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)\frac{N^K_{\ell+i+1}(u)}{N^K_{i+1}(u)}du\\
&-\frac{\|p\|_{\mathrm{Lip}}}{\underline{p}}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)\\
&\qquad\qquad\times\Big(\beta_{\ell+i+1}\Big(\frac{u}{\log(K)}\Big)-\beta_{i+1}\Big(\frac{u}{\log(K)}\Big)\Big)\frac{N^K_{\ell+i+1}(u)}{N^K_{i+1}(u)}du\\
&+\frac{\|p\|_{\mathrm{Lip}}}{\underline{p}}\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)\\
&\qquad\qquad\times\Big(M_{\ell+i+1}\Big(\frac{u}{\log(K)}\Big)-M_{i+1}\Big(\frac{u}{\log(K)}\Big)\Big)\frac{N^K_{\ell+i+1}(u)}{N^K_{i+1}(u)}du\\
&+\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)\\
&\qquad\qquad\times\Big(\Delta_K M^K_{\ell+i}\Big(\frac{u}{\log(K)}\Big)-\Delta_K M^K_{i}\Big(\frac{u}{\log(K)}\Big)\Big)\frac{N^K_{\ell+i+1}(u)}{N^K_{i+1}(u)}du,
\end{aligned}
$$

where

$$
C_0(K,L)=C+\frac{C\overline{G}(L)}{\delta_K K^a}.
$$

Now, using (3.2), we have on the event $\Omega_K(L)$ (recall (3.5)) whose probability tends to 1 (by Lemma 3.2) that

$$
\begin{aligned}
&\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)\\
&\qquad\qquad\times\Big|\Delta_K M^K_{\ell+i}\Big(\frac{u}{\log(K)}\Big)-\Delta_K M^K_{i}\Big(\frac{u}{\log(K)}\Big)\Big|\frac{N^K_{\ell+i+1}(u)}{N^K_{i+1}(u)}du\\
&\qquad\leqslant 2\overline{p}\log(K)(t-s)\varepsilon_K\sum_{\ell\in\mathbb{Z}}h_K G(h_k\ell)e^{Lh_K|\ell|}\leqslant 2\overline{p}\log(K)\varepsilon_K\overline{G}(L)(t-s),
\end{aligned}
$$

using (2.5) again. Similarly, using (3.7), we have that on the event $\Omega_K(L)$, with a probability tending to 1,

$$
\begin{aligned}
&\int_{(s\wedge\theta_K)\log(K)}^{(t\wedge\theta_K)\log(K)}\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K)G(h_K\ell)\\
&\qquad\qquad\times\Big|M_{\ell+i+1}\Big(\frac{u}{\log(K)}\Big)-M_{i+1}\Big(\frac{u}{\log(K)}\Big)\Big|\frac{N^K_{\ell+i+1}(u)}{N^K_{i+1}(u)}du\\
&\qquad\qquad\qquad\leqslant\overline{p}\log(K)\varepsilon_K\overline{G}(L)(t-s).
\end{aligned}
$$

To conclude, we use the inequality $e^x(3-x)\leqslant e^2$ for all $x\in\mathbb{R}$ to deduce that

$$\frac{N^K_{\ell+i+1}(u)}{N^K_{i+1}(u)}\left[3-\log(K)\left(\beta_{\ell+i+1}\Big(\frac{u}{\log(K)}\Big)-\beta_{i+1}\Big(\frac{u}{\log(K)}\Big)\right)\right]\leqslant e^2.$$

Combining the four previous inequalities, we deduce that on the event $\Omega_K(L)$,

$$\begin{aligned}g^K_i(t)-g^K_i(s)&\leqslant C(K,L)(t-s)\\&\quad+\overline{p}\log(K)\int_{(s\wedge\theta_K)}^{(t\wedge\theta_K)}\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor}h_KG(h_K\ell)\left[g^K_{\ell+i}(v)-g^K_i(v)\right]^+\,e^{h_KL|\ell|}dv,\end{aligned}$$

where $[x]^+=x\vee 0$ is the positive part of x and

$$C(K,L)=C+\frac{C\overline{G}(L)}{\delta_KK^a}+2\overline{p}\log(K)\varepsilon_K\overline{G}(L)+\frac{\|p\|_{\mathrm{Lip}}}{\underline{p}}\overline{p}e^2.$$

Thus, for all $t\leqslant\theta_K(L)\wedge T$, and on the event $\Omega_K(L)$,

$$\frac{dg^K_i(t)}{dt}\leqslant C(K,L)+\overline{p}\log(K)\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor}h_KG(h_K\ell)\left[g^K_{\ell+i}(t)-g^K_i(t)\right]^+\,e^{h_KL|\ell|}.$$

Step 2. — We next provide an upper bound on $g^K_i(t)$. To this end, we will use the maximum principle for $\omega\in\Omega_K(L)$ fixed. Defining $\widetilde{g}^K_i(t)=g^K_i(t)-2C(K,L)t$, we deduce that for any $t\leqslant\theta_K(L)\wedge T$,

$$\frac{d\widetilde{g}^K_i(t)}{dt}<\overline{p}\log(K)\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor}h_KG(h_K\ell)\left[\widetilde{g}^K_{\ell+i}(t)-\widetilde{g}^K_i(t)\right]^+\,e^{h_KL|\ell|}.\tag{3.14}$$

Let us introduce

$$(i_K,t_K)=(i_K(\omega),t_K(\omega))=\operatorname{argmax}_{i\in\{0,\dots,1/\delta_K-1\},t\in[0,\theta_K(\omega)\wedge T]}\ \widetilde{g}^K_i(t)$$

and let us prove that

$$t_K=0.$$

By contradiction, if we assume that $t_K>0$, then the right term of (3.14) is non-positive for $i=i_K$ and then the left term is negative, contradicting the fact that $\widetilde{g}^K_{i_K}(t)$ is maximal for $t=t_K$. Hence, we have proved that, almost surely on the event $\Omega_K(L)$, for all $t\leqslant\theta_K(L)\wedge T$ and $0\leqslant i\leqslant 1/\delta_K-1$,

$$\begin{aligned}g^K_i(t)=\widetilde{g}^K_i(t)+2C(K,L)t&\leqslant\max_{0\leqslant j\leqslant 1/\delta_K-1}\widetilde{g}^K_j(0)+2C(K,L)t\\&=\max_{0\leqslant j\leqslant 1/\delta_K-1}g^K_j(0)+2C(K,L)t.\end{aligned}$$

