Convergence of population processes with small and frequent mutations to the canonical equation of adaptive dynamics

Nicolas Champagnat, Vincent Hass

Convergence of population processes with small and frequent mutations to the canonical equation of adaptive dynamics

Nicolas Champagnat[*], Vincent Hass[†1]

July 7, 2024

[*] *Université de Lorraine, CNRS, Inria, IECL, F-54000 Nancy, France*
[†] *Université de Franche-Comté, IUT Nord Franche-Comté, F-90000 Belfort, France*

Abstract. In this article, a stochastic individual-based model describing Darwinian evolution of asexual, phenotypic trait-structured population, is studied. We consider a large population with constant population size characterised by a resampling rate modeling competition pressure driving selection and a mutation rate where mutations occur during life. In this model, the population state at fixed time is given as a measure on the space of phenotypes and the evolution of the population is described by a continuous time, measure-valued Markov process. We investigate the asymptotic behaviour of the system, where mutations are frequent, in the double simultaneous limit of large population ($K \to +\infty$) and small mutational effects ($\sigma_K \to 0$) proving convergence to an ODE known as the canonical equation of adaptive dynamics. This result holds only for a certain range of σ_K parameters (as a function of K) which must be small enough but not too small either. The canonical equation describes the evolution in time of the dominant trait in the population driven by a fitness gradient. This result is based on an slow-fast asymptotic analysis. We use an averaging method, inspired by Kurtz [57], which exploits a martingale approach and compactness-uniqueness arguments. The contribution of the fast component, which converges to the centered Fleming-Viot process, is obtained by averaging according to its invariant measure, recently characterised in [18].

Keywords. Adaptive dynamics, Canonical equation, Individual-based model, Measure-valued Markov process, Slow-fast asymptotic analysis, Averaging method, Centered Fleming-Viot process.

MSC subject classification. Primary 60B10, 60G44, 60G57, 92D10, 92D25, 92D40; Secondary 60F10, 60G10, 60J35, 60J60, 60J68.

1 Introduction

In this article we study, at the individual level and in the interplay between ecology and Darwinian evolution, a population model, structured by a 1−dimensional quantitative phenotypic

[1]Corresponding author at: IUT Nord Franche-Comté, Département Génie Civil et Construction Durable, 19 Avenue du Maréchal Juin, 90 016 Belfort, Cedex, France.
E-mail addresses: nicolas.champagnat@inria.fr (Nicolas Champagnat), vincent.hass@univ-fcomte.fr (Vincent Hass).

trait. The Darwinian evolution is based on three basic mechanisms. Firstly, *heredity* which allows the transmission of the individual phenotypic characteristics from one generation to another. Secondly, a source of variation in the individual phenotypic characteristics: in our case it is only *mutations*. Finally, a *selection* mechanism which can result from interaction between individuals in the population such as competition. Our model is an *individual-based model* (in short, IBM) which involves a finite and asexual population with constant population size in which each individual's birth, death and mutation events are described. IBMs were first introduced in ecology as a tool to describe local interactions or complex phenomena at the individual level [73, 12, 13, 33, 52, 28]. Ecological studies using IBMs are mainly numerical and the models are mostly posed in discrete space as systems of interacting particles [73] and more rarely in continuous space [12, 13, 33]. Many IBMs (with non-constant population size) have been proposed and studied in the context of Darwinian evolution by the biology community [68, 31, 41] and the mathematical community [16, 17, 5]. Others are dispersal models in spatially structured populations where the trait is viewed as a spatial location and mutations as dispersal [12, 13, 62, 46]. Other models, structured in age, were developed in [22] and studied mathematically in [74, 82], or structured in age and traits in [42, 66, 67].

We consider an IBM with fixed population size, so that births and deaths occur simultaneously in so-called *resampling* (or swap phenomenon) events. The mutation and resampling rate of an individual depends on its phenotype. When a mutation occurs, the new mutant trait is close to its parent's one yielding a slow variation of the trait. In population genetics, the Wright-Fisher model (and its extensions with selection, mutation or immigration), Cannings model or the Moran model are interested in the evolution of allele frequencies according to various mechanisms [38]. In [43, 27, 26, 37], the authors construct the Fleming-Viot process as a scaling limit of large population from the Moran process. Several extensions exist, including frequency-dependent selection, recombination, other reproduction mechanisms [39, 51]. In particular, the last article provides a bridge between population genetics and eco-evolutionary models. Others, based on Kermack-McKendrick's model, are interested in epidemiological questions [70, 71].

The aim of this article is to describe the evolutionary dynamics of the dominant trait, at the population level, on a long time scale, as a solution to an ordinary differential equation (in short, ODE) called *Canonical Equation of the Adaptive Dynamics* (in short, CEAD). Canonical equations are well-known tools in evolutionary biology, used to predict the evolutionary fate of ecological communities. More precisely, such equations describe the evolution of dominant traits in a biological population as driven by mutations and a fitness gradient which describes the strength of selection that pushes the population to locally increase its fitness [85, 61, 81, 30].

Fitness measures the selective value of a given individual in a given environment including the population under consideration itself. This individual can be any (fictitious) mutant individual that can be born in the population at the time under consideration. The way to construct a fitness landscape depends on the ecological context [69]. For continuous time homogeneous Markov models as studied below, it is the instantaneous growth rate (birth rate minus death rate) of the individual considered in the environment considered. If we further assume that the population constituting the environment is in a stationary state, then we can consider that the fitness of a given mutant individual in this population governs the possibility of invasion of the descendants of this mutant in the population.

The canonical equation has been studied and derived in different contexts such as game theory [54] and quantitative genetics [61]. In the branch of evolutionary biology called *adap-*

tive dynamics [55, 72, 65], the CEAD has been introduced heuristically in [30]. The theory of adaptive dynamics studies the links between ecology and Darwinian evolution, more precisely, it investigates the effects of the ecological aspects of population dynamics on the evolutionary process, and so describes the population dynamics on the phenotypic level instead of the genotypic level. The theory of adaptive dynamics is based on biological assumptions of *rare* and *small mutations* and of *large population* under which the CEAD was proposed.

Two mathematical approaches were developed to give a proper mathematical justification of this theory: a *deterministic* one, and a *stochastic* one. All these approaches are based on the use of IBMs and different combinations of the three previous biological assumptions and consider a parameter scaling under which the population distribution over the trait space concentrates to DIRAC masses, i.e. to subpopulations in which all the individuals have the same trait. The canonical equation corresponds to the motion of DIRAC masses. Let us introduce three parameters corresponding to the different biological assumptions: p for the mutation probability, σ for the mutation size and K for the population size. In the sequel, we denote $p_K := p(K)$ and $\sigma_K := \sigma(K)$ to emphasise the dependence of K.

In the deterministic approach, [46] first establishes that, under the asymptotic of large population ($K \to +\infty$), the IBM converges in law to a deterministic process which is a weak solution of a partial differential equation. Then by adding the small mutation assumption ($\sigma \to 0$) and an appropriate time scaling $1/\sigma$, [32, 78, 64] establish the convergence to a version of the canonical equation different from the one of [30] and described by a HAMILTON-JACOBI equation with constraint.

The stochastic approach was developed in [16, 20]. In [16] it is proved that the IBM converges, for finite dimensional distributions, under the double simultaneous asymptotic of large population ($K \to +\infty$) and rare mutation ($p_K \to 0$) to a stochastic process: the *TSS* for *Trait Substitution Sequence* introduced in [68, Section 6.4]. This is a pure jump MARKOV process in the trait space where the population is at all times monomorphic and where the jumps describe the invasion and then the fixation of a mutant y in a monomorphic resident population of trait x. By adding the small mutation assumption ($\sigma \to 0$) to the TSS, its convergence to the canonical equation proposed in [30] is established in [20]. The time scale involved in observing the canonical equation phenomenon is $1/K\sigma_K^2 p_K$. Extensions of these results were obtained in [19] for chemostat models, in [35, 66] for age-structured populations models, in [63] for spatial-structured populations models and [2] which studies the simultaneous application of the three limits ($K \to +\infty, p_K \to 0, \sigma_K \to 0$) in order to determine precisely the range of parameters leading to the canonical equation. Note that another approach, studied in [23, 36], consider models of large population ($K \to +\infty$) scaling with rare-but-not-too-rare mutations (with power-law mutation rates $p_K := K^{-\alpha} \to 0$). This specific mutational scale implies that negligible subpopulations of size K^β, $\beta \in (0,1)$ may have a strong contribution to evolution. A similar scaling was also studied in [34, 15, 21, 11] with fixed mutation size. Building on these works, [76] recently proposed a derivation of a CEAD assuming small mutations.

Despite their success, the proposed approaches are criticised by biologists [83, 79]. Among the biological assumptions of adaptive dynamics, the assumption of rare mutations is the most critised as unrealistic [83]. The rate of molecular mutation is relatively well known and generally involves several nucleotide substitutions per generation. The adaptive dynamics response is based on the fact that only non-synonymous mutations (changing phenotypes) producing viable individuals should be taken into account [1] [68, Section 6.4]. Since only a small fraction of DNA codes for proteins and many mutations produce non-functional proteins, and thus non-viable

individuals, it is not unreasonable to assume that mutations are rare, but probably not as rare as assumed in the stochastic approach [16, 20, 2]. Another criticism of stochastic approaches is that the phenomenon of the canonical equation takes place on a too long evolutionary time scale. In order to solve these problems, we propose to apply a double simultaneous asymptotic of small mutations ($\sigma_K \to 0$) and large population ($K \to +\infty$), but *frequent mutations* ($p \equiv 1$). After conveniently scaling the population state, this leads to a *slow-fast dynamics* where the phenomenon of the canonical equation is visible on a time scale $1/K\sigma_K^2$. So, there is some consistency between the previous stochastic approaches developed in [20, 2] and ours: it is exactly the same canonical equation. However in our situation, the CEAD is visible on a shorter evolutionary scale than theirs, and therefore is biologically more reasonable. Note however that the restrictions on σ_K in [2] are in a completely different regime than ours ($\sigma_K \gg K^{-1/2}$, as opposed to $\sigma_K \ll K^{-2}$ here). Hence it is possible to choose p_K in [2] in such a way that the CEAD appears on a faster time scale than in our work, where the fastest scale is in K^3.

As the mutation size parameter is small, the population distribution tends to be close to a Dirac mass. Our aim is to describe the evolution of the support of this Dirac mass on a large time scale. The mean trait appears as the natural slow component. It will be proved to act on the time scale $1/K\sigma_K^2$. In our case, the fast component acts on the time scale K and is given by the dynamics of the centered and dilated distribution of traits corresponding to a discrete version of the centered Fleming-Viot process [18]. Since the centered Fleming-Viot process is ergodic [18], we expect the centered and dilated distribution of traits to stabilise on the slow time scale, hence the distribution of traits should stabilise to a Dirac mass. Therefore, the dynamics of the dominant trait corresponds to that of the mean trait.

The reason for not considering the same IBM as in [16, 20, 2] is because it involves three time scales, a slow one corresponding to the dynamics of the mean trait acting on the evolutionary time scale $1/K\sigma_K^2$, a fast one corresponding to the dynamics of the centered, dilated distribution of traits acting on the evolutionary time scale K and a very fast one corresponding to the population size dynamics acting on the ecological time scale 1. For simplicity, we focus here on a model with constant population size to reduce the number of time scales to two. We expect our results to extend to general IBMs as in [16, 20, 2] and we leave this for future works.

To prove convergence in the framework of slow-fast dynamics (also called *stochastic singular perturbations*), different techniques can be used.

Firstly, the method of the *perturbed test function* initially proposed by Papanicolaou, Stroock and Varadhan in [75] identifies the generator of the limit process with a martingale approach whose idea is the following. If we consider a family of stochastic processes $((X^\varepsilon, Y^\varepsilon))_{\varepsilon>0}$ of generator L^ε where X^ε is the *slow component* and Y^ε is the *fast component* in the form $Y^\varepsilon(t) = Y(t/\varepsilon)$ where Y is a Markov process, the slow-fast problem consists in identifying the limit process of X^ε using ergodicity properties of the fast dynamics. Assuming that the family $(X^\varepsilon)_{\varepsilon>0}$ is tight, we consider X a limiting value. We can expect to characterise X with a martingale problem derived from the martingale problem of $(X^\varepsilon, Y^\varepsilon)$ provided that the solution to this problem is unique. We would then obtain the convergence of X^ε to X in law. However, for multiscale singular problems, the convergence of $L^\varepsilon\varphi((X_t^\varepsilon, Y_t^\varepsilon))$ when $\varepsilon \to 0$, for a test function φ depending only on the X component, does not (in general) take place because it contains diverging terms in ε. To overcome this difficulty, firstly the idea is to decompose the generator L^ε in the following form $L^\varepsilon = \frac{1}{\varepsilon}L_1 + L_0$ where L_1 is the infinitesimal generator of the

MARKOV process Y in the variable y and L_0 the operator of the slow component depending on slow and fast variables. Secondly, the idea is to perturb the initial test function $\varphi(x)$ into a test function $\varphi^\varepsilon(x,y) := \varphi(x) + \varepsilon\varphi_1(x,y)$ such that

$$L^\varepsilon\varphi^\varepsilon = \left[\frac{1}{\varepsilon}L_1 + L_0\right](\varphi + \varepsilon\varphi_1) = \left(L_1\varphi_1 + L_0\varphi - \overline{L}_0\varphi\right) + \overline{L}_0\varphi + \varepsilon L_0\varphi_1 \tag{1}$$

because of $L_1\varphi = 0$ and where $\overline{L}_0$ is the operator of the limit slow component *averaged* by the unique invariant probability measure of the limit fast component Y. Provided that φ_1 solves the POISSON equation (with respect to L_1 and the variable y)

$$L_1\varphi_1 + L_0\varphi - \overline{L}_0\varphi = 0, \tag{2}$$

we deduce that $L^\varepsilon\varphi^\varepsilon = \overline{L}_0\varphi + O(\varepsilon)$. Then, $\varphi^\varepsilon(x,y) \to \varphi(x)$ and $L^\varepsilon\varphi^\varepsilon(x,y) \to \overline{L}_0\varphi(x)$ when $\varepsilon \to 0$ as expected.

The perturbed test function method has been extended in [10, 58, 59] and in the books [44, Section 6.3], [60]. Many other references apply the perturbed test function strategy in various settings: in finance [45] in transport problems as in [75] or in [25] where the tightness of the fast component is established using its occupation measure. Similar methods are used in homogenisation [75, 7, 77] (see also [6] which exploits spectral and semi-group properties in addition), and in stochastic stability and control [60].

Finally, an important method is the stochastic averaging, in a generic framework, that was proposed by KURTZ in [57]. Many approaches [3, 67, 53, 24, 49, 14, 4, 8] including ours, are based on it. The main idea consists in exploiting the *occupation measure* Γ^ε of the fast component Y^ε which is formally defined for all $t \geqslant 0$, for any Borelian B by

$$\Gamma^\varepsilon([0,t]\times B) = \int_0^t \mathbb{1}_B(Y_s^\varepsilon)\,\mathrm{d}s,$$

and to establish the convergence of the couple $(X^\varepsilon, \Gamma^\varepsilon)$ when $\varepsilon \to 0$ using *compactness-uniqueness* arguments. Proceeding in this way allows us to escape the difficulties created by the fluctuations of the fast component and avoids to obtain tightness result of the "slow-fast" couple $(X^\varepsilon, Y^\varepsilon)$. The proof can be divided into four steps.

The first one consists in establishing uniform tightness of the sequence of laws of the couple $(X^\varepsilon, \Gamma^\varepsilon)$. The second one consists in establishing a martingale problem for any limiting value of the family of laws of the couple. The third one is based on establishing the uniqueness of the limiting value of Γ^ε expressed in terms of the limit X of X^ε which is given. The characterisation of the limiting value Γ of Γ^ε is usually based on ergodicity arguments. Finally, the last step is to establish the uniqueness of the martingale problem for the slow component limit X when Γ is given as above.

The KURTZ approach seems to be better adapted to our situation than the perturbed test function method. Indeed, it is delicate to find the good class of test function φ_1 satisfying (2) and the computations are difficult because of the moment terms (see Section 2.3). In addition, the remainder terms generated by (1) are difficult to control and the fast component Y^ε does not take the form $Y(t/\varepsilon)$. However, KURTZ's approach can neither be implemented directly because of difficulties inherent to our model described in Section 2. Therefore, our result required a complete reworking of the classical arguments. In particular, in our case, we do not have uniform moment estimates but only a uniform control, in probability, of the second moment of the fast variable up

to a stopping time $\check{\tau}^K$ where the diameter of the support of the fast component becomes large.

This paper is organised as follows. In Section 2, we define our trait structured IBM, state the central theorem about the CEAD characterising the limit model which consists of an ODE ruling the dynamics of the limit slow component. We establish in this section the sketch and outline of the main proof and the difficulties encountered. We give also some prospects that let us believe that we can improve our main result by relaxing some assumptions. The rest of this paper is devoted to prove the central theorem of Section 2. In Section 3 some approximations of the infinitesimal generator of the slow-fast process are proved. In Section 4, we establish estimates of moments, in particular of the sixth and second order moment up to time $\check{\tau}^K$. We also prove the convergence of $\check{\tau}^K$ to $+\infty$ when $K \to +\infty$ exploiting coupling arguments between the moment of order 2 and a certain biased random walk for which large deviations estimates are established. In Section 5, we prove the tightness of the couple "slow-occupation measure fast" in the torus case. In Section 6, we establish, in the torus case, the uniqueness of the limit occupation measure which is described by the unique invariant probability measure of the centered FLEMING-VIOT process. This result is used in Section 7 to characterise the limit slow component as the unique solution of the CEAD with values in the torus. Thanks to the non-explosion of this ODE and choosing the torus large enough we are able to conclude the proof of our main result given in Section 2.

2 A trait structured IBM

Let us describe the microscopic model which models population evolution, in a Darwinian sense, at the individual level.

2.1 Parameters and assumptions of the model

We consider a discrete population of constant size in continuous time in which the survival and reproductive capacity of each individual is characterised by a continuous quantitative phenotypic trait $x \in \mathbb{R}$, i.e. an overall characteristic subject to selection such as body size at maturity. The individuals reproduce asexually during their lives, i.e. the reproduction scheme is assumed clonal simultaneously with death of another individual, with frequency-dependent rates. Each birth of an individual occurs simultaneously with the death of another individual in the population.

We are interested in approximating the long-term dynamics of a large population. We assume that the number of individuals alive at each time $t \geqslant 0$ is always equal to K. Let us denote by $x_1(t), \cdots, x_K(t)$, the phenotypic trait values of these individuals at time t. A mutation occurs during life at the individual level at rate 1: in this sense mutations are considered *frequent*. Each mutation has an amplitude of the order of magnitude $\sigma_K \in (0, +\infty)$. Small σ_K means mutations have a small phenotypic effect, i.e. evolution acts slowly on the individual phenotypic characteristics. The state of the population at time t, can be described by the finite point measure on $\mathbb{R}$ rescaled by K

$$\nu_t^K := \frac{1}{K} \sum_{i=1}^{K} \delta_{x_i(t)}$$

where δ_x is the DIRAC measure at x.

For all $x, y \in \mathbb{R}$, let us introduce the following biological parameters:

- $b(x,y) \in \mathbb{R}_+$ is the *resampling* rate, i.e. the rate of simultaneous birth of an individual holding trait y and death of an individual holding trait x and it can be interpreted as modeling a competitive pressure driving selection.
- $\theta(x) \in \mathbb{R}_+$ is the rate of mutation of an individual holding trait x.
- $m(x, \mathrm{d}h)$ is the mutation law of the scaled mutation step h, born from an individual with trait x, where the mutant trait is given by $y := x + \sigma_K h \in \mathbb{R}$. It is a probability measure on $\mathbb{R}$.

Let us also introduce the following notations, used throughout this paper:

- $\beta(x) := \frac{\theta(x)}{b(x,x)}$, the ratio between the mutation rate and the resampling rate in a monomorphic population with trait x, which can be interpreted (in the scaling limit considered below) as the mean number of mutations between two resampling events.
- $\mathrm{Fit}(y,x) := b(x,y) - b(y,x)$ is the adaptive value or *fitness* of a mutant individual with trait y in the population of $(K-1)$ individuals of trait x. The fitness can be interpreted as the initial growth rate of a mutant individual y in a resident monomorphic population with trait x.
  Indeed, the total birth rate of a mutant individual y in this population is $b(x,y)(K-1)$ and the total death rate of a mutant individual y in this population is $b(y,x)(K-1)$. Hence, the (initial) growth rate of the mutant population is $\mathrm{Fit}(y,x)(K-1)$.

Let $\mathcal{M}_1(\mathbb{R})$ be the set of probability measures on $\mathbb{R}$, endowed with the topology of weak convergence making it a Polish space [9]. For a measurable real bounded function f, and a measure $\nu \in \mathcal{M}_1(\mathbb{R})$, we denote $\langle f, \nu \rangle := \int_{\mathbb{R}} f(x)\nu(\mathrm{d}x)$. We denote by id the identity function. We denote also $\mathbb{N} := \{0, 1, 2, \cdots\}$ and $\mathbb{N}^\star := \mathbb{N} \setminus \{0\}$. If $\mathcal{I}$ is an interval of $\mathbb{R}$, then for all $\ell \in \mathbb{N}$, $q \in \mathbb{N}^\star$, we denote by $\mathscr{C}^\ell(\mathcal{I}^q, \mathbb{R})$ the space of functions of class $\mathscr{C}^\ell$ from $\mathcal{I}^q$ to $\mathbb{R}$ and by $\mathscr{C}^\ell_b(\mathcal{I}^q, \mathbb{R})$ the space of functions of class $\mathscr{C}^\ell(\mathcal{I}^q, \mathbb{R})$ with bounded derivatives. Finally, we define for all $K \in \mathbb{N}^\star$,

$$\mathcal{M}_{1,K}(\mathbb{R}) := \left\{ \frac{1}{K} \sum_{i=1}^{K} \delta_{x_i} \,\middle|\, (x_i)_{1 \leqslant i \leqslant K} \in \mathbb{R}^K \right\}$$

the set of probability measures on $\mathbb{R}$ of K atoms of mass $1/K$.

Assumptions. Let us denote by **(A)** the following two assumptions:

(A1) The maps $b : \mathbb{R}^2 \to \mathbb{R}_+$, and $\theta : \mathbb{R} \to \mathbb{R}_+$ are respectively in $\mathscr{C}^2_b(\mathbb{R}^2, \mathbb{R})$ and $\mathscr{C}^2_b(\mathbb{R}, \mathbb{R})$ and there exists $0 < \underline{b}, \overline{b}, \underline{\theta}, \overline{\theta} < \infty$ such that:

$$\underline{b} \leqslant b(\cdot,\cdot) \leqslant \overline{b}, \qquad \text{and} \qquad \underline{\theta} \leqslant \theta(\cdot) \leqslant \overline{\theta}.$$

(A2) (a) There exists $A_m \in (0, +\infty)$ such that the law $m(x, \mathrm{d}h)$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$ with density $m(x,h)$ centered and satisfies

$$\forall x \in \mathbb{R},\ \forall |h| \geqslant A_m, \quad m(x,h) = 0.$$

(b) For all $\alpha \in \mathbb{R}$, for all $\ell \in \mathbb{N}$, we denote by

$$m_\ell(\alpha) := \int_{\mathbb{R}} |h|^\ell \, m(\alpha, h) \, \mathrm{d}h$$

the ℓ^{th} moment of the mutation law. We assume for $\ell \in \{2, 4, 6\}$ that m_ℓ is LIPSCHITZ and there exists $\underline{m}_\ell, \overline{m}_\ell \in (0, +\infty)$ such that for all $\alpha \in \mathbb{R}$, $\underline{m}_\ell \leqslant m_\ell(\alpha) \leqslant \overline{m}_\ell$.

Let us now give the infinitesimal generator L^K of the $\mathcal{M}_{1,K}(\mathbb{R})-$valued MARKOV process $\nu^K := \left(\nu_t^K\right)_{t \geqslant 0}$ describing the ecological dynamics of the population with resampling. The generator L^K is defined for any bounded measurable map ϕ from $\mathcal{M}_{1,K}(\mathbb{R})$ to $\mathbb{R}$, by

$$\begin{aligned} L^K \phi(\nu) = K \int_{\mathbb{R}} \nu(\mathrm{d}x) \int_{\mathbb{R}} \nu(\mathrm{d}y) b(x,y) \left[\phi\left(\nu - \frac{\delta_x}{K} + \frac{\delta_y}{K}\right) - \phi(\nu) \right] \\ + K \int_{\mathbb{R}} \theta(x) \nu(\mathrm{d}x) \int_{\mathbb{R}} m(x,h) \left[\phi\left(\nu - \frac{\delta_x}{K} + \frac{\delta_{x+\sigma_K h}}{K}\right) - \phi(\nu) \right] \mathrm{d}h. \end{aligned} \tag{3}$$

The first term describes the resampling event of one individual by another and the second term describes the effect of mutations over the lifetime of individuals.

2.2 Main result

The main result of this article is the following theorem. For all $\nu \in \mathcal{M}_1(\mathbb{R})$, let us denote by $\mathrm{Diam}\left(\mathrm{Supp}\, \nu\right)$ the diameter of the support of ν.

Theorem 2.1. *Assume* **(A)** *and for all* $K \in \mathbb{N}^\star$, $\sigma_K \in (0, +\infty)$, *for some* $x_0 \in \mathbb{R}$, *for all* $\ell \in \{1, 2, 3\}$,

$$\left\langle \mathrm{id}, \nu_0^K \right\rangle := x_0, \qquad \text{and} \qquad \left\langle (\mathrm{id} - x_0)^{2\ell}, \nu_0^K \right\rangle \leqslant C_{2\ell}^\star K^\ell \sigma_K^{2\ell} \tag{4}$$

for some $C_{2\ell}^\star > 0$. *Assume that* $K \to +\infty$, $\sigma_K \to 0$ *such that*

$$\exists\, \varepsilon > 0, \quad \forall C > 0, \qquad K^{-C\log(K)} \ll \sigma_K \ll K^{-(2+\varepsilon)}. \tag{5}$$

Then, for the SKOROHOD *topology, the sequence of the mean trait processes*

$$\left(\left(\left\langle \mathrm{id}, \nu_{t/K\sigma_K^2}^K \right\rangle \right)_{t \geqslant 0} \right)_{K \in \mathbb{N}^\star}$$

converges in law in $\mathbb{D}\left(\mathbb{R}_+, \mathbb{R}\right)$ *when* $K \to +\infty$, $\sigma_K \to 0$, *to the unique deterministic process* $(\zeta_t)_{t \geqslant 0} \in \mathscr{C}^0\left(\mathbb{R}_+, \mathbb{R}\right)$ *with initial condition* $\zeta_0 := x_0$, *which is solution to the* Canonical Equation of Adaptive Dynamics (CEAD)*:*

$$\dot{y} = \partial_1 \mathrm{Fit}(y, y) \times \frac{\beta(y) m_2(y)}{2}.$$

In addition, we have the following support concentration property : for all $T > 0$,

$$\sup_{t \in [0,T]} \mathrm{Diam}\left(\mathrm{Supp}\, \nu_{t/K\sigma_K^2}^K\right) \leqslant \frac{1}{K} \tag{6}$$

holds with probability which tends to 1 *when* $K \to +\infty$.

The canonical equation is composed of two terms. The first term: the fitness gradient, describes the strength of selection that pushes the population to increase its adaptive value locally. The second term describes the effect of mutations. Note that (6) describes, in some sense, when $K \to +\infty$, the convergence of the population distribution $\left(\nu^K_{t/K\sigma_K^2}\right)$ to a DIRAC mass.