Step 3. — We next provide an upper bound on $|\Delta_K \beta_i^K|$ which will allow us to conclude the proof. By Proposition 3.3, for $t \leqslant T$,

$$\begin{aligned}\Delta_K \beta_i^K(t\wedge\theta_K) &= g_i^K(t) - \frac{\|p\|_{\mathrm{Lip}}}{\underline{p}} A_{i+1}^K(t\wedge\theta_K) + \Delta_K M_i^K(t\wedge\theta_K)\\ &\leqslant \max_{0\leqslant j\leqslant 1/\delta_K} g_j^K(0) + 2C(K,L)t + \frac{\|p\|_{\mathrm{Lip}}}{\underline{p}}\Big(\max_{0\leqslant j\leqslant 1/\delta_K-1}\beta_i^K(0)+C_1 t\Big)+\varepsilon_K.\end{aligned}$$

A similar argument applied to $(\beta_i^K(t\wedge\theta_K)-\beta_{i-1}^K(t\wedge\theta_K))/\delta_K$ gives the converse inequality, so, recalling that $A_i^K(0)=\beta_i^K(0)$, we finally obtain that there exists a constant C independent of K, i, t and L such that, almost surely on the event $\Omega_K(L)$, for all $i\in\{0,\dots,1/\delta_K-1\}$ and $t\leqslant T$,

$$\begin{aligned}|\Delta_K\beta_i^K(t\wedge\theta_K)| \leqslant C\Big[\max_{0\leqslant j\leqslant 1/\delta_K-1}\big(|\Delta_K\beta_i^K(0)|+\beta_i^K(0)\big)+1+T\\ +\overline{G}(L)T\Big(\frac{1}{\delta_K K^a}+\varepsilon_K\log(K)\Big)\Big].\end{aligned}$$

Finally, defining $\widetilde{\Omega}_K$ as the event of probability converging to 1 where

$$\max_{0\leqslant j\leqslant 1/\delta_K-1}|\Delta_K\beta_i^K(0)|+\beta_i^K(0)\leqslant A+\|\beta_0\|_\infty+1,$$

where the constant A comes from Assumption 2.2, on the event $\Omega_K(L)\cap\widetilde{\Omega}_K$, we have for $t\leqslant T$,

$$|\Delta_K\beta_i^K(t\wedge\theta_K)|\leqslant C\Big[A+\|\beta_0\|_\infty+2+T+\overline{G}(L)T\Big(\frac{1}{\delta_K K^a}+\varepsilon_K\log(K)\Big)\Big].$$

To conclude the proof of Proposition 3.6, we first fix $T>0$, set

$$L_T = C(A+\|\beta_0\|_\infty+3+T)$$

and choose K large enough such that $C\,T\,\overline{G}(L)\big(\frac{1}{\delta_K K^a}+\varepsilon_K\log(K)\big)<1$. Then

$$\mathbb{P}(\tau_K(L_T)>T)\geqslant\mathbb{P}(\Omega_K(L_T)\cap\widetilde{\Omega}_K)\xrightarrow[K\to\infty]{}1.$$

Combining the last estimate with Lemma 3.5 ends the proof of Proposition 3.6. $\square$

3.4. Proof of Theorem 2.4. — In a first step, we prove that the sequence of laws of $(\widetilde{\beta}_t^K, t\in[0,T])_K$ is tight in $\mathcal{P}(\mathbb{D}([0,T],\mathcal{C}(\mathbb{T},\mathbb{R})))$. We will then check that it is actually C-tight.

Let us recall that the random functions $\widetilde{\beta}^K\in\mathbb{D}([0,T],\mathcal{C}(\mathbb{T},\mathbb{R}))$ are defined in (2.6) as follows. For all $x\in\mathbb{T}$, let $i\in\{0,\dots,1/\delta_K-1\}$ be such that $x\in[i\delta_K,(i+1)\delta_K)$, and set

$$\widetilde{\beta}_t^K(x):=\widetilde{\beta}^K(t,x)=\beta_i^K(t)\Big(1-\frac{x}{\delta_K}+i\Big)+\beta_{i+1}^K(t)\Big(\frac{x}{\delta_K}-i\Big),$$

where, by convention, $\beta_{1/\delta_K}^K(t)=\beta_0^K(t)$. Let us recall that the proof of Theorem 2.4 is based on the criterion of Jakubowski [22] recalled in Section 3. Our goal is to prove Conditions (i) and (ii) therein.

Let us first prove (i). By Ascoli's theorem, we know that a compact set K_ε is a set of equi-continuous and equi-bounded functions. By Corollary 3.4 and Proposition 3.6, we have, on the event $\{\theta_K(L_T)>T\}$ of probability converging to 1 when K tends to infinity, that, for all $x\in\mathbb{T}$ and all $t\in[0,T]$,

$$\widetilde{\beta}_t^K(x)=(x-i\delta_K)\Delta_K\beta_i^K(t)+\beta_i^K(t)\leqslant L_T\delta_K+C(T),$$

so the sequence $(\widetilde{\beta}_t^K,t\in[0,T])_K$ is equi-bounded. Furthermore, recall that, by (2.22), for $x,y\in\mathbb{T}$,

$$\Big|\widetilde{\beta}_t^K(x)-\widetilde{\beta}_t^K(y)\Big|=\rho(x,y)\sup_{0\leqslant j\leqslant 1/\delta_K-1}\big|\Delta_K\beta_j^K(t)\big|\leqslant L_T\rho(x,y).$$

We deduce that the sequence is equi-continuous and (i) is proved.

Let us now prove (ii), i.e., that for all $f\in\mathcal{C}(\mathbb{T},\mathbb{R})$, the sequence of laws of the real-valued processes

$$\begin{aligned}X_f^K(t)=\int_{\mathbb{T}}\widetilde{\beta}^K(t,x)f(x)dx=\sum_{i=0}^{1/\delta_K-1}\bigg[\beta_i^K(t)\int_{i\delta_K}^{(i+1)\delta_K}\Big(1+i-\frac{x}{\delta_K}\Big)f(x)dx\\ +\beta_{i+1}^K(t)\int_{i\delta_K}^{(i+1)\delta_K}\Big(\frac{x}{\delta_K}-i\Big)f(x)dx\bigg]\end{aligned}$$

is tight. Recalling that $\beta_i^K(t)=A_i^K(t)+M_i^K(t)$, the process X_f^K is a local semi-martingale with Doob-Meyer decomposition $X_f^K=A_f^K+M_f^K$, where

$$A_f^K(t)=\sum_i\bigg[A_i^K(t)\int_{i\delta_K}^{(i+1)\delta_K}\Big(1+i-\frac{x}{\delta_K}\Big)f(x)dx+A_{i+1}^K(t)\int_{i\delta_K}^{(i+1)\delta_K}\Big(\frac{x}{\delta_K}-i\Big)f(x)dx\bigg]$$

and M_f^K is defined similarly using $M_i^K(t)$ instead of $A_i^K(t)$.