Remark 2.2. (i) *We conjecture that the assumption* $\sigma_K \ll K^{-(2+\varepsilon)}$ *given by* (5) *is too restrictive and that it can be weakened by* $\sigma_K \ll K^{-(\frac{3}{2}+\varepsilon)}$. *This conjecture is supported by the discussion in* Section 2.3.5 *below.*

(ii) *We actually prove* Theorem 2.1 *(see* Section 4.4.3*) under a slightly weaker assumption than the left bound of* (5)*: for a universal constant* C_0 *depending only on* b, θ, m *and not on* K*,*

$$\exists\, \varepsilon > 0, \qquad K^{-\varepsilon^2 C_0 \log(K)} \ll \sigma_K \ll K^{-(2+\varepsilon)}.$$

(iii) *Assumption* **(A2)(a)** *is only used in* Section 4.4 *and could be weakened with some tail bounds (typically exponential) on* $m(x,\cdot)$.

The aim of the rest of the paper is to prove Theorem 2.1.

2.3 Sketch and outline of the proof

Note that, the convergence result of Theorem 2.1 takes place on the time scale $1/K\sigma_K^2$. The idea of the proof is based on slow-fast asymptotic analysis techniques developed by KURTZ [57]. Let us begin by introducing the slow and fast dynamics involved in our model, then the difficulties encountered and finally the outline of the proof.

2.3.1 Slow-fast asymptotic analysis

Let $a, \alpha \in \mathbb{R}$ and $\mathcal{B}(\mathbb{R})$ the BOREL $\sigma-$field on $\mathbb{R}$. Let us define respectively by τ_α and h_a, the translation operator of vector α and the homothety of ratio a. For all $x \in \mathbb{R}$, for all $A \in \mathcal{B}(\mathbb{R})$,

$$(h_a \circ \tau_\alpha) \sharp \delta_x(A) = \delta_{a(x+\alpha)}(A) \qquad \text{and} \qquad (\tau_\alpha \circ h_a) \sharp \delta_x(A) = \delta_{ax+\alpha}(A).$$

Note that for all $K \in \mathbb{N}^\star$, any $\nu \in \mathcal{M}_{1,K}(\mathbb{R})$ has all its moments finite. We define for all $K \in \mathbb{N}^\star$, $\mathcal{M}^c_{1,K}(\mathbb{R})$ the set of centered probability measures of $\mathcal{M}_{1,K}(\mathbb{R})$. From the population process $\left(\nu^K_{t/K\sigma_K^2}\right)_{t\geqslant 0}$, we introduce two evolutionary dynamics:

- The slow dynamics, with value in $\mathbb{R}$, corresponds to that of the mean trait defined by
$$z^K_t := \left\langle \mathrm{id}, \nu^K_{t/K\sigma_K^2} \right\rangle.$$
- The fast dynamics, with value in $\mathcal{M}^c_{1,K}(\mathbb{R})$, corresponds to that of the centered and dilated distribution of traits defined by
$$\mu^K_t := \left(h_{\frac{1}{\sigma_K\sqrt{K}}} \circ \tau_{-z^K_t}\right) \sharp\, \nu^K_{t/K\sigma_K^2} = \frac{1}{K}\sum_{i=1}^K \delta_{\frac{1}{\sigma_K\sqrt{K}}(x_i(t)-z^K_t)}. \tag{7}$$

Note that $\nu^K_t = \left(\tau_{z^K_{tK\sigma_K^2}} \circ h_{\sigma_K\sqrt{K}}\right) \sharp\, \mu^K_{tK\sigma_K^2}$.

For all $\ell \in \mathbb{N}$ and $\nu \in \mathcal{M}_1(\mathbb{R})$, we denote by $M_\ell(\nu) := \left\langle |\mathrm{id}|^\ell, \nu \right\rangle$ the ℓ^{th} moment of the measure ν. Let us consider for all $\ell \in \mathbb{N}^\star$,

$$\mathcal{M}_1^{c,\ell}(\mathbb{R}) := \left\{ \mu \in \mathcal{M}_1(\mathbb{R}) \;\middle|\; M_\ell(\mu) < \infty, \langle \mathrm{id}, \mu \rangle = 0 \right\}$$

and the convention $M_\ell(\mu) = +\infty$ if $\mu \notin \mathcal{M}_1^{c,\ell}(\mathbb{R})$. Let us denote by $\mathcal{B}(E, \mathbb{R})$ the set of BOREL functions on the metric space E.

In the following, for all $K \in \mathbb{N}^\star$, we study the couple of processes $\left(\left(z_t^K, \mu_t^K\right)\right)_{t \geqslant 0}$. Denoting $\mathcal{L}^K$ the infinitesimal generator of the process $\left(\left(z_t^K, \mu_t^K\right)\right)_{t \geqslant 0}$, computed in Proposition 3.1, we will check the following assertions, proved respectively in Propositions 3.2 and 3.3.

(i) If we consider a function $f : (z, \mu) \mapsto f(z) \in \mathscr{C}_b^2(\mathbb{R}, \mathbb{R})$, the generator $\mathcal{L}^K$ satisfies the following decomposition

$$\mathcal{L}^K f(z, \mu) = \mathcal{L}_{\text{SLOW}} f(z, \mu) + O\left(\frac{1}{K^2} + \frac{M_2(\mu)}{K} + \sigma_K \sqrt{K} M_3(\mu)\right) \tag{8}$$

where the operator $\mathcal{L}_{\text{SLOW}}$, of the limit slow component, is defined from $\mathscr{C}_b^1(\mathbb{R}, \mathbb{R})$ to $\mathcal{B}\left(\mathbb{R} \times \mathcal{M}_1^{c,2}(\mathbb{R}), \mathbb{R}\right)$ by

$$\mathcal{L}_{\text{SLOW}} f(z, \mu) = f'(z) M_2(\mu) \, \partial_1 \mathrm{Fit}(z, z). \tag{9}$$

(ii) If we consider a function $F_\varphi : (z, \mu) \mapsto F_\varphi(\mu) := F(\langle \varphi, \mu \rangle)$ with $F \in \mathscr{C}^3(\mathbb{R}, \mathbb{R})$ and $\varphi \in \mathscr{C}_b^3(\mathbb{R}, \mathbb{R})$, the generator $\mathcal{L}^K$ satisfies the following decomposition

$$\mathcal{L}^K F_\varphi(z, \mu) = \frac{\theta(z) m_2(z)}{K^2 \sigma_K^2} \left[\mathcal{L}_{\text{FVc}}^{\lambda(z)} F_\varphi(z, \mu) + O\left(\frac{1}{\sqrt{K}} + \sigma_K K^{\frac{3}{2}} M_2(\mu) + \frac{M_3(\mu)}{K}\right)\right] \tag{10}$$

where $\mathcal{L}_{\text{FVc}}^{\lambda(z)}$ is the generator of the *centered* FLEMING-VIOT *process* with resampling rate $\lambda(z) := \frac{b(z,z)}{\theta(z) m_2(z)}$, as studied in [18] and whose definition is recalled below. For all $F \in \mathscr{C}^2(\mathbb{R}, \mathbb{R})$ and $g \in \mathscr{C}_b^2(\mathbb{R}, \mathbb{R})$,

$$\begin{aligned} \mathcal{L}_{\text{FVc}}^{\lambda} F_g(\mu) := {} & F'(\langle g, \mu \rangle) \left(\left\langle \frac{g''}{2}, \mu \right\rangle + \lambda \left[\langle g'', \mu \rangle M_2(\mu) - 2 \langle g' \times \mathrm{id}, \mu \rangle \right] \right) \\ & + \lambda F''(\langle g, \mu \rangle) \left[\langle g^2, \mu \rangle - \langle g, \mu \rangle^2 + \langle g', \mu \rangle^2 M_2(\mu) - 2 \langle g', \mu \rangle \langle g \times \mathrm{id}, \mu \rangle \right]. \end{aligned} \tag{11}$$

From (10) and (8), note that the fast component $\left(\mu^K_{K\sigma_K^2 t}\right)_{t \geqslant 0} = \left(\left(h_{\frac{1}{\sigma_K \sqrt{K}}} \circ \tau_{-\langle \mathrm{id}, \nu_t^K \rangle}\right) \sharp \, \nu_t^K\right)_{t \geqslant 0}$ moves on the evolutionary time scale K whereas the slow component $\left(z^K_{K\sigma_K^2 t}\right)_{t \geqslant 0} = \left(\langle \mathrm{id}, \nu_t^K \rangle\right)_{t \geqslant 0}$ moves over a much longer evolutionary time scale $1/K\sigma_K^2$. The different time scales involved in this model can be represented as in Figure 2.1. It follows that the fast component will be the first to stabilise in its equilibrium state. Note that the operator of the slow component (9) depends on the fast variable through the second-order moment $M_2(\mu)$. This is a standard difficulty in slow-fast analysis, usually solved by assuming that once the fast component is stabilised in its

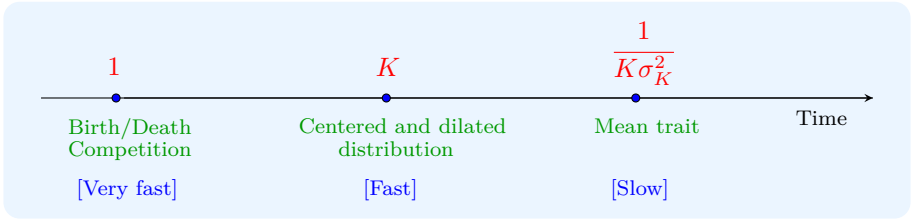

Figure 2.1: Different time scales involved in our individual-based model expected in the original time scale of the process $\left(\nu_t^K\right)_{t\geqslant 0}$.

equilibrium state, $M_2(\mu)$ can be characterised in terms of the slow component leading to an autonomous slow dynamics in the limit $K \to +\infty$.

As the fast limit component is driven by a centered Fleming-Viot process, we expect that it will inherit ergodicity properties as stated in [18, Section 4]. This reference establishes the existence of a unique invariant probability measure π^λ for the centered Fleming-Viot process and characterises it. In particular, we expect that the averaging principle applied to the drift coefficient in (9) will lead to the averaged drift

$$\partial_1 \mathrm{Fit}(z_t, z_t) \int_{\mathcal{M}_1(\mathbb{R})} M_2(\mu)\, \pi^{\lambda(z_t)}(\mathrm{d}\mu) = \frac{\partial_1 \mathrm{Fit}(z_t, z_t)}{2\lambda(z_t)} = \partial_1 \mathrm{Fit}(z_t, z_t)\, \frac{\beta(z_t)\, m_2(z_t)}{2}$$

where $(z_t)_{t\geqslant 0}$ is the limit slow component (see [18, Corollary 4.16] for the computation of the integral). Formally, by replacing $M_2(\mu)$ by its mean value $\int_{\mathcal{M}_1(\mathbb{R})} M_2(\mu)\, \pi^{\lambda(z_t)}(\mathrm{d}\mu)$ in (9), we obtain that (see details in Section 7)

$$\mathcal{L}_{\mathrm{CEAD}} f(z,\mu) = f'(z)\partial_1 \mathrm{Fit}(z,z)\frac{\beta(z)m_2(z)}{2}$$

and we recognise the generator of the announced CEAD.

2.3.2 An example

The goal of this section is to illustrate numerically the phenomenon of the CEAD. We have performed numerical simulations of the process ν^K for a simple model from the biological literature. The simulation is based on the `IBMPopSim` package developed by Giorgi et al in [50].

The biological model is adapted from Kisdi [56] for which the trait space is $\mathcal{X} := [0,4]$ (and not $\mathbb{R}$ as previously) and the parameters are:

$$\theta(x) := 1, \quad c(x,y) := 1 - \frac{1}{1 + 1.2\exp(-4(x-y))}, \quad b(x,y) := (4-y) \times c(x,y) \tag{12}$$

and $m(x, \mathrm{d}h)$ is a Gaussian distribution $\mathcal{N}\left(x, \sigma_K^2\right)$ conditional on the mutant trait remaining in $\mathcal{X}$.

Let us recall that $b(x,y)$ is a resampling rate, i.e. the rate of simultaneous death of an individual with trait x and birth of an individual with trait y and it can be interpreted as

modelling a competitive pressure driving selection. Here, it is composed of two terms which both have a biological interpretation. Note that the first term $c(x, y)$ depends only on $x - y$ and tends to 0 when $x - y \to +\infty$ and to 1 when $x - y \to -\infty$. So, this function models an *asymmetric competition* in favour of high traits: x feels little competition from a smaller trait y, but strong competition from a higher trait y. Since the second term $4 - y$ only depends on the reproducing trait and is favourable to small traits, evolution is expected to select an optimal trait.

We consider the previous model with three scaling parameters: K (for the population size) and σ_K (for the size of mutation) as in Section 2.1, and a parameter p_K giving the probability of mutation at each birth event. This amouts to replace in Section 2.1 the mutation kernel $m(x, h)\mathrm{d}h$ by $p_K m(x, h)\mathrm{d}h + (1 - p_K)\delta_x(\mathrm{d}h)$. We performed simulations of Kisdi's model, starting with a monomorphic initial condition of trait 1 and varying the parameters K, p_K and σ_K. Some of these simulations are presented in the following figures and show a wide variety of evolutionary behaviours.

Note that, when K and σ_K are not too large and p_K not too small, the population evolves according to a relatively stable general scenario as in Figures 2.3, 2.4 and 2.5: it initially remains concentrated around the trait value equal to 1 that progressively moves towards a trait $x^\star$ close to 3.5 corresponding to the optimal trait value of this population, in the sense that it is best adapted to survive and reproduce. Once this point is reached, the population stabilises in its steady state. In other words, Figures 2.4 and 2.5 show the long-term behaviour of the mean trait of the population, i.e. the one of the slow component. Figure 2.6 shows the behaviour of the fast component by zooming in the initial dynamics of Figure 2.5. Note that Figure 2.2 illustrates the approach developed by [20].

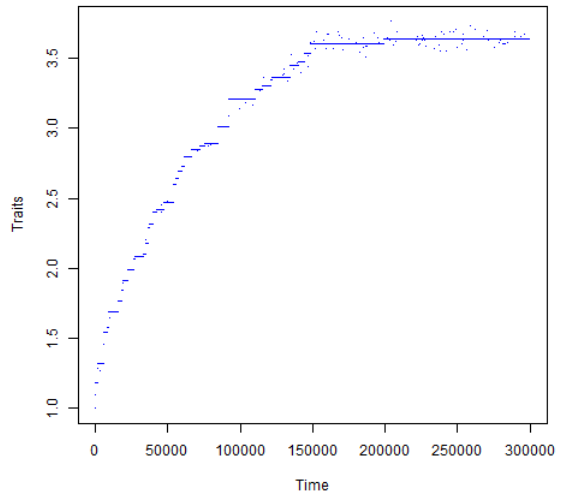

Figure 2.2: Numerical simulation of the trait distribution of the microscopic model with parameters (12) and $K = 300$, $p_K = 0.00003$, $\sigma_K = 0.05$.

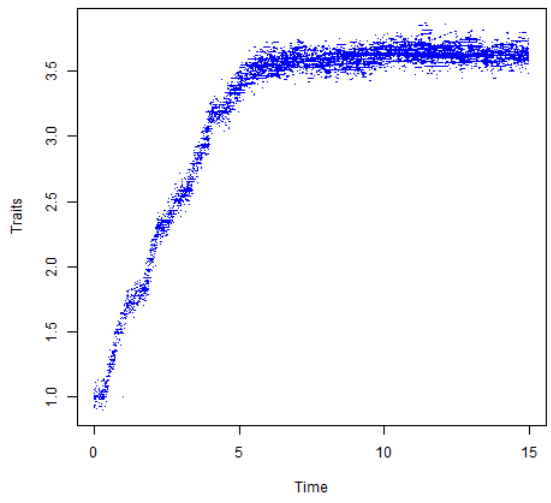

Figure 2.3: Numerical simulation of the trait distribution of the microscopic model with parameters (12) and $K = 300$, $p_K = 1$, $\sigma_K = 0.05$.

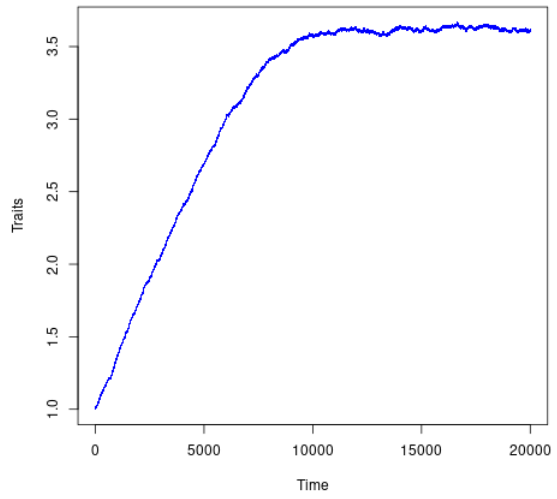

Figure 2.4: Numerical simulation of the trait distribution of the microscopic model with parameters (12) and $K = 300$, $p_K = 1$, $\sigma_K = 0.001$.

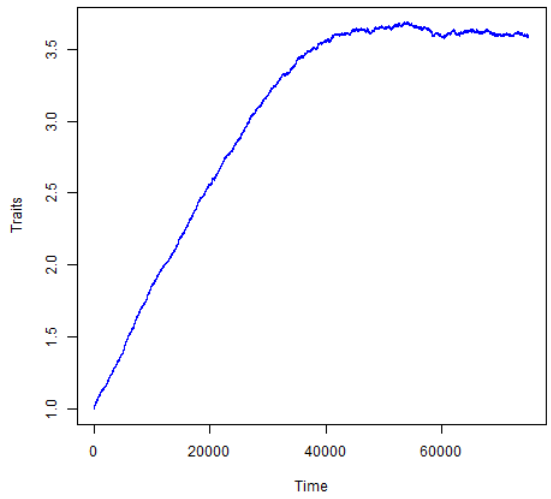

Figure 2.5: Numerical simulation of the trait distribution of the microscopic model with parameters (12) and $K = 300$, $p_K = 1$, $\sigma_K = 0.0005$.

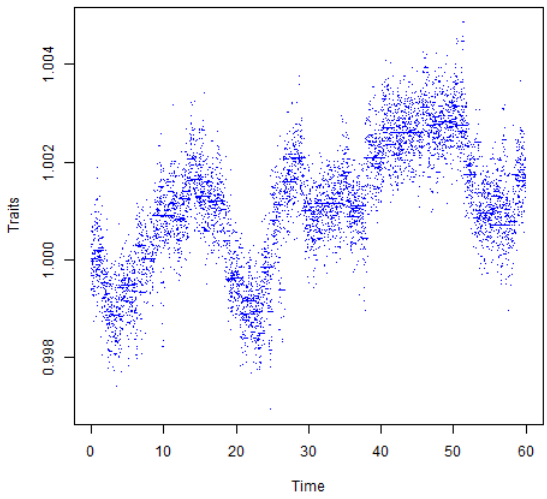

Figure 2.6: Numerical simulation of the trait distribution of the microscopic model with parameters (12) and $K = 300$, $p_K = 1$, $\sigma_K = 0.0005$.

2.3.3 Difficulties encountered

The slow-fast analysis relies on a stochastic averaging result that exploits tightness arguments. The classical approach to prove tightness of sequences of laws [5, 46] requires to have uniform pathwise estimates on the moments of the process. In our situation, we note the presence of moments in (8) and (10). However, we could not establish control in expectation of $\sup_{0\leqslant t\leqslant T} M_2\left(\mu_t^K\right)$ for all $T \geqslant 0$ because of the long time scale which does not allow to exploit the martingale problem associated to the decomposition of $M_2\left(\mu_t^K\right)$. This is an important difference, for example with [67] which studies similar forms of processes (individual-based models). Instead, we will first use fine pathwise estimates and expectation bounds on $M_2\left(\mu_t^K\right)$ up to the stopping time τ^K defined for all $K \in \mathbb{N}^\star$ by

$$\tau^K := \check{\tau}^K \wedge \widehat{\tau}^K. \tag{13}$$

where

$$\widehat{\tau}^K := \inf\left\{t \geqslant 0 \,\middle|\, M_2\left(\mu_t^K\right) \geqslant K^\varepsilon\right\} \tag{14}$$

with $\varepsilon > 0$ given in Theorem 2.1 and where

$$\check{\tau}^K := \inf\left\{t \geqslant 0 \,\middle|\, \mathrm{Diam}\left(\mathrm{Supp}\,\mu_t^K\right) > \frac{1}{\sigma_K K^{\frac{3+\varepsilon}{2}}}\right\} \tag{15}$$

which allows to control the diameter of the support of the centered and dilated distribution of the traits μ^K. An important difficulty arises in Lemma 3.6 from the presence of third order moment in the Doob semi-martingale decomposition of $M_2\left(\mu_t^K\right)$:

$$\begin{aligned} M_2\left(\mu_t^K\right) = M_2\left(\mu_0^K\right) &- \frac{1}{K^2\sigma_K^2}\int_0^t \left\{2b\left(z_s^K, z_s^K\right) M_2\left(\mu_s^K\right) - \theta\left(z_s^K\right) m_2\left(z_s^K\right)\right\} \mathrm{d}s \\ &+ \frac{1}{K^2\sigma_K^2}\int_0^t O\left(\frac{1}{K} + \sigma_K K M_2\left(\mu_s^K\right) + \sigma_K K^{\frac{3}{2}} M_3\left(\mu_s^K\right)\right) \mathrm{d}s + M_t^{K,P_{\mathrm{id}^2,1}} \end{aligned}$$

where $\left(M_t^{K,P_{\mathrm{id}^2,1}}\right)_{t\geqslant 0}$ is a local-martingale. This leads to introduce the stopping time $\check{\tau}^K$ which guarantees that the error term in $M_3\left(\mu_t^K\right)$ remains under control. This leads to a new difficulty: establishing that $\check{\tau}^K$ tends to $+\infty$ in probability when $K \to +\infty$. To prove this, we use pathwise estimates on $M_2\left(\mu_t^K\right)$ up to time $\check{\tau}^K$ to estimate the different transitions of $M_2\left(\mu_t^K\right)$ between thresholds of the form $3^{\ell-1}K^{\frac{\varepsilon}{2}}$ and $3^{\ell+1}K^{\frac{\varepsilon}{2}}$ and to construct a coupling between these transitions and biased random walks. This allows us to use large deviations results on random walks and estimates on the exit from an attracting domain (see e.g. [47, 29]) to prove that the stopping time $\widehat{\tau}^K$, defined by (14), converges to $+\infty$ in probability when $K \to +\infty$. This in turn implies that $\check{\tau}^K \to +\infty$ in probability when $K \to +\infty$ (see Section 2.3.5 for a discussion of the method used for this step).

The implementation of the slow-fast method of Kurtz [57] is done in two steps. To establish the tightness of the sequence of laws of the slow component, we exploit criteria developed by Ethier-Kurtz [40, Theorems 3.9.1 and 3.9.4] by restricting ourselves to the torus case. This strategy allows to overcome the difficulty related to the verification of the compact containment condition, on the real line, of the slow component stopped at the stopping time τ^K. The second

step consists in characterising each accumulation point of the sequence of laws of the slow dynamics and the occupation measure of the fast dynamics. For this, we need to check that any measure γ in $\mathcal{M}_1(\mathcal{M}_1(\mathbb{R}))$ satisfying

$$\int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}^{\lambda}_{\mathrm{FVc}}\phi(\mu)\gamma(\mathrm{d}\mu) = 0 \tag{16}$$

for a certain class $\mathcal{F}$ of functions ϕ must be π^{λ}. Equation (10) suggests to take $\mathcal{F}$ as the set of functions of the form $F_{\varphi}(\mu) := F(\langle\varphi,\mu\rangle)$. However, the last property seems hard to prove for this choice of $\mathcal{F}$. Instead, we adapt a result of Dawson [26, Theorem 2.7.1] that applies to the set $\mathcal{F}$ of so-called *polynomial functions* of the form

$$P_{f,n}(\mu) := \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, \cdots, x_n)\, \mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_n).$$

However the duality property used by Dawson in [26] does not hold in our case. In [18], it is proved only a weak duality relation involving stopping times. This difficulty will be solved by proving that the measure γ in (16) gives mass only to measures having its first four moments finite. This will allow us to characterise the limit fast component and hence the limit slow component on the torus, as solution to an ODE. Since this ODE is non-explosive, we choose the torus large enough to conclude the proof on the real line.

2.3.4 Outline of the proof

In Section 3, we begin by giving in Proposition 3.1 an approximation of the infinitesimal generator of the slow-fast process $\left(\left(z_t^K, \mu_t^K\right)\right)_{t\geqslant 0}$ for a class of test functions on $\mathbb{R}\times\mathcal{M}^c_{1,K}(\mathbb{R})$, large enough to be convergence determining and we characterise martingales associated to our process. We give also martingale problems of the slow-fast process for polynomials in μ where the main results are given in Lemma 3.4 for bounded test functions and Lemmas 3.6 and 3.7 for unbounded test functions. In Section 4, we prove some moment estimates. In Section 4.1, we give estimates of the moment of order 6. The most important result is given by Lemma 4.2 because consequences are useful for all results in Section 4 as well as in Theorem 5.1. In Section 4.2, we establish some inequalities on $M_2\left(\mu_t^K\right)$ and we control its bracket that we will use in Section 4.3 to prove $\mathbb{P}-$a.s. that $M_2\left(\mu_t^K\right)$ takes superlinear (of the order of $K^{C\log(K)}$) long time before hitting K^{ε}. The fact that $\tau^K \to +\infty$ in probability when $K \to +\infty$ is proved in Section 4.3 using the technical Lemma 4.9. Section 4.4 is dedicated to prove this lemma by constructing a coupling between the moment of order 2 and a biased random walk. The rest of the proof deals with the compactness-uniqueness argument associated to our slow-fast problem. Firstly, in Section 5, we establish uniform tightness of the sequence of laws of $\left(\left(z^K, \Gamma^K\right)\right)_{K\in\mathbb{N}^\star}$ stopped at time $\check{\tau}^K$ in the torus case where Γ^K designates the occupation measure of the process μ^K. The main result of section is given by Theorem 5.1. In Proposition 6.3 of Section 6, we identify and characterise in a unique way the limiting distribution of the fast component on the torus. Section 6.1 is devoted to proving this result by exploiting the key Lemma 6.4 whose proof is given in Section 6.2. In Lemma 7.1 of Section 7.1, we proceed similarly for the slow component. Thanks to the uniqueness on the torus of the limit slow component, we deduce in Section 7.2 the announced result of Theorem 2.1.

2.3.5 Prospects

In the first part of this section, we explain the origin of exponent 2 in the inequality $\sigma_K \ll K^{-(2+\varepsilon)}$ of Assumption (5). In the second part, it is explained how Assumption (5) may be improved into the assumption $\sigma_K \ll K^{-\left(\frac{3}{2}+\varepsilon\right)}$ by using estimates for moments of order 2ℓ. Finally, in the last part of this section, we explain the difficulties in obtaining these moment estimates. Note that the assumption $\sigma_K \ll K^{-\frac{3}{2}}$ is the best we can expect because of the error term $\sigma_K K^{\frac{3}{2}}$ in (10).