We use Aldous and Rebolledo criteria (see for example Joffe-Métivier [23]) to prove the tightness of the sequence (X_f^K). Let S be a stopping time for the filtration of the underlying Poisson point measures, a.s. in $[0,T]$. We need to estimate for $\alpha>0$, the quantity $\mathbb{P}(|A_f^K((S+\alpha)\wedge T)-A_f^K(S)|>\eta)$ for $\eta>0$. From (2.11), we deduce

$$\begin{aligned}A_f^K((S+\alpha)\wedge T)-A_f^K(S)=\sum_i\bigg\{\bigg(\int_{i\delta_K}^{(i+1)\delta_K}\Big(1+i-\frac{x}{\delta_K}\Big)f(x)dx\bigg)\\ \bigg[\frac{1}{\log(K)}\int_{S\log(K)}^{((S+\alpha)\wedge T)\log(K)}\Big(b(i\delta_K)N_i^K(s)\log\big(1+1/N_i^K(s)\big)\\ +d(i\delta_K)N_i^K(s)\log\big(1-1/N_i^K(s)\big)\Big)ds\\ +\frac{1}{\log(K)}\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor}h_Kp((\ell+i)\delta_K)G(h_K\ell)\\ \times\int_{S\log(K)}^{((S+\alpha)\wedge T)\log(K)}N_{\ell+i}^K(s)\log\big(1+1/N_i^K(s)\big)\,ds\bigg]\end{aligned}$$

$$
\begin{aligned}
+ \Big(\int_{i\delta_K}^{(i+1)\delta_K} (\frac{x}{\delta_K} - i) f(x)dx\Big) \\
\times \Big[\frac{1}{\log(K)} \int_{S\log(K)}^{((S+\alpha)\wedge T)\log(K)} \Big(b((i+1)\delta_K) N_{i+1}^K(s) \log\big(1 + 1/N_{i+1}^K(s)\big) \\
+ d((i+1)\delta_K) N_{i+1}^K(s) \log\big(1 - 1/N_{i+1}^K(s)\big)\Big) ds \\
+ \frac{1}{\log(K)} \sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K - 1 - \lfloor 1/2\delta_K\rfloor} h_K p((\ell+i+1)\delta_K) G(h_K\ell) \\
\times \int_{S\log(K)}^{((S+\alpha)\wedge T)\log(K)} N_{\ell+i+1}^K(s) \log\big(1 + 1/N_{i+1}^K(s)\big)\, ds\Big]\Big\}.
\end{aligned}
$$

Using (3.12) and the definition of $\theta_K(L_T)$, proceeding as in the proof of Proposition 3.6 we have

$$
\begin{aligned}
&\mathbb{E}(|A_f^K((S+\alpha)\wedge\theta_K(L_T)\wedge T) - A_f^K(S\wedge\theta_K(L_T))|) \\
&\quad\leqslant \frac{C}{\log(K)} \sum_{i=0}^{1/\delta_K - 1} \int_{i\delta_K}^{(i+1)\delta_K} |f(x)| dx \Big\{ 2(\overline{b} + \overline{d})\, \alpha \log(K) \\
&+ \sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K - 1 - \lfloor 1/2\delta_K\rfloor} h_K p((\ell+i)\delta_K) G(h_K\ell) \times \mathbb{E}\Big(\int_{(S\wedge\theta_K(L_T))\log(K)}^{((S+\alpha)\wedge\theta_K(L_T)\wedge T)\log(K)} \frac{N_{\ell+i}^K(s)}{N_i^K(s)} ds\Big) \\
&\quad + \sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K - 1 - \lfloor 1/2\delta_K\rfloor} h_K p((\ell+i+1)\delta_K) G(h_K\ell) \\
&\qquad \times \mathbb{E}\Big(\int_{(S\wedge\theta_K(L_T))\log(K)}^{((S+\alpha)\wedge\theta_K(L_T)\wedge T)\log(K)} \frac{N_{\ell+i+1}^K(s)}{N_{i+1}^K(s)} ds\Big)\Big\} \\
&\quad\leqslant \alpha C\, \big[2(\overline{b} + \overline{d}) + 2\overline{p}\,\overline{G}(L)\big]\, \|f\|_\infty.
\end{aligned}
$$

By Proposition 3.6, $\theta_K(L_T) > T$ with probability converging to 1, so we deduce from Markov's inequality that, for all $\varepsilon > 0$ and $\eta > 0$, there exists α such that,

$$
\limsup_{K\to+\infty} \sup_S \mathbb{P}(|A_f^K((S+\alpha)\wedge T) - A_f^K(S)| > \eta) \leqslant \varepsilon,
$$

where the supremum is taken over all stopping times $S \leqslant T$. This is Aldous criterion for $A_f^K(t)$.

It remains to prove a similar property replacing A_f^K by $\langle M_f^K\rangle$. This can be done similarly using (2.16). Computations are actually simpler by Lemma 3.1.

Hence we have proved that the sequence of laws of $(\widetilde{\beta}_t^K, t \in [0,T])_K$ is relatively compact in $\mathcal{P}(\mathbb{D}([0,T], \mathcal{C}(\mathbb{T},\mathbb{R})))$. We prove below that this sequence is actually C-tight and that the Lipschitz estimate (2.8) is satisfied with the value of L_T of Proposition 3.6.