(a) Origin of the exponent 2 in Assumption (5). On the one hand, as for all $t \geqslant 0$ and $K \in \mathbb{N}^\star$, $\mathrm{Diam}\left(\mathrm{Supp}\,\mu_t^K\right)^2 \leqslant \left(2\max\left\{|x| \,\middle|\, x \in \mathrm{Supp}\,\mu_t^K\right\}\right)^2 \leqslant 4KM_2\left(\mu_t^K\right)$, we deduce that

$$\begin{aligned}\mathbb{P}\left(\check{\tau}^K < \widehat{\tau}^K \wedge T\right) &\leqslant \mathbb{P}\left(\exists t < T \wedge \widehat{\tau}^K, \mathrm{Diam}\left(\mathrm{Supp}\,\mu_t^K\right) > \frac{1}{\sigma_K K^{\frac{3+\varepsilon}{2}}}\right) \\ &\leqslant \mathbb{P}\left(\exists t < T \wedge \widehat{\tau}^K, M_2\left(\mu_t^K\right) > \frac{1}{4\sigma_K^2 K^{4+\varepsilon}}\right)\end{aligned}$$

On the other hand, note that thanks to the definition (14) of the stopping time $\widehat{\tau}^K$,

$$\mathbb{P}\left(\exists t < T \wedge \widehat{\tau}^K, M_2\left(\mu_t^K\right) > K^\varepsilon\right) = 0.$$

Hence, Assumption (5) implies that $\mathbb{P}\left(\check{\tau}^K < \widehat{\tau}^K \wedge T\right) = 0$ for K large enough. Using this relation, we prove in Section 4.3 that $\mathbb{P}\left(\tau^K < T\right) \leqslant \mathbb{P}\left(\widehat{\tau}^K \leqslant T \wedge \check{\tau}^K\right)$. Hence, to establish that $\tau^K \to +\infty$ when $K \to +\infty$ in probability, we need to prove that

$$\lim_{K\to+\infty} \mathbb{P}\left(\widehat{\tau}^K \leqslant T \wedge \check{\tau}^K\right) = \lim_{K\to+\infty} \mathbb{P}\left(\sup_{0\leqslant t\leqslant T\wedge\check{\tau}^K} M_2\left(\mu_t^K\right) \geqslant K^\varepsilon\right) = 0$$

as established in Section 4.4.

(b) Improvement of Assumption (5) with moments of order 2ℓ. Assume now that $\sigma_K \ll K^{-\left(\frac{3}{2}+\varepsilon\right)}$. Then, there exists $\ell \in \mathbb{N}^\star$, $\widetilde{\varepsilon} > 0$ such that

$$\forall C > 0, \qquad K^{-C\log(K)} \ll \sigma_K \ll K^{-\left(\frac{3}{2}+\frac{1}{2\ell}+\frac{\ell+1}{2\ell}\widetilde{\varepsilon}\right)} \tag{17}$$

and we consider the stopping times

$$\widehat{\tau}_{\mathrm{bis}}^K := \inf\left\{t \geqslant 0 \,\middle|\, M_2\left(\mu_t^K\right) \geqslant K^{\widetilde{\varepsilon}}\right\}$$

and

$$\check{\tau}_{\mathrm{bis}}^K := \inf\left\{t \geqslant 0 \,\middle|\, \mathrm{Diam}\left(\mathrm{Supp}\,\mu_t^K\right) > \frac{1}{\sigma_K K^{\frac{3+\widetilde{\varepsilon}}{2}}}\right\}.$$

Using now that for all $t \geqslant 0$, for all $K \in \mathbb{N}^\star$, for all $\ell \in \mathbb{N}^\star$, $\mathrm{Diam}\left(\mathrm{Supp}\,\mu_t^K\right)^{2\ell} \leqslant 2^{2\ell} K M_{2\ell}\left(\mu_t^K\right)$, we deduce that

$$\begin{aligned}\mathbb{P}\left(\exists t < T \wedge \widehat{\tau}_{\mathrm{bis}}^K, \mathrm{Diam}\left(\mathrm{Supp}\,\mu_t^K\right) > \frac{1}{\sigma_K K^{\frac{3+\widetilde{\varepsilon}}{2}}}\right) \\ \leqslant \mathbb{P}\left(\exists t < T \wedge \widehat{\tau}_{\mathrm{bis}}^K, M_{2\ell}\left(\mu_t^K\right) > \frac{1}{2^{2\ell}\sigma_K^{2\ell} K^{3\ell+1+\ell\widetilde{\varepsilon}}}\right),\end{aligned}$$

which is zero for K large enough by (17). As previously, if we can prove for all $T \geqslant 0$ that

$$\lim_{K\to+\infty} \mathbb{P}\left(\widehat{\tau}_{\text{bis}}^K \leqslant T \wedge \check{\tau}_{\text{bis}}^K\right) = \lim_{K\to+\infty} \mathbb{P}\left(\sup_{0\leqslant t\leqslant T\wedge\check{\tau}_{\text{bis}}^K} M_{2\ell}\left(\mu_t^K\right) \geqslant K^{\widetilde{\varepsilon}}\right) = 0, \tag{18}$$

then, we can conclude as before.

(c) Difficulties encountered. The main difficulty consists in proving (18). For this we could seek for an extension of Lemma 4.2 to the framework of moments of order 2ℓ. Lemma 4.2 relies on the inequality for all $t \geqslant 0$, K large enough,

$$\begin{aligned}
&\frac{3}{4}M_6\left(\mu_{t\wedge\check{\tau}^K}\right) + 3M_4\left(\mu_{t\wedge\check{\tau}^K}^K\right) M_2\left(\mu_{t\wedge\check{\tau}^K}^K\right) + \frac{3}{2}M_2^3\left(\mu_{t\wedge\check{\tau}^K}^K\right) \\
&\leqslant \frac{3}{4}M_6\left(\mu_0^K\right) + 7M_4\left(\mu_0^K\right) M_2\left(\mu_0^K\right) + \frac{3}{2}M_2^3\left(\mu_{t\wedge\check{\tau}^K}\right) \\
&- \frac{C_1}{K^2\sigma_K^2}\int_0^{t\wedge\check{\tau}^K}\left(\frac{3}{4}M_6\left(\mu_{t\wedge\check{\tau}^K}\right) + 3M_4\left(\mu_s^K\right) M_2\left(\mu_s^K\right) + \frac{3}{2}M_2^3\left(\mu_s^K\right) - C_2\right)\mathrm{d}s + \mathrm{Mart}_{t\wedge\check{\tau}^K}
\end{aligned}$$

for some constants $C_1, C_2 > 0$ and where $(\mathrm{Mart}_{t\wedge\check{\tau}^K})_{t\geqslant 0}$ is a martingale. However a similar calculation does not seem to be successful for moments of order 8 or more.

3 Infinitesimal generator approximation

Let us begin this section by giving the generator of the couple of process $\left(z^K, \mu^K\right)$. Then, we give an approximation of the previous formula for a class of test functions which is convergence determining. Finally, we give an extension to this result in the case of polynomials in μ.

3.1 Generators and martingales

For all $K \in \mathbb{N}^\star$, let us introduce the filtration $\left(\mathcal{F}_t^K\right)_{t\geqslant 0}$ defined by $\mathcal{F}_t^K := \sigma\left(z_s^K, \mu_s^K \,\middle|\, s \leqslant t\right)$.

Proposition 3.1. *The infinitesimal generator of the $\mathbb{R}\times\mathcal{M}_{1,K}^c(\mathbb{R})-$valued Markov process $\left(z^K, \mu^K\right)$, defined for all bounded measurable function Φ from $\mathbb{R}\times\mathcal{M}_{1,K}^c(\mathbb{R})$ to $\mathbb{R}$, is given by*

$$\begin{aligned}
\mathcal{L}^K\Phi(z,\mu) = &\frac{K}{K\sigma_K^2}\int_{\mathbb{R}}\mu(\mathrm{d}x)\int_{\mathbb{R}}\mu(\mathrm{d}y)b\left(\sigma_K\sqrt{K}x+z, \sigma_K\sqrt{K}y+z\right) \\
&\times\left[\Phi\left(z+\frac{\sigma_K\sqrt{K}}{K}(y-x), \tau_{-\frac{y-x}{K}}\sharp\left[\mu-\frac{\delta_x}{K}+\frac{\delta_y}{K}\right]\right)-\Phi(z,\mu)\right] \\
&+\frac{K}{K\sigma_K^2}\int_{\mathbb{R}}\mu(\mathrm{d}x)\theta\left(\sigma_K\sqrt{K}x+z\right)\int_{\mathbb{R}}m\left(\sigma_K\sqrt{K}x+z,h\right) \\
&\times\left[\Phi\left(z+\frac{\sigma_K\sqrt{K}h}{K^{\frac{3}{2}}}, \tau_{-\frac{h}{K^{3/2}}}\sharp\left[\mu-\frac{\delta_x}{K}+\frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K}\right]\right)-\Phi(z,\mu)\right]\mathrm{d}h.
\end{aligned}$$

Moreover, for all $K \in \mathbb{N}^\star$, for all bounded measurable function Φ from $\mathbb{R}\times\mathcal{M}_{1,K}^c(\mathbb{R})$ to $\mathbb{R}$, the following process $\left(M_t^{K,\Phi}\right)_{t\geqslant 0}$ defined by

$$M_t^{K,\Phi} := \Phi\left(z_t^K, \mu_t^K\right) - \Phi\left(z_0^K, \mu_0^K\right) - \int_0^t \mathcal{L}^K\Phi\left(z_s^K, \mu_s^K\right)\mathrm{d}s \tag{19}$$

is a $\left(\mathcal{F}_t^K\right)_{t\geqslant 0}-$*martingale, square integrable, with quadratic variation:*

$$\begin{aligned}\left\langle M_t^{K,\Phi}\right\rangle_t &= \frac{K}{K\sigma_K^2}\int_0^t \mathrm{d}s \int_{\mathbb{R}} \mu_s^K(\mathrm{d}x)\int_{\mathbb{R}} \mu_s^K(\mathrm{d}y) b\left(\sigma_K\sqrt{K}x+z_s^K, \sigma_K\sqrt{K}y+z_s^K\right)\\ &\qquad\times\left[\Phi\left(z_s^K+\frac{\sigma_K\sqrt{K}}{K}(y-x), \tau_{-\frac{y-x}{K}}\sharp\left[\mu_s^K-\frac{\delta_x}{K}+\frac{\delta_y}{K}\right]\right)-\Phi\left(z_s^K,\mu_s^K\right)\right]^2\\ &+\frac{K}{K\sigma_K^2}\int_0^t \mathrm{d}s\int_{\mathbb{R}}\mu_s^K(\mathrm{d}x)\theta\left(\sigma_K\sqrt{K}x+z_s^K\right)\int_{\mathbb{R}} m\left(\sigma_K\sqrt{K}x+z_s^K,h\right)\\ &\qquad\times\left[\Phi\left(\mu_s^K+\frac{\sigma_K\sqrt{K}h}{K^{\frac{3}{2}}}, \tau_{-\frac{h}{K^{3/2}}}\sharp\left[\mu_s^K-\frac{\delta_x}{K}+\frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K}\right]\right)-\Phi\left(\mu_s^K,\mu_s^K\right)\right]^2\mathrm{d}h.\end{aligned}$$

Note that the factor $1/K\sigma_K^2$ corresponds to the time scaling of ν_t^K used to define z_t^K and μ_t^K.

Proof. Step 1. About the generators. On the one hand, from (3), note that for all $\Phi:\mathbb{R}\times\mathcal{M}_{1,K}^c(\mathbb{R},\mathbb{R})\to\mathbb{R}$ and $\phi:\mathcal{M}_{1,K}(\mathbb{R})\to\mathbb{R}$ such that $\phi(\nu)=\Phi\left(\langle \mathrm{id},\nu\rangle, \left(h_{\frac{1}{\sigma_K\sqrt{K}}}\circ\tau_{-\langle \mathrm{id},\nu\rangle}\right)\sharp\nu\right)$,

$$\mathcal{L}^K\Phi(z,\mu)=\frac{1}{K\sigma_K^2}\times L^K\phi\left(\left(\tau_z\circ h_{\sigma_K\sqrt{K}}\right)\sharp\,\mu\right). \tag{20}$$

On the other hand, note that for all $\nu\in\mathcal{M}_{1,K}(\mathbb{R})$, $u,v\in\mathbb{R}$, $\psi:\mathbb{R}\to\mathbb{R}$ and $\Psi:\mathcal{M}_{1,K}(\mathbb{R})\to\mathbb{R}$ measurable bounded functions, we have

$$\begin{aligned}&\int_{\mathbb{R}}\psi(x)\Psi\left(\nu+\frac{\delta_{x+\sigma_K\sqrt{K}u}}{K}-\frac{\delta_v}{K}\right)\nu(\mathrm{d}x)\\ &\quad=\int_{\mathbb{R}}\psi\left(\sigma_K\sqrt{K}y+\langle \mathrm{id},\nu\rangle\right)\Psi\left(\left(\tau_{\langle \mathrm{id},\nu\rangle}\circ h_{\sigma_K\sqrt{K}}\right)\sharp\left[\mu+\frac{\delta_{y+u}}{K}-\frac{\delta_v}{K}\right]\right)\mu(\mathrm{d}y).\end{aligned} \tag{21}$$

Finally, by noting that for all $u\in\mathbb{R}$ the mean trait of $\left(\tau_{\langle \mathrm{id},\nu\rangle}\circ h_{\sigma_K\sqrt{K}}\right)\sharp\left[\mu-\frac{\delta_x}{K}+\frac{\delta_u}{K}\right]$ is

$$\langle \mathrm{id},\nu\rangle+\frac{\sigma_K\sqrt{K}}{K}(u-x)$$

and the centered and dilated distribution of traits of $\left(\tau_{\langle \mathrm{id},\nu\rangle}\circ h_{\sigma_K\sqrt{K}}\right)\sharp\left[\mu-\frac{\delta_x}{K}+\frac{\delta_u}{K}\right]$ is

$$\tau_{-\frac{u-x}{K}}\sharp\left[\mu-\frac{\delta_x}{K}+\frac{\delta_u}{K}\right],$$

the announced result follows from (20) and (21).

Step 2. About the martingales. The martingale property follows from classical arguments since Φ and $\mathcal{L}^K\Phi$ are bounded [40, Chapter 4]. To compute the bracket, we proceed according to the following classical method (see e.g. [46]). We apply (19) replacing Φ by Φ^2 to obtain the martingale M^{K,Φ^2}. Then, we apply the Itô formula to compute $\Phi^2\left(z_t^K,\mu_t^K\right)$ from (19). We deduce that

$$\mathrm{Mart}_t:=\Phi^2\left(z_t^K,\mu_t^K\right)-\Phi^2\left(z_0^K,\mu_0^K\right)-2\int_0^t\Phi\left(z_s^K,\mu_s^K\right)\mathcal{L}^K\Phi\left(z_s^K,\mu_s^K\right)\mathrm{d}s-\left\langle M^{K,\Phi}\right\rangle_t$$

is a martingale. Hence, the martingale $M_t^{K,\Phi^2} - \mathrm{Mart}_t$ has finite variation, so it is null [48, Theorem 4.1], leading to

$$\left\langle M^{K,\Phi} \right\rangle_t = \int_0^t \left[\mathcal{L}^K \Phi^2 \left(z_s^K, \mu_s^K \right) - 2\Phi \left(z_s^K, \mu_s^K \right) \mathcal{L}^K \Phi \left(z_s^K, \mu_s^K \right) \right] \mathrm{d}s$$

then the announced conclusion. □

3.2 Asymptotic expansions

Recall that, a set $S \subset \mathscr{C}_b^0(\mathbb{R}, \mathbb{R})$ is called $\mathcal{M}_{1,K}^c(\mathbb{R}) -$ *convergence determining* if whenever $(P_n)_{n\in\mathbb{N}} \in \mathcal{M}_{1,K}^c(\mathbb{R})^{\mathbb{N}}$, $P \in \mathcal{M}_{1,K}^c(\mathbb{R})$ and

$$\lim_{n\to+\infty} \langle f, P_n \rangle = \langle f, P \rangle$$

for all $f \in S$, we have that $(P_n)_{n\in\mathbb{N}}$ converges weakly to P [40, Chapter 3, Section 4, p.112]. As developed in [26, Theorem 3.2.6], the class of functions on $\mathcal{M}_{1,K}^c(\mathbb{R})$,

$$\mathbb{F}_b^2 := \left\{ F_\varphi \,\middle|\, F_\varphi(\mu) := F\left(\langle \varphi, \mu \rangle\right), F \in \mathscr{C}^2(\mathbb{R}, \mathbb{R}), \varphi \in \mathscr{C}_b^2(\mathbb{R}, \mathbb{R}) \right\}$$

is $\mathcal{M}_{1,K}^c(\mathbb{R})$ −convergence determining.

In the following sections, we prove the assertions (i) and (ii) of Section 2.3.1.

3.2.1 Slow component

Proposition 3.2. *For all $f \in \mathscr{C}_b^2(\mathbb{R}, \mathbb{R})$, understood as a function of (z, μ) which depends only on z, the infinitesimal generator $\mathcal{L}^K$ defined in* Proposition 3.1*, satisfies the following decomposition*

$$\mathcal{L}^K f(z, \mu) = \mathcal{L}_{\mathrm{SLOW}} f(z, \mu) + O\left(\frac{1}{K^2} + \frac{M_2(\mu)}{K} + \sigma_K \sqrt{K} M_3(\mu) \right)$$

where the operator $\mathcal{L}_{\mathrm{SLOW}}$ is given by (9)*.*

Proof. From Proposition 3.1 for the choice of test functions $\Phi(z, \mu) := f(z)$ we obtain that $\mathcal{L}^K f(z, \mu) = (\mathbf{A})^K + (\mathbf{B})^K$ where

$$(\mathbf{A})^K := \frac{K}{K\sigma_K^2} \int_{\mathbb{R}} \mu(\mathrm{d}x) \int_{\mathbb{R}} \mu(\mathrm{d}y) b\left(\sigma_K\sqrt{K}x + z, \sigma_K\sqrt{K}y + z\right), \\ \times \left[f\left(z + \frac{\sigma_K\sqrt{K}}{K}(y - x) \right) - f(z) \right]$$

$$(\mathbf{B})^K := \frac{K}{K\sigma_K^2} \int_{\mathbb{R}} \mu(\mathrm{d}x) \theta\left(\sigma_K\sqrt{K}x + z\right) \int_{\mathbb{R}} m\left(\sigma_K\sqrt{K}x + z, h\right) \\ \times \left[f\left(z + \frac{\sigma_K\sqrt{K}}{K^{\frac{3}{2}}} h \right) - f(z) \right] \mathrm{d}h.$$

Now, we want to decompose and study $(\mathbf{A})^K$. Denoting $(\mathbf{C})^K_{x,y}(z) := f\left(z + \frac{\sigma_K\sqrt{K}}{K}(y-x)\right) - f(z)$, from Taylor's formula, note that $(\mathbf{A})^K = (\mathbf{A})^K_1 + (\mathbf{A})^K_2 + (\mathbf{A})^K_3$ where

$$
\begin{aligned}
(\mathbf{A})^K_1 &:= \frac{K}{K\sigma_K^2}\int_{\mathbb{R}}\mu(\mathrm{d}x)\int_{\mathbb{R}}\mu(\mathrm{d}y)b\left(z,z\right)(\mathbf{C})^K_{x,y}(z),\\
(\mathbf{A})^K_2 &:= \frac{K}{K\sigma_K^2}\times\sigma_K\sqrt{K}\int_{\mathbb{R}}\mu(\mathrm{d}x)\int_{\mathbb{R}}\mu(\mathrm{d}y)\left(x\partial_1 b\left(z,z\right)+y\partial_2 b(z,z)\right)(\mathbf{C})^K_{x,y}(z),\\
(\mathbf{A})^K_3 &:= \frac{K}{K\sigma_K^2}\int_{\mathbb{R}}\mu(\mathrm{d}x)\int_{\mathbb{R}}\mu(\mathrm{d}y)\Big[b\left(\sigma_K\sqrt{K}x+z,\sigma_K\sqrt{K}y+z\right)-b(z,z),\\
&\qquad\qquad -\sigma_K\sqrt{K}\left(x\partial_1 b\left(z,z\right)+y\partial_2 b(z,z)\right)\Big](\mathbf{C})^K_{x,y}(z).
\end{aligned}
$$

Using that $(\mathbf{C})^K_{x,y}(z) = f'(z)\frac{\sigma_K\sqrt{K}}{K}(x-y) + O\left(\frac{\sigma_K^2}{K}|y-x|^2\right)$, we deduce that

$$
(\mathbf{A})^K_1 = O\left(\frac{M_2\left(\mu\right)}{K}\right),\quad (\mathbf{A})^K_2 = f'(z)\left(\partial_2 b(z,z)-\partial_1 b(z,z)\right)M_2\left(\mu\right)+O\left(\frac{\sigma_K\sqrt{K}M_3\left(\mu\right)}{K}\right),
$$

and noting that $\left|(\mathbf{C})^K_{x,y}\right| \leqslant C_1\frac{\sigma_K\sqrt{K}}{K}|y-x|$ for some constant $C_1>0$ and

$$
\begin{aligned}
&\left|b\left(\sigma_K\sqrt{K}x+z,\sigma_K\sqrt{K}y+z\right)-b(z,z)-\sigma_K\sqrt{K}\left(x\partial_1 b\left(z,z\right)+y\partial_2 b(z,z)\right)\right|\\
&\qquad\qquad\leqslant C_2\sigma_K^2 K\left(x^2+y^2\right)
\end{aligned}
$$

for some constant $C_2>0$, we deduce that $(\mathbf{A})^K_3 = O\left(\sigma_K\sqrt{K}M_3\left(\mu\right)\right)$. Therefore,

$$
(\mathbf{A})^K = f'(z)\left(\partial_2 b(z,z)-\partial_1 b(z,z)\right)M_2\left(\mu\right)+O\left(\frac{M_2\left(\mu\right)}{K}+\sigma_K\sqrt{K}M_3\left(\mu\right)\right).
$$

As previously, we obtain that

$$
(\mathbf{B})^K = O\left(\frac{\overline{\theta}\overline{m}_2}{K^2}\right).
$$

and the announced result follows. □

3.2.2 Fast component

Proposition 3.3. *For all $F\in\mathscr{C}^3(\mathbb{R},\mathbb{R})$ and $\varphi\in\mathscr{C}^3_b(\mathbb{R},\mathbb{R})$, the infinitesimal generator $\mathcal{L}^K$ defined in* Proposition 3.1*, satisfies the following decomposition*

$$
\mathcal{L}^K F_\varphi\left(z,\mu\right) = \frac{\theta(z)m_2(z)}{K^2\sigma_K^2}\left[\mathcal{L}^{\lambda(z)}_{\mathrm{FVc}}F_\varphi(z,\mu)+O\left(\frac{1}{\sqrt{K}}+\sigma_K K^{\frac{3}{2}}M_2\left(\mu\right)+\frac{M_3(\mu)}{K}\right)\right]
$$

where the operator $\mathcal{L}^{\lambda(z)}_{\mathrm{FVc}}$ is given by (11) *and F_φ is understood as a function of (z,μ) which depends only on μ.*

Proof. Noting that

$$\begin{aligned}(\mathbf{D})_{x,y}^K &:= \left\langle \varphi, \tau_{-\frac{y-x}{K}} \sharp \left[\mu - \frac{\delta_x}{K} + \frac{\delta_y}{K}\right]\right\rangle - \langle \varphi, \mu\rangle \\ &= \left\langle \varphi \circ \tau_{-\frac{y-x}{K}}, \mu\right\rangle + \frac{1}{K}\left\{\varphi\left(y - \frac{y-x}{K}\right) - \varphi\left(x - \frac{y-x}{K}\right)\right\} - \langle \varphi, \mu\rangle,\end{aligned}$$

setting $(\mathbf{C})_{x,y}^K := F\left(\langle \varphi, \mu\rangle + (\mathbf{D})_{x,y}^K\right) - F\left(\langle \varphi, \mu\rangle\right)$ and from Proposition 3.1 we obtain that $\mathcal{L}^K F_\varphi(z,\mu) = (\mathbf{A})^K + (\mathbf{B})^K$ where

$$\begin{aligned}(\mathbf{A})^K &:= \frac{K}{K\sigma_K^2}\int_{\mathbb{R}} \mu(\mathrm{d}x)\int_{\mathbb{R}} \mu(\mathrm{d}y) b\left(\sigma_K\sqrt{K}x + z, \sigma_K\sqrt{K}y + z\right)(\mathbf{C})_{x,y}^K, \\ (\mathbf{B})^K &:= \frac{K}{K\sigma_K^2}\int_{\mathbb{R}} \mu(\mathrm{d}x)\theta\left(\sigma_K\sqrt{K}x + z\right)\int_{\mathbb{R}} m\left(\sigma_K\sqrt{K}x + z, h\right)(\mathbf{C})_{x,x+h/\sqrt{K}}^K \mathrm{d}h.\end{aligned}$$

Step 1. Decomposition and study of $(\mathbf{A})^K$. From Taylor's formula, we obtain that $(\mathbf{A})^K = (\mathbf{A})_1^K + (\mathbf{A})_2^K$ where

$$\begin{aligned}(\mathbf{A})_1^K &:= \frac{K}{K\sigma_K^2}\int_{\mathbb{R}} \mu(\mathrm{d}x)\int_{\mathbb{R}} \mu(\mathrm{d}y) b(z,z)(\mathbf{C})_{x,y}^K, \\ (\mathbf{A})_2^K &:= \frac{K}{K\sigma_K^2}\int_{\mathbb{R}} \mu(\mathrm{d}x)\int_{\mathbb{R}} \mu(\mathrm{d}y)\left[b\left(\sigma_K\sqrt{K}x + z, \sigma_K\sqrt{K}y + z\right) - b(z,z)\right](\mathbf{C})_{x,y}^K.\end{aligned}$$

Denoting $(\mathbf{E})_{x,y}^K := \frac{1}{K}\left(\varphi(y) - \varphi(x) - (y-x)\langle \varphi', \mu\rangle\right)$ and from Taylor's formula again, we obtain that $(\mathbf{A})_1^K = (\mathbf{A})_{11}^K + (\mathbf{A})_{12}^K + (\mathbf{A})_{13}^K + (\mathbf{A})_{14}^K$ where

$$\begin{aligned}(\mathbf{A})_{11}^K &:= \frac{K}{K\sigma_K^2}\int_{\mathbb{R}} \mu(\mathrm{d}x)\int_{\mathbb{R}} \mu(\mathrm{d}y) b(z,z) F'\left(\langle \varphi, \mu\rangle\right)(\mathbf{D})_{x,y}^K, \\ (\mathbf{A})_{12}^K &:= \frac{K}{2K\sigma_K^2}\int_{\mathbb{R}} \mu(\mathrm{d}x)\int_{\mathbb{R}} \mu(\mathrm{d}y) b(z,z) F''\left(\langle \varphi, \mu\rangle\right)\left[(\mathbf{E})_{x,y}^K\right]^2, \\ (\mathbf{A})_{13}^K &:= \frac{K}{2K\sigma_K^2}\int_{\mathbb{R}} \mu(\mathrm{d}x)\int_{\mathbb{R}} \mu(\mathrm{d}y) b(z,z) F''\left(\langle \varphi, \mu\rangle\right)\left[\left[(\mathbf{D})_{x,y}^K\right]^2 - \left[(\mathbf{E})_{x,y}^K\right]^2\right], \\ (\mathbf{A})_{14}^K &:= \frac{K}{K\sigma_K^2}\int_{\mathbb{R}} \mu(\mathrm{d}x)\int_{\mathbb{R}} \mu(\mathrm{d}y) b(z,z) \\ &\qquad \times\left((\mathbf{C})_{x,y}^K - F'\left(\langle \varphi, \mu\rangle\right)(\mathbf{D})_{x,y}^K - \frac{F''\left(\langle \varphi, \mu\rangle\right)}{2}\left[(\mathbf{D})_{x,y}^K\right]^2\right)\end{aligned}$$