Since $|\beta_i^K(t)-\beta_i^K(t-)|\leqslant C/\log(K)$ for any K, i and t, we have

$$\lim_K \mathbb{P}(\sup_{t\leqslant T}\|\widetilde{\beta}_t^K-\widetilde{\beta}_{t-}^K\|_\infty>\varepsilon)=0.$$

Then, we deduce from Proposition 3.26 in Jacod-Shiryaev [21, p. 315] that, for all $f\in\mathcal{C}(\mathbb{T},\mathbb{R})$, the sequence of laws of X_f^K defined in (3.1) is C-tight. We proceed by contradiction to deduce that $(\widetilde{\beta}^K)$ is also C-tight: if this is not true, there exists an event Ω_1 of positive probability such that, for all $\omega\in\Omega_1$, there exists $t_0(\omega)$, $\alpha(\omega)$ and a ball $B(\omega)\subset\mathbb{T}$ of positive radius such that, for all $x\in B$,

$$|\widetilde{\beta}_{t_0}^K(x)-\widetilde{\beta}_{t_0-}^K(x)|>\alpha. \tag{3.15}$$

Therefore, there exists non-random $\alpha>0$ and $\varepsilon>0$ and $i\in\{0,1,\dots,\lfloor 1/\varepsilon\rfloor-1\}$ and an event $\Omega_2\subset\Omega_1$ of positive probability such that (3.15) holds true for all $x\in[i\varepsilon,(i+1)\varepsilon]$ and for this non-random α. Now, we define $f_i\in\mathcal{C}(\mathbb{T},\mathbb{R})$ with support in $[i\varepsilon,(i+1)\varepsilon]$ and positive on $(i\varepsilon,(i+1)\varepsilon)$. Then, for all $\omega\in\Omega_2$,

$$\liminf_{K\to+\infty}|X_{f_i}^K(t_0)-X_{f_i}^K(t_0-)|\geqslant\alpha\inf_{x\in[(i+1/3)\varepsilon,(i+2/3)\varepsilon]}|f_i(x)|>0.$$

This is a contradiction with the C-tightness of $(X_{f_i}^K)$.

We now prove the Lipschitz estimate for β. Using the Skorohod representation theorem, we can construct copies $\widehat{\beta}^K$ of $\widetilde{\beta}^K$ and $\widehat{\beta}$ of β in the Polish space $\mathbb{D}([0,T],\mathcal{C}(\mathbb{T},\mathbb{R}))$, such that $(\widehat{\beta}^K)$ converges (up to a subsequence) almost surely for the $\mathbb{L}^\infty$ norm on $[0,T]$ to $\widehat{\beta}$. We then define

$$\widehat{\tau}^K=\inf_{i\neq j\in\{0,\dots,1/\delta_K-1\}}\inf\Big\{t\geqslant 0:\frac{|\widehat{\beta}^K(t,i\delta_K)-\widehat{\beta}^K(t,j\delta_K)|}{\rho(i\delta_K,j\delta_K)}>L_T\Big\}.$$

Then $\widehat{\tau}_K$ is distributed as τ_K in (2.21). It then follows from Proposition 3.6 that $\widehat{\tau}_K>T$ with probability converging to 1. Hence, for all $x\neq y\in\mathbb{T}$, almost surely

$$|\widehat{\beta}(t,x)-\widehat{\beta}(t,y)|=\lim_{K\to+\infty}|\widehat{\beta}^K(t,x)-\widehat{\beta}^K(t,y)|\leqslant L_T\,\rho(x,y).$$

By continuity of $\widehat{\beta}$ we deduce that this property holds, almost surely, for all $x,y\in\mathbb{T}$.

4. Identification of the limit as a viscosity solution of a Hamilton-Jacobi equation

Theorem 2.3 will be deduced from Theorems 2.4 and 2.5. In the previous section, we proved Theorem 2.4, i.e., the C-tightness of the laws of $(\widetilde{\beta}_t^K,t\in[0,T])_K$ for all $T>0$. Hence, the sequence of laws of $(\widetilde{\beta}_t^K,t\in[0,T])_K$ admits at least one limiting value. Our aim is now to prove Theorem 2.5, i.e., to identify the limiting path as the unique viscosity solution of the Hamilton-Jacobi equation (2.7).

Let β be distributed as a limiting value of the laws of $(\widetilde{\beta}_t^K,t\in[0,T])_K$. By Theorem 2.4, β belongs to $C([0,T]\times\mathbb{T},[0,+\infty))$. In the sequel and with an abuse of notation, we denote again by $(\widetilde{\beta}_t^K,t\in[0,T])_K$ the subsequence that converges in distribution to β.

We define $\widetilde{A}^K$ as the piecewise affine interpolation of the A_i^K as we did for $\widetilde{\beta}^K$ in (2.6). It follows from Lemma 3.1 that $\widetilde{A}^K - \widetilde{\beta}^K$ converges in law, and thus in probability, to 0. Therefore $(\widetilde{A}^K - \widetilde{\beta}^K, \widetilde{\beta}^K)$ converges in law to $(0,\beta)$ and thus $(\widetilde{A}^K, \widetilde{\beta}^K)$ converges in law to (β,β). We apply the Skorokhod's representation theorem to the random variables $(\widetilde{\beta}^K, \widetilde{A}^K, \beta)$ on the Polish space $\mathbb{D}([0,T],\mathcal{C}(\mathbb{T}))\times\mathcal{C}([0,T]\times\mathbb{T})^2$: there exist a new probability space $(\widehat{\Omega},\widehat{\mathcal{A}},\widehat{\mathbb{P}})$, and random processes $\widehat{\beta}^K, \widehat{A}^K, \widehat{\beta}$ and $\widehat{M}^K = \widehat{\beta}^K - \widehat{A}^K$ on this space such that $(\widehat{A}^K, \widehat{\beta}^K)$ has the same distribution as $(\widetilde{A}^K, \widetilde{\beta}^K)$ and converges almost surely to $(\widehat{\beta},\widehat{\beta})$, where $\widehat{\beta}$ has the same distribution as β. Let us denote by $\widehat{\Omega}_0$ the event where the convergence holds.

We also define for the value L_T introduced in Proposition 3.6

$$\widehat{\Omega}_K = \Big\{\omega\in\widehat{\Omega} : \sup_{x\in\mathbb{T}}\sup_{t\in[0,T]}|\widehat{M}^K(t,x)|\leqslant\varepsilon'_K, \|\widehat{\beta}^K\|_{\mathrm{Lip}}\leqslant L_T, \inf_{(t,x)\in[0,T]\times\mathbb{T}}\widehat{\beta}^K(t,x)\geqslant a\Big\},$$

where $\varepsilon'_K = \delta_K^{-1/2}K^{-a/2}$ converges to 0 by (2.3). It follows from Lemma 3.1, Proposition 3.6 and Lemma 3.5 that $\mathbb{P}(\widehat{\Omega}_K)\to 1$ when $K\to+\infty$.