From Taylor's formula again and the centered condition of μ, we obtain that

$$\begin{aligned}(\mathbf{A})_{11}^K &= \frac{Kb(z,z)}{K\sigma_K^2}\int_{\mathbb{R}} \mu(\mathrm{d}x)\int_{\mathbb{R}} \mu(\mathrm{d}y) F'\left(\langle \varphi, \mu\rangle\right)\left[(\mathbf{E})_{x,y}^K - \frac{1}{K^2}(y-x)\left(\varphi'(y) - \varphi'(x)\right)\right. \\ &\qquad \left. + \frac{(y-x)^2\langle \varphi'', \mu\rangle}{2K^2} + O\left(\frac{|y-x|^2 + |y-x|^3}{K^3}\right)\right] \\ &= \frac{b(z,z)}{K^2\sigma_K^2} F'\left(\langle \varphi, \mu\rangle\right)\left[-2\langle \varphi' \times \mathrm{id}, \mu\rangle + \langle \varphi'', \mu\rangle M_2(\mu)\right] + \frac{1}{K^2\sigma_K^2} \times O\left(\frac{M_2(\mu) + M_3(\mu)}{K}\right)\end{aligned}$$

In a straightforward way, we obtain that

$$(\mathbf{A})_{12}^K = \frac{b(z,z)}{K^2\sigma_K^2} F''\left(\langle\varphi,\mu\rangle\right)\left[\langle\varphi^2,\mu\rangle - \langle\varphi,\mu\rangle^2 + M_2(\mu)\langle\varphi',\mu\rangle^2 - 2\langle\varphi\times\mathrm{id},\mu\rangle\langle\varphi',\mu\rangle\right].$$

As φ is bounded and so F too, we deduce that there exists two constants $C_1, C_2 > 0$ such that

$$\begin{aligned}\left|\left[(\mathbf{D})_{x,y}^K\right]^2 - \left[(\mathbf{E})_{x,y}^K\right]^2\right| &= \left|(\mathbf{D})_{x,y}^K - (\mathbf{E})_{x,y}^K\right|\left|(\mathbf{D})_{x,y}^K + (\mathbf{E})_{x,y}^K\right| \\ &\leqslant C_1\frac{|y-x| + (y-x)^2}{K^2}\left(\frac{1}{K} + \frac{|y-x|}{K}\right),\end{aligned}$$

$$\left|(\mathbf{C})_{x,y}^K - F'\left(\langle\varphi,\mu\rangle\right)(\mathbf{D})_{x,y}^K - \frac{F''\left(\langle\varphi,\mu\rangle\right)}{2}\left[(\mathbf{D})_{x,y}^K\right]^2\right| \leqslant C_2\frac{1+|y-x|^3}{K^3}.$$

Hence,

$$(\mathbf{A})_{13}^K = \frac{1}{K^2\sigma_K^2}\times O\left(\frac{1+M_3(\mu)}{K}\right) \qquad \text{and} \qquad (\mathbf{A})_{14}^K = \frac{1}{K^2\sigma_K^2}\times O\left(\frac{1+M_3(\mu)}{K}\right).$$

Finally, as $\left|b\left(\sigma_K\sqrt{K}x + z, \sigma_K\sqrt{K}y + z\right) - b(z,z)\right| \leqslant C_3\sigma_K\sqrt{K}\left(|x| + |y|\right)$ and $\left|(\mathbf{C})_{x,y}^K\right| \leqslant C_4\frac{1+|y-x|}{K}$ for some constants $C_3, C_4 > 0$, we deduce that

$$(\mathbf{A})_2^K = \frac{1}{K^2\sigma_K^2}\times O\left(\sigma_K K^{\frac{3}{2}}\left[M_1(\mu) + M_2(\mu)\right]\right).$$

Therefore, using Hölder's inequalities to bound $M_1(\mu) \leqslant \sqrt{M_2(\mu)} \leqslant 1 + M_2(\mu)$ and $M_2(\mu) \leqslant M_3^{2/3}(\mu) \leqslant 1 + M_3(\mu)$ we have that

$$\begin{aligned}(\mathbf{A})^K = &\frac{b(z,z)}{K^2\sigma_K^2}F'\left(\langle\varphi,\mu\rangle\right)\left[-2\langle\varphi'\times\mathrm{id},\mu\rangle + \langle\varphi'',\mu\rangle M_2(\mu)\right] \\ &+ \frac{b(z,z)}{K^2\sigma_K^2}F''\left(\langle\varphi,\mu\rangle\right)\left[\langle\varphi^2,\mu\rangle - \langle\varphi,\mu\rangle^2 + M_2(\mu)\langle\varphi',\mu\rangle^2 - 2\langle\varphi\times\mathrm{id},\mu\rangle\langle\varphi',\mu\rangle\right] \\ &+ \frac{1}{K^2\sigma_K^2}\times O\left(\frac{1}{K} + \sigma_K K^{\frac{3}{2}}\left[1 + M_2(\mu)\right] + \frac{M_3(\mu)}{K}\right).\end{aligned}$$

Step 2. Conclusion. In similar way to Step 1, we obtain that

$$(\mathbf{B})^K = \frac{\theta(z)m_2(z)}{K^2\sigma_K^2}F'\left(\langle\varphi,\mu\rangle\right)\left\langle\frac{\varphi''}{2},\mu\right\rangle + \frac{1}{K^2\sigma_K^2}\times O\left(\frac{1}{\sqrt{K}} + \sigma_K K M_1(\mu)\right).$$

Since, by Assumption (5), $\sigma_K K^{\frac{3}{2}} \ll \frac{1}{\sqrt{K}}$, the announced result follows. □

3.3 Generators and martingales in the case of polynomials in μ

3.3.1 For bounded test functions

In this section, we extend in Lemma 3.4 the result of Proposition 3.1 to test functions of the form

$$P_{f,n}(\mu) := \langle f, \mu^n\rangle := \int_{\mathbb{R}}\cdots\int_{\mathbb{R}} f(x_1,\cdots,x_n)\,\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_n) \tag{22}$$

with $n \in \mathbb{N}^\star$, $\mu \in \mathcal{M}_{1,K}^c(\mathbb{R})$, $f \in \mathscr{C}_b^3(\mathbb{R}^n, \mathbb{R})$ and where μ^n is the $n-$fold product measure of μ. In Lemmas 3.6 and 3.7 we extend this result to the case of specific unbounded test functions.

For all $n \in \mathbb{N}^\star$, we denote by $\mathbf{1} \in \mathbb{R}^n$, the vector whose coordinates are all 1 and by Δ the Laplacian operator on $\mathbb{R}^n$. Let us introduce, for all $n \in \mathbb{N}^\star$, $\lambda > 0$ and $f \in \mathscr{C}_b^2(\mathbb{R}^n, \mathbb{R})$, the operator $B_\lambda^{(n)}$ defined by

$$B_\lambda^{(n)} f(x) := \frac{1}{2}\Delta f(x) - 2\lambda \left(\nabla f(x) \cdot \mathbf{1}\right)(x \cdot \mathbf{1}), \qquad x \in \mathbb{R}^n. \tag{23}$$

Let us consider for all $n \in \mathbb{N}^\star$ for all $i, j \in \{1, \cdots, n\}$,

- $\Phi_{i,j} : \mathscr{C}_b^2(\mathbb{R}^n, \mathbb{R}) \longrightarrow \mathscr{C}_b^2(\mathbb{R}^{n-1}, \mathbb{R})$, with $i \neq j$, is the function obtained from f by inserting the variable x_i between x_{j-1} and x_j when $i < j$ and by inserting the variable x_{i-1} between x_{j-1} and x_j when $i > j$:

$$\begin{array}{ll} \Phi_{i,j} f(x_1, \cdots, x_{n-1}) = f(x_1, \cdots, x_{j-1}, x_i, x_j, x_{j+1}, \cdots, x_{n-1}) & i < j \\ \Phi_{i,j} f(x_1, \cdots, x_{n-1}) = f(x_1, \cdots, x_{j-1}, x_{i-1}, x_j, x_{j+1}, \cdots, x_{n-1}) & i > j \end{array}. \tag{24}$$

- $K_{i,j} : \mathscr{C}_b^2(\mathbb{R}^n, \mathbb{R}) \longrightarrow \mathscr{C}^2(\mathbb{R}^{n+1}, \mathbb{R})$ is defined as

$$K_{i,j} f(x_1, \cdots, x_n, x_{n+1}) := \partial_{ij}^2 f(x_1, \cdots, x_n) x_{n+1}^2. \tag{25}$$

We recall from [18, Definition 2.8] that the extended generator (in the sense of Dynkin, see (27) below) $\mathcal{L}_{\mathrm{FVc}}^\lambda$ of the centered Fleming-Viot process with resampling rate λ is defined for any $n \in \mathbb{N}^\star$, for any test functions $f \in \mathscr{C}_b^2(\mathbb{R}^n, \mathbb{R})$ by

$$\mathcal{L}_{\mathrm{FVc}}^\lambda P_{f,n}(\mu) = \left\langle B_\lambda^{(n)} f, \mu^n \right\rangle + \lambda \sum_{\substack{i,j=1 \\ i \neq j}}^n \left[\left\langle \Phi_{i,j} f, \mu^{n-1} \right\rangle - \langle f, \mu^n \rangle\right] + \lambda \sum_{i,j=1}^n \left\langle K_{i,j} f, \mu^{n+1} \right\rangle. \tag{26}$$

From [18, Definition 2.8] and denoting

$$\widetilde{\Omega} := \left\{ X \in \mathscr{C}^0\left([0, +\infty), \mathcal{M}_1^{c,2}(\mathbb{R})\right) \;\middle|\; \forall T > 0, \sup_{0 \leqslant t \leqslant T} M_2(X_t) < \infty \right\},$$

we recall that a probability measure $\mathbb{P}_\mu$ on $\widetilde{\Omega}$ is said to solve the *centered Fleming-Viot martingale problem for polynomials* with initial condition $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$, if the canonical process $(X_t)_{t \geqslant 0}$ on $\widetilde{\Omega}$ satisfies $\mathbb{P}_\mu(X_0 = \mu) = 1$ and, for all $n \in \mathbb{N}^\star$ and $f \in \mathscr{C}_b^2(\mathbb{R}^n, \mathbb{R})$,

$$\widehat{M}_t^{P_{f,n}} := P_{f,n}(X_t) - P_{f,n}(X_0) - \int_0^t \mathcal{L}_{\mathrm{FVc}}^\lambda P_{f,n}(X_s)\, \mathrm{d}s \tag{27}$$

is a $\mathbb{P}_\mu-$martingale.

Lemma 3.4. *The infinitesimal generator $\mathcal{L}^K$ of the $\mathbb{R} \times \mathcal{M}_{1,K}^c(\mathbb{R})-$valued Markov process $\left(z^K, \mu^K\right)$ given by* Proposition 3.1, *satisfies for all $n \in \mathbb{N}^\star$ and $f \in \mathscr{C}_b^3(\mathbb{R}^n, \mathbb{R})$, the following relations:*

$$\mathcal{L}^K P_{f,n}(z, \mu) = \frac{\theta(z) m_2(z)}{K^2 \sigma_K^2} \mathcal{L}_{\mathrm{FVc}}^{\lambda(z)} P_{f,n}(\mu) + \frac{1}{K^2 \sigma_K^2} \times O\left(\frac{1}{\sqrt{K}} + \sigma_K K^{\frac{3}{2}} M_2(\mu) + \frac{M_3(\mu)}{K}\right)$$

where $\lambda(z) := \frac{b(z,z)}{\theta(z)m_2(z)}$. *Moreover, for all* $K, n \in \mathbb{N}^\star$ *and test functions* $f \in \mathscr{C}_b^3(\mathbb{R}^n, \mathbb{R})$, *the process* $\left(M_t^{K,P_{f,n}}\right)_{t\geqslant 0}$ *defined by*

$$M_t^{K,P_{f,n}} := P_{f,n}\left(\mu_t^K\right) - P_{f,n}\left(\mu_0^K\right) - \int_0^t \mathcal{L}^K P_{f,n}\left(z_s^K, \mu_s^K\right) \mathrm{d}s$$

is a square integrale martingale started at 0.

The proof of this result is given in Section 3.4.

3.3.2 For some unbounded test functions

The following results are particular extensions of Lemma 3.4 when the test functions are no longer bounded but the processes are stopped at time $\check{\tau}^K$ Then, we give in Lemmas 3.6 and 3.7 Doob's semi-martingale decomposition of moments of order 2 and those whose degree is 6.

Proposition 3.5. *For all* $n \in \mathbb{N}^\star$, *for all* $f \in \mathscr{C}^3(\mathbb{R}^n, \mathbb{R})$, *the process* $\left(M_{t\wedge\check{\tau}^K}^{K,P_{f,n}}\right)_{t\geqslant 0}$ *defined by*

$$\begin{aligned}
M_{t\wedge\check{\tau}^K}^{K,P_{f,n}} &:= P_{f,n}\left(\mu_{t\wedge\check{\tau}^K}^K\right) - P_{f,n}\left(\mu_0^K\right) - \frac{K}{K\sigma_K^2}\int_0^{t\wedge\check{\tau}^K}\int_{\mathbb{R}}\mu(\mathrm{d}x)\int_{\mathbb{R}}\mu(\mathrm{d}y) \\
&\quad \times b\left(\sigma_K\sqrt{K}x + z_s^K, \sigma_K\sqrt{K}x + z_s^K\right)\left[P_{f\circ\tau_{-\frac{y-x}{K}},n}\left(\mu - \frac{\delta_x}{K} + \frac{\delta_y}{K}\right) - P_{f,n}(\mu)\right]\mathrm{d}s \\
&\quad + \frac{K}{K\sigma_K^2}\int_0^{t\wedge\check{\tau}^K}\int_{\mathbb{R}}\mu(\mathrm{d}x)\theta\left(\sigma_K\sqrt{K}x + z_s^K\right)\int_{\mathbb{R}} m\left(\sigma_K\sqrt{K}x + z_s^K, h\right)\mathrm{d}h \\
&\quad \times\left[P_{f\circ\tau_{-\frac{h}{K^{\frac{3}{2}}}},n}\left(\mu - \frac{\delta_x}{K} + \frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K}\right) - P_{f,n}(\mu)\right]\mathrm{d}s
\end{aligned}$$

is a bounded martingale.