Notice that, for K fixed, both $\widehat{A}^K$ and $\widehat{\beta}^K$ are càdlàg functions almost surely with values in the set of piecewise affine functions of x between the grid points $i\delta_K$ (this set of càdlàg functions is a measurable subset of $\mathbb{D}([0,T],\mathcal{C}(\mathbb{T}))$). We set

$$\widehat{A}_i^K(t)=\widehat{A}^K(t,i\delta_K),\quad \widehat{\beta}_i^K(t)=\widehat{\beta}^K(t,i\delta_K)\quad\text{and}\quad \widehat{N}_i^K(t)=K^{\widehat{\beta}_i^K(t/\log(K))}$$

for all $i\in\{0,\dots,1/\delta_K-1\}$. In addition, the processes $\widehat{A}_i^K$ and $\widehat{N}_i^K$ satisfy almost surely the relation (2.11) for all $t\in[0,T]$, because (2.11) involves measurable functions of the processes $\widehat{A}_i^K$ and $\widehat{N}_i^K$. We define $\widehat{\Omega}_1$ as the event where these almost-sure properties hold true and we notice that the set

$$\Omega_0=\widehat{\Omega}_0\cap\widehat{\Omega}_1\cap\limsup_{K\to+\infty}\widehat{\Omega}_K$$

has probability 1.

To prove that β is a viscosity sub-solution of Equation (2.7), we work ω by ω in Ω_0. Let $\omega\in\Omega_0$ and $T>0$ and consider a smooth function $\varphi:[0,T]\times\mathbb{T}$ (depending on ω) such that $\widehat{\beta}(\omega)-\varphi$ attains a strict global maximum on $[0,T]\times\mathbb{T}$ at the point $(\bar{t}(\omega),\overline{x}(\omega))$ such that $\bar{t}(\omega)>0$. We will prove that

$$\frac{\partial}{\partial t}\varphi(\bar{t},\overline{x})\leqslant b(\overline{x})-d(\overline{x})+p(x)\int_{\mathbb{R}}G(h)e^{h\partial_x\varphi(\bar{t},\overline{x})}dh.$$

Since $\widehat{A}^K(\omega)$ is continuous and converges in $L^\infty([0,T]\times\mathbb{T})$ to β, there exists for K large enough a local maximum of $\widehat{A}^K(\omega)-\varphi$ on $[0,T]\times\mathbb{T}$ at a point $(t_K(\omega),x_K(\omega))$ such that $(t_K(\omega),x_K(\omega))\to(\bar{t}(\omega),\overline{x}(\omega))$ as $K\to\infty$. Assume K is large enough so that $t_K(\omega)>0$.

From now on, we will omit the dependencies with respect to $\omega\in\Omega_0$ to avoid heavy notation.

Defining $i_K \in \{0,\dots,1/\delta_K-1\}$ such that $i_K\delta_K \leqslant x_K < (i_K+1)\delta_K$, we have

$$\begin{aligned}
\frac{\partial}{\partial t}\widehat{A}^K(t_K,x_K) &= \Big(1-\frac{x_K}{\delta_K}+i_K\Big)\frac{d}{dt}\widehat{A}^K_{i_K}(t_K) + \Big(\frac{x_K}{\delta_K}-i_K\Big)\frac{d}{dt}\widehat{A}^K_{i_K+1}(t_K)\\
&= \Big(1-\frac{x_K}{\delta_K}+i_K\Big)\widehat{N}^K_{i_K}(t_K\log(K))\Big(b(i_K\delta_K)\log\big(1+1/\widehat{N}^K_{i_K}(t_K\log(K))\big)\\
&\qquad\qquad + d(i_K\delta_K)\log\big(1-1/\widehat{N}^K_{i_K}(t_K\log(K))\big)\Big)\\
&\quad + \Big(\frac{x_K}{\delta_K}-i_K\Big)\widehat{N}^K_{i_K+1}(t_K\log(K))\Big(b\big((i_K+1)\delta_K\big)\log\big(1+1/\widehat{N}^K_{i_K+1}(t_K\log(K))\big)\\
&\qquad\qquad + d\big((i_K+1)\delta_K\big)\log\big(1-1/\widehat{N}^K_{i_K+1}(t_K\log(K))\big)\Big)\\
&\quad + \Big(1-\frac{x_K}{\delta_K}+i_K\Big)\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i_K)\delta_K)G(h_K\ell)\\
&\qquad\qquad \times \widehat{N}^K_{\ell+i_K}(t_K\log(K))\log\big(1+1/\widehat{N}^K_{i_K}(t_K\log(K))\big)\\
&\quad + \Big(\frac{x_K}{\delta_K}-i_K\Big)\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i_K+1)\delta_K)G(h_K\ell)\\
&\qquad\qquad \times \widehat{N}^K_{\ell+i_K+1}(t_K\log(K))\log\big(1+1/\widehat{N}^K_{i_K+1}(t_K\log(K))\big).
\end{aligned}$$

Using that for all $x\geqslant -1/2$,

$$b\log(1+x)+d\log(1-x)\leqslant (b-d)x,$$

and using $\widehat{\beta}^K(t,x)\geqslant a$ for K large enough (by definition of Ω_0), we deduce that

$$\begin{aligned}
\frac{\partial}{\partial t}\widehat{A}^K(t_K,x_K) &\leqslant \Big(1-\frac{x_K}{\delta_K}+i_K\Big)\big(b(i_K\delta_K)-d(i_K\delta_K)\big)\Big(1+\frac{C}{K^a}\Big)\\
&\quad + \Big(\frac{x_K}{\delta_K}-i_K\Big)\big(b((i_K+1)\delta_K)-d((i_K+1)\delta_K)\big)\Big(1+\frac{C}{K^a}\Big)\\
&\quad + \Big(1-\frac{x_K}{\delta_K}+i_K\Big)\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i_K)\delta_K)G(h_K\ell)\\
&\qquad\qquad \times e^{\log(K)(\widehat{\beta}^K_{\ell+i_K}(t_K)-\widehat{\beta}^K_{i_K}(t_K))}\Big(1+\frac{C}{K^a}\Big)\\
&\quad + \Big(\frac{x_K}{\delta_K}-i_K\Big)\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K p((\ell+i_K+1)\delta_K)G(h_K\ell)\\
&\qquad\qquad \times e^{\log(K)(\widehat{\beta}^K_{\ell+i_K+1}(t_K)-\widehat{\beta}^K_{i_K+1}(t_K))}\Big(1+\frac{C}{K^a}\Big).
\end{aligned}$$

We next use the fact that μ, b and d are C^1 functions in $\mathbb{T}$ to obtain, modifying the constant C if necessary,

$$\begin{aligned}
\frac{\partial}{\partial t}\widehat{A}^K(t_K,x_K) &\leqslant \big(b(x_K)-d(x_K)+C\delta_K\big)\Big(1+\frac{C}{K^a}\Big)\\
&\quad + (1-\frac{x_K}{\delta_K}+i_K)\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K\big(p(x_K)+C\delta_K(|\ell|+1)\big)G(h_K\ell)\\
&\qquad\qquad \times e^{\log(K)(\widehat{\beta}^K_{\ell+i_K}(t_K)-\widehat{\beta}^K_{i_K}(t_K))}\Big(1+\frac{C}{K^a}\Big)
\end{aligned}\tag{4.1}$$

$$+\,(\frac{x_K}{\delta_K}-i_K)\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K\big(p(x_K)+C\delta_K(|\ell|+1)\big)G(h_K\ell)\\ \times e^{\log(K)(\widehat{\beta}^K_{\ell+i_K+1}(t_K)-\widehat{\beta}^K_{i_K+1}(t_K))}\Big(1+\frac{C}{K^a}\Big).$$

Since (t_K,x_K) is a maximum point of $\widehat{A}^K-\varphi$, we deduce that

$$\begin{aligned}\widehat{\beta}^K_j(t_K)-\widehat{\beta}^K_{i_K}(t_K)&=\widehat{\beta}^K(t_K,j\delta_K)-\widehat{\beta}^K(t_K,x_K)-\Big(\widehat{\beta}^K_{i_K}(t_K)-\widehat{\beta}^K(t_K,x_K)\Big)\\ &\leqslant \varphi(t_K,j\delta_K)-\varphi(t_K,x_K)+\widehat{M}^K(t_K,j\delta_K)-\widehat{M}^K(t_K,x_K)\\ &\qquad -\Big(\widehat{\beta}^K_{i_K}(t_K)-\widehat{\beta}^K(t_K,x_K)\Big)\end{aligned}$$

and similarly for $\widehat{\beta}^K_{j+1}(t_K)-\widehat{\beta}^K_{i_K+1}(t_K)$. In addition,

$$\varphi(t_K,j\delta_K)-\varphi(t_K,x_K)\leqslant (j-i_K)\delta_K\partial_x\varphi(t_K,x_K)+O(|x_K-i_K\delta_K|)+O(|j-i_K|^2\delta_K^2).$$

Therefore, since $\omega\in\limsup\widehat{\Omega}_K$, there exists a subsequence in K (still denoted K) along which

$$\begin{aligned}\widehat{\beta}^K_j(t_K)-\widehat{\beta}^K_{i_K}(t_K)&\leqslant (j-i_K)\delta_K\partial_x\varphi(t_K,x_K)+C\left(|x_K-i_K\delta_K|+|j-i_K|^2\delta_K^2+\varepsilon_K\right)\\ &\leqslant (j-i_K)\delta_K\partial_x\varphi(t_K,x_K)+C(|j-i_K|^2\delta_K^2+\delta_K+\varepsilon_K)\end{aligned}$$

and

$$\widehat{\beta}^K_{j+1}(t_K)-\widehat{\beta}^K_{i_K+1}(t_K)\leqslant (j-i_K)\delta_K\partial_x\varphi(t_K,x_K)+C(|j-i_K|^2\delta_K^2+\delta_K+\varepsilon_K).$$

Combining these inequalities with (4.1), we obtain

$$\begin{aligned}\frac{\partial}{\partial t}\varphi(t_K,x_K)\leqslant\frac{\partial}{\partial t}\widehat{A}^K(t_K,x_K)&\leqslant\big(b(x_K)-d(x_K)+C\delta_K\big)\Big(1+\frac{C}{K^a}\Big)\\ &+\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K(p(x_K)+C\delta_K(|\ell|+1))G(h_K\ell)\\ &\qquad\times e^{h_K\ell\partial_x\varphi(t_K,x_K)+C\frac{(h_K\ell)^2}{\log(K)}+o(1)}\Big(1+\frac{C}{K^a}\Big).\end{aligned}\tag{4.2}$$

Note that the first inequality above comes from the fact that $(t_K,x_K)\in(0,T]\times\mathbb{T}$ is a maximum point of $\widehat{A}^K-\varphi$ in $[0,T]\times\mathbb{T}$, with the maximum being attained possibly on $\{T\}\times\mathbb{T}$. Since $p(x_K)\to p(\overline{x})$ when $K\to+\infty$, to prove the convergence of the sum in the right-hand side of (4.2), it is sufficient to study the convergence of

$$S=\sum_{\ell=-\lfloor 1/2\delta_K\rfloor}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K G\big(h_K\ell\big)e^{h_K\ell\partial_x\varphi(t_K,x_K)+C(h_K\ell)^2/\log(K)}.$$

Recall from Assumption 2.1 that G is continuous and that there exists $R>0$ such that G is nonincreasing on $[R,+\infty)$ and nondecreasing on $(-\infty,-R]$. We first notice that

$$S_0=\sum_{\ell=\lfloor -R/h_K\rfloor}^{\lfloor R/h_K\rfloor} h_K G\big(h_K\ell\big)e^{h_K\ell\partial_x\varphi(t_K,x_K)+C(h_K\ell)^2/\log(K)}$$

is a Riemann sum which converges to $\int_{-R}^{R} G(y)e^{y\partial_x\varphi(\overline{t},\overline{x})}dy$. Hence we only have to deal with the remainder $S-S_0$. We detail the analysis for

$$S_+ = \sum_{\ell=\lfloor R/h_K\rfloor+1}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K G\big(h_K\ell\big)e^{h_K\ell\partial_x\varphi(t_K,x_K)+C(h_K\ell)^2/\log(K)}.$$

A similar computation applies to the lower tail.

For all $\varepsilon>0$, there exists K_0 such that for $K\geqslant K_0$, $|\partial_x\varphi(t_K,x_K)-\partial_x\varphi(\overline{t},\overline{x})|\leqslant\varepsilon$. Hence, setting $q=\partial_x\varphi(\overline{t},\overline{x})$ and recalling that G is nonincreasing on $[R,+\infty)$, for K large enough,

$$\begin{aligned}
S_+ &\leqslant \sum_{\ell=\lfloor R/h_K\rfloor+1}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} h_K G\big(h_K\ell\big)e^{qh_K\ell+C\frac{(h_K\ell)^2}{\log(K)}+\varepsilon h_K|\ell|}\\
&\leqslant \sum_{\ell=\lfloor R/h_K\rfloor+1}^{1/\delta_K-1-\lfloor 1/2\delta_K\rfloor} \int_{h_K(\ell-1)}^{h_K\ell} G\big(h_K\ell\big)e^{qh_K\ell+C\frac{(h_K\ell)^2}{\log(K)}+\varepsilon h_K|\ell|}\,dy\\
&\leqslant \int_R^{h_K/2\delta_K} G(y)e^{qy+|q|h_K+C\frac{y^2+h_K(2|y|+h_K)}{\log(K)}+\varepsilon(|y|+h_K)}\,dy\\
&\leqslant \int_R^{\log(K)} G(y)e^{qy+C\frac{y^2+2h_K|y|}{\log(K)}+\varepsilon|y|}(1+\varepsilon)\,dy.
\end{aligned}$$

Observing that, for $|y|\leqslant(\log(K))^{1/3}$,

$$\frac{y^2+2h_K|y|}{\log(K)}\leqslant(\log(K))^{-1/3}+2\delta_K(\log(K))^{1/3}\longrightarrow 0$$

when $K\to+\infty$ and that, for $|y|\leqslant\log(K)$, $(y^2+2h_K|y|)/\log(K)\leqslant y+2h_K$, we can decompose the domain of integration as

$$[R,\log(K)]=[R,(\log(K))^{1/3})\cup[(\log(K))^{1/3},\log(K)]$$

to deduce that, for K large enough,

$$\begin{aligned}
S_+ \leqslant (1+2\varepsilon)\int_R^{(\log(K))^{1/3}} G(y)e^{qy+\varepsilon|y|}dy+(1+2\varepsilon)\int_{(\log(K))^{1/3}}^{\log(K)} G(y)e^{(q+C)y+\varepsilon|y|}dy\\
\leqslant (1+2\varepsilon)\int_R^{+\infty} G(y)e^{qy+\varepsilon|y|}dy+\frac{1+2\varepsilon}{e^{(\log(K))^{1/3}}}\int_{\mathbb{R}} G(y)e^{(q+C+2)|y|}dy.
\end{aligned}$$

For the second inequality, we used that, for $\varepsilon<1$, $e^{\varepsilon|y|}\leqslant e^{2|y|-(\log(K))^{1/3}}$ for all $y\geqslant(\log(K))^{1/3}$. Now, by dominated convergence,

$$\int_R^{+\infty} G(y)e^{qy+\varepsilon|y|}dy\xrightarrow[\varepsilon\to 0]{}\int_R^{+\infty} G(y)e^{qy}dy.$$

To conclude, recalling that $q=\partial_x\varphi(\overline{t},\overline{x})$, we have proved that

$$\limsup_{K\to+\infty} S\leqslant\int_{\mathbb{R}} G(y)e^{y\partial_x\varphi(\overline{t},\overline{x})}dy.$$

Therefore,

$$\frac{\partial}{\partial t}\varphi(\overline{t},\overline{x})\leqslant b(\overline{x})-d(\overline{x})+p(\overline{x})\int_{\mathbb{R}} G(h)e^{h\partial_x\varphi(\overline{t},\overline{x})}dh.$$

We conclude that β is a viscosity sub-solution of (2.7) in $(0,T]\times\mathbb{T}$.

Following similar arguments, we can prove that β is a viscosity super-solution, and hence a viscosity solution of (2.7) in $(0,T]\times\mathbb{T}$. The result then follows from uniqueness of a Lipschitz viscosity solution of (2.7) [2].

References

[1] V. Bansaye & S. Méléard – *Stochastic models for structured populations. Scaling limits and long time behavior*, Math. Biosciences Institute Lect. Series. Stochastics in Biological Systems, vol. 1.4, Springer, Cham, 2015.

[2] G. Barles – *Solutions de viscosité des équations de Hamilton-Jacobi*, Math. & Applications, vol. 17, Springer-Verlag, Paris, 1994.

[3] G. Barles, S. Mirrahimi & B. Perthame – "Concentration in Lotka-Volterra parabolic or integral equations: a general convergence result", *Methods Appl. Anal.* **16** (2009), no. 3, p. 321–340.

[4] J. Berestycki, E. Brunet, J. W. Harris, S. C. Harris & M. I. Roberts – "Growth rates of the population in a branching Brownian motion with an inhomogeneous breeding potential", *Stochastic Process. Appl.* **125** (2015), no. 5, p. 2096–2145.

[5] J. D. Biggins – "Branching out", in *Probability and mathematical genetics*, London Math. Soc. Lecture Note Ser., vol. 378, Cambridge University Press, Cambridge, 2010, p. 113–134.

[6] S. Billiard, P. Collet, R. Ferrière, S. Méléard & V. C. Tran – "Stochastic dynamics for adaptation and evolution of microorganisms", in *European Congress of Math.*, European Mathematical Society, Zürich, 2018, p. 525–550.

[7] J. Blath, T. Paul & A. Tóbiás – "A stochastic adaptive dynamics model for bacterial populations with mutation, dormancy and transfer", *ALEA Lat. Am. J. Probab. Math. Stat.* **20** (2023), no. 1, p. 313–357.

[8] A. Bovier, L. Coquille & C. Smadi – "Crossing a fitness valley as a metastable transition in a stochastic population model", *Ann. Appl. Probab.* **29** (2019), no. 6, p. 3541–3589.

[9] A. Callegaro & M. I. Roberts – "A spatially- dependent fragmentation process", 2021, `arXiv:2103.09761`.

[10] V. Calvez, S. Figueroa Iglesias, H. Hivert, S. Méléard, A. Melnykova & S. Nordmann – "Horizontal gene transfer: numerical comparison between stochastic and deterministic approaches", in *CEMRACS 2018*, ESAIM Proc. Surveys, vol. 67, EDP Sci., Les Ulis, 2020, p. 135–160.

[11] N. Champagnat – "A microscopic interpretation for adaptive dynamics trait substitution sequence models", *Stochastic Process. Appl.* **116** (2006), no. 8, p. 1127–1160.

[12] N. Champagnat & B. Henry – "A probabilistic approach to Dirac concentration in nonlocal models of adaptation with several resources", *Ann. Appl. Probab.* **29** (2019), no. 4, p. 2175–2216.

[13] N. Champagnat & S. Méléard – "Polymorphic evolution sequence and evolutionary branching", *Probab. Theory Relat. Fields* **151** (2011), no. 1-2, p. 45–94.

[14] N. Champagnat, S. Méléard & V. C. Tran – "Stochastic analysis of emergence of evolutionary cyclic behavior in population dynamics with transfer", *Ann. Appl. Probab.* **31** (2021), no. 4, p. 1820–1867.

[15] L. Coquille, A. Kraut & C. Smadi – "Stochastic individual-based models with power law mutation rate on a general finite trait space", *Electron. J. Probab.* **26** (2021), article no. 123 (37 pages).

[16] U. Dieckmann & R. Law – "The dynamical theory of coevolution: a derivation from stochastic ecological processes", *J. Math. Biol.* **34** (1996), no. 5-6, p. 579–612.

[17] O. Diekmann, P.-E. Jabin, S. Mischler & B. Perthame – "The dynamics of adaptation: An illuminating example and a Hamilton–Jacobi approach", *Theoret. Population Biol.* **67** (2005), no. 4, p. 257–271.

[18] R. Durrett & J. Mayberry – "Traveling waves of selective sweeps", *Ann. Appl. Probab.* **21** (2011), no. 2, p. 699–744.

[19] R. Forien, J. Garnier & F. Patout – "Ancestral lineages in mutation selection equilibria with moving optimum", *Bull. Math. Biol.* **84** (2022), no. 9, article no. 93 (43 pages).

[20] P.-E. Jabin – "Small populations corrections for selection-mutation models", *Netw. Heterog. Media* **7** (2012), no. 4, p. 805–836.

[21] J. Jacod & A. N. Shiryaev – *Limit theorems for stochastic processes*, second ed., Grundlehren Math. Wissen., vol. 288, Springer-Verlag, Berlin, 2003.

[22] A. Jakubowski – "On the Skorokhod topology", *Ann. Inst. H. Poincaré Probab. Statist.* **22** (1986), no. 3, p. 263–285.

[23] A. Joffe & M. Métivier – "Weak convergence of sequences of semimartingales with applications to multitype branching processes", *Adv. in Appl. Probab.* **18** (1986), no. 1, p. 20–65.

[24] P. J. Keeling & J. D. Palmer – "Horizontal gene transfer in eukaryotic evolution", *Nat. Rev. Genet.* (2008), no. 8, p. 605–618.

[25] B. R. Levin & F. M. Stewart – "The population biology of bacterial plasmids: a priori conditions for the existence of mobilizable nonconjugative factors", *Genetics* **94** (1980), no. 2, p. 425–443.

[26] A. Lorz, S. Mirrahimi & B. Perthame – "Dirac mass dynamics in multidimensional nonlocal parabolic equations", *Comm. Partial Differential Equations* **36** (2011), no. 6, p. 1071–1098.

[27] P. Maillard, G. Raoul & J. Tourniaire – "Spatial dynamics of a population in a heterogeneous environment", 2021, `arXiv:2105.06985`.

[28] B. Mallein – "Maximal displacement of a branching random walk in time-inhomogeneous environment", *Stochastic Process. Appl.* **125** (2015), no. 10, p. 3958–4019.

[29] J. Metz, S. Geritz, G. Meszéna, F. Jacobs & J. V. Heerwaarden – "Adaptative dynamics: a geometrical study of the consequences of nearly faithful reproduction", in *Proc. Colloquium (Amsterdam, Jan. 1995)* (S. V. Strien & S. Verduyn Lunel, eds.), Stochastic and Spatial Structures of Dynamical Systems, vol. 45, North-Holland, 1996, p. 183–231.

[30] S. Mirrahimi, G. Barles, B. Perthame & P. E. Souganidis – "A singular Hamilton-Jacobi equation modeling the tail problem", *SIAM J. Math. Anal.* **44** (2012), no. 6, p. 4297–4319.

[31] H. Ochman, J. G. Lawrence & E. A. Groisman – "Lateral gene transfer and the nature of bacterial innovation", *Nature* **405** (2000), p. 299–304.

[32] B. Perthame & G. Barles – "Dirac concentrations in Lotka-Volterra parabolic PDEs", *Indiana Univ. Math. J.* **57** (2008), no. 7, p. 3275–3301.

[33] B. Perthame & M. Gauduchon – "Survival thresholds and mortality rates in adaptive dynamics: conciliating deterministic and stochastic simulations", *Math. Med. Biol.* **27** (2010), no. 3, p. 195–210.

[34] D. Waxman & S. Gavrilets – "20 questions on adaptive dynamics", *J. Evol. Biol.* **18** (2005), no. 5, p. 1139–1154.

Manuscript received 16th June 2022
accepted 6th September 2023

Nicolas Champagnat, Université de Lorraine, CNRS, Inria, IECL
F-54000 Nancy, France
E-mail : `nicolas.champagnat@inria.fr`
Url : `http://nchampagnat.perso.math.cnrs.fr/`

Sylvie Méléard, École Polytechnique & Institut Universitaire de France, CNRS, Institut polytechnique de Paris
route de Saclay, 91128 Palaiseau Cedex, France
E-mail : `sylvie.meleard@polytechnique.edu`
Url : `https://sites.google.com/view/sylvie-meleard/accueil`

Sepideh Mirrahimi, Institut Montpelliérain Alexander Grothendieck, Univ. Montpellier, CNRS
Place Eugène Bataillon, 34090 Montpellier, France
E-mail : `sepideh.mirrahimi@umontpellier.fr`
Url : `https://imag.umontpellier.fr/~mirrahimi/`

Viet Chi Tran, LAMA, Univ Gustave Eiffel, Univ Paris Est Creteil, CNRS
F-77454 Marne-la-Vallée, France
E-mail : `chi.tran@univ-eiffel.fr`
Url : `https://perso.math.u-pem.fr/tran.viet-chi/`