Proof. Thanks to the stopping time $\check{\tau}^K$, $\left(M_{t\wedge\check{\tau}^K}^{K,P_{f,n}}\right)_{t\geqslant 0}$ is bounded. Hence the martingality of this process follows from Lemma 3.4. □

(a) Moment of order 2. The following result will be useful in Section 4.

Lemma 3.6. *For all* $K \in \mathbb{N}^\star$, *the process* $\left(M_{t\wedge\check{\tau}^K}^{K,P_{\mathrm{id}^2,1}}\right)_{t\geqslant 0}$ *defined in* Proposition 3.5 *satisfies for all* $t \geqslant 0$

$$\begin{aligned}
M_{t\wedge\check{\tau}^K}^{K,P_{\mathrm{id}^2,1}} &= M_2\left(\mu_{t\wedge\check{\tau}^K}^K\right) - M_2\left(\mu_0^K\right) \\
&\quad + \frac{1}{K^2\sigma_K^2}\int_0^{t\wedge\check{\tau}^K}\left\{2b\left(z_s^K, z_s^K\right)M_2\left(\mu_s^K\right) - \theta\left(z_s^K\right)m_2\left(z_s^K\right)\right\}\mathrm{d}s \\
&\quad + \frac{1}{K^2\sigma_K^2}\int_0^{t\wedge\check{\tau}^K} O\left(\frac{1}{K} + \sigma_K\sqrt{K}M_1\left(\mu_s^K\right) + \sigma_K K^{\frac{3}{2}}M_3\left(\mu_s^K\right)\right)\mathrm{d}s
\end{aligned}$$

is a square integrable martingale with quadratic variation:

$$\begin{aligned}\left\langle M^{K,P_{\mathrm{id}^2,1}}\right\rangle_{t\wedge\check{\tau}^K} &= \frac{2}{K^2\sigma_K^2}\int_0^{t\wedge\check{\tau}^K} b\left(z_s^K, z_s^K\right)\left\{M_4\left(\mu_s^K\right) - M_2^2\left(\mu_s^K\right)\right\}\mathrm{d}s \\ &+ \frac{1}{K^2\sigma_K^2}\int_0^{t\wedge\check{\tau}^K} O\left(\frac{\sigma_K\sqrt{K}}{K^2} + \frac{M_2\left(\mu_s^K\right)}{K} + \frac{\sigma_K\sqrt{K}M_3\left(\mu_s^K\right)}{K} + \sigma_K\sqrt{K}M_5\left(\mu_s^K\right)\right)\mathrm{d}s.\end{aligned} \tag{28}$$

Proof. Step 1. Approximation of the Doob *decomposition.* Note that for all $x, y \in \mathbb{R}$,

$$P_{\mathrm{id}^2\circ\tau_{-\frac{y-x}{K}},1}\left(\mu - \frac{\delta_x}{K} + \frac{\delta_y}{K}\right) - P_{\mathrm{id}^2,1}(\mu) = \frac{1}{K}\left(y^2 - x^2\right) - \frac{1}{K^2}(y-x)^2.$$

From Taylor's formula, it follows that

$$\begin{aligned}&\frac{K}{K\sigma_K^2}\int_0^t \mathrm{d}s\int_{\mathbb{R}}\mu(\mathrm{d}x)\int_{\mathbb{R}}\mu(\mathrm{d}y)b\left(\sigma_K\sqrt{K}x + z_s^K, \sigma_K\sqrt{K}y + z_s^K\right) \\ &\qquad\qquad\times\left[P_{\mathrm{id}^2\circ\tau_{-\frac{y-x}{K}},1}\left(\mu - \frac{\delta_x}{K} + \frac{\delta_y}{K}\right) - P_{\mathrm{id}^2,1}(\mu)\right] \\ &= \frac{1}{K^2\sigma_K^2}\int_0^t\left[-2b\left(z_s^K, z_s^K\right)M_2\left(\mu_s^K\right) + O\left(\sigma_K K^{\frac{3}{2}}M_3\left(\mu_s^K\right)\right)\right]\mathrm{d}s\end{aligned}$$

and

$$\begin{aligned}&\frac{K}{K\sigma_K^2}\int_0^t \mathrm{d}s\int_{\mathbb{R}}\mu(\mathrm{d}x)\int_{\mathbb{R}} m\left(\sigma_K\sqrt{K}x + z_s^K, h\right)\theta\left(\sigma_K\sqrt{K}x + z_s^K\right) \\ &\qquad\qquad\times\left[P_{\mathrm{id}^2\circ\tau_{-\frac{h}{K^{3/2}}},1}\left(\mu - \frac{\delta_x}{K} + \frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K}\right) - P_{\mathrm{id}^2,1}(\mu)\right]\mathrm{d}h \\ &= \frac{1}{K^2\sigma_K^2}\int_0^t\left[\theta\left(z_s^K\right)m_2\left(z_s^K\right) + O\left(\frac{1}{K} + \sigma_K\sqrt{K}M_1\left(\mu_s^K\right)\right)\right]\mathrm{d}s.\end{aligned}$$

We deduce the first announced result.

Step 2. Quadratic variation. In similar way to the proof of Proposition 3.1, we have

$$\begin{aligned}\left\langle M^{K,P_{\mathrm{id}^2,1}}\right\rangle_t &= \frac{K}{K\sigma_K^2}\int_0^t \mathrm{d}s\int_{\mathbb{R}}\mu_s^K(\mathrm{d}x)\int_{\mathbb{R}}\mu_s^K(\mathrm{d}y)b\left(\sigma_K\sqrt{K}x + z_s^K, \sigma_K\sqrt{K}y + z_s^K\right) \\ &\qquad\qquad\times\left[\frac{1}{K}\left(y^2 - x^2\right) - \frac{1}{K^2}(y-x)^2\right]^2 \\ &+ \frac{K}{K\sigma_K^2}\int_0^t \mathrm{d}s\int_{\mathbb{R}}\mu_s^K(\mathrm{d}x)\int_{\mathbb{R}} m\left(\sigma_K\sqrt{K}x + z_s^K, h\right)\theta\left(\sigma_K\sqrt{K}x + z_s^K\right) \\ &\qquad\qquad\times\left[\frac{h^2}{K^2} - \frac{h^2}{K^3} + \frac{2xh}{K^{\frac{3}{2}}}\right]^2\mathrm{d}h.\end{aligned}$$

The announced result follows from Taylor's formula and straightforward computations. □

(b) Moments of degree 6. Let us denote for all $\ell \in \mathbb{N}$, $\widetilde{M}_\ell(\mu) := \left\langle \mathrm{id}^\ell, \mu\right\rangle$.

Lemma 3.7. *Let us consider the functions $f, g : \mathbb{R}^2 \to \mathbb{R}$ and $h : \mathbb{R}^3 \to \mathbb{R}$ respectively defined by $f(u,v) := u^4v^2$, $g(u,v) := u^3v^3$ and $h(u,v,w) := u^2v^2w^2$. For all $K \in \mathbb{N}^\star$, the processes $\left(M_{t\wedge\check{\tau}^K}^{K,P_{\mathrm{id}^6,1}}\right)_{t\geqslant 0}$, $\left(M_{t\wedge\check{\tau}^K}^{K,P_{f,2}}\right)_{t\geqslant 0}$, $\left(M_{t\wedge\check{\tau}^K}^{K,P_{g,2}}\right)_{t\geqslant 0}$ and $\left(M_{t\wedge\check{\tau}^K}^{K,P_{h,3}}\right)_{t\geqslant 0}$ satisfiy for all $t \geqslant 0$*

$$
\begin{aligned}
\bullet\ M_{t\wedge\check{\tau}^K}^{K,P_{\mathrm{id}^6,1}} &= M_6\left(\mu_{t\wedge\check{\tau}^K}^K\right) - M_6\left(\mu_0^K\right) + \frac{3}{K^2\sigma_K^2}\int_0^{t\wedge\check{\tau}^K} \Big\{b\left(z_s^K, z_s^K\right)\Big[4M_6\left(\mu_s^K\right) \\
&\qquad -10M_4\left(\mu_s^K\right)M_2\left(\mu_s^K\right)\Big] - 5\theta\left(z_s^K\right)m_2\left(z_s^K\right)M_4\left(\mu_s^K\right)\Big\}\,\mathrm{d}s \\
&\quad + \frac{1}{K^2\sigma_K^2}\int_0^{t\wedge\check{\tau}^K} O\left(\frac{M_3\left(\mu_s^K\right)}{\sqrt{K}} + \frac{1+M_4\left(\mu_s^K\right)}{K} + \sigma_K\sqrt{K}M_5\left(\mu_s^K\right),\right. \\
&\qquad \left. + \frac{M_6\left(\mu_s^K\right)}{K^2} + \sigma_K K^{\frac{3}{2}}M_7\left(\mu_s^K\right)\right)\mathrm{d}s \\
\bullet\ M_{t\wedge\check{\tau}^K}^{K,P_{f,2}} &= M_4\left(\mu_{t\wedge\check{\tau}^K}^K\right)M_2\left(\mu_{t\wedge\check{\tau}^K}^K\right) - M_4\left(\mu_0^K\right)M_2\left(\mu_0^K\right) \\
&\quad + \frac{1}{K^2\sigma_K^2}\int_0^{t\wedge\check{\tau}^K} \Big\{2b\left(z_s^K, z_s^K\right)\Big[6M_4\left(\mu_s^K\right)M_2\left(\mu_s^K\right) - M_6\left(\mu_s^K\right) + 4\widetilde{M}_3^2\left(\mu_s^K\right) \\
&\qquad -6M_2^3\left(\mu_s^K\right)\Big] - \theta\left(z_s^K\right)m_2\left(z_s^K\right)\Big[6M_2^2\left(\mu_s^K\right) + M_4\left(\mu_s^K\right)\Big]\Big\}\,\mathrm{d}s \\
&\quad + \frac{1}{K^2\sigma_K^2}\int_0^{t\wedge\check{\tau}^K} O\left(\frac{M_6\left(\mu_s^K\right)}{K^3} + \sigma_K K^{\frac{3}{2}}M_7\left(\mu_s^K\right)\right)\mathrm{d}s, \\
\bullet\ M_{t\wedge\check{\tau}^K}^{K,P_{g,2}} &= \widetilde{M}_3^2\left(\mu_{t\wedge\check{\tau}^K}^K\right) - \widetilde{M}_3^2\left(\mu_0^K\right) + \frac{1}{K^2\sigma_K^2}\int_0^{t\wedge\check{\tau}^K} \Big\{b\left(z_s^K, z_s^K\right)\Big[-2M_6\left(\mu_s^K\right) \\
&\qquad + 14\widetilde{M}_3^2\left(\mu_s^K\right) + 12M_4\left(\mu_s^K\right)M_2\left(\mu_s^K\right) - 18M_2^3\left(\mu_s^K\right)\Big]\Big\}\,\mathrm{d}s \\
&\quad + \frac{1}{K^2\sigma_K^2}\int_0^{t\wedge\check{\tau}^K} O\left(\frac{M_6\left(\mu_s^K\right)}{K^3} + \sigma_K K^{\frac{3}{2}}M_7\left(\mu_s^K\right)\right)\mathrm{d}s, \\
\bullet\ M_{t\wedge\check{\tau}^K}^{K,P_{h,3}} &= M_2^3\left(\mu_{t\wedge\check{\tau}^K}^K\right) - M_2^3\left(\mu_0^K\right) + \frac{1}{K^2\sigma_K^2}\int_0^{t\wedge\check{\tau}^K} \Big\{b\left(z_s^K, z_s^K\right)\Big[12M_2^3\left(\mu_s^K\right) \\
&\qquad - 6M_4\left(\mu_s^K\right)M_2\left(\mu_s^K\right)\Big] - 3\theta\left(z_s^K\right)m_2\left(z_s^K\right)M_4\left(\mu_s^K\right)\Big\}\,\mathrm{d}s \\
&\quad + \frac{1}{K^2\sigma_K^2}\int_0^{t\wedge\check{\tau}^K} O\left(\frac{M_6\left(\mu_s^K\right)}{K^3} + \sigma_K K^{\frac{3}{2}}M_7\left(\mu_s^K\right)\right)\mathrm{d}s.
\end{aligned}
$$

Proof. The approximation proposed is obtained in a similar way to the proof of Lemma 3.6 □

3.4 Proof of Lemma 3.4

From Proposition 3.1, for the choice of test functions $\Phi(z,\mu) := P_{f,n}(\mu)$ given by (22) and noting that for all $x, y \in \mathbb{R}$,

$$
\left\langle f, \left(\tau_{-\frac{y-x}{K}} \sharp \left[\mu - \frac{\delta_x}{K} + \frac{\delta_y}{K}\right]\right)^n\right\rangle = \left\langle f \circ \tau_{-\frac{y-x}{K}}, \left[\mu - \frac{\delta_x}{K} + \frac{\delta_y}{K}\right]^n\right\rangle,
$$

we obtain that $\mathcal{L}^K P_{f,n}(z,\mu) = \frac{K}{K\sigma_K^2}\left((\mathbf{A})^{\mathrm{K}} + (\mathbf{B})^{\mathrm{K}}\right)$ where

$$(\mathbf{A})^K := \int_{\mathbb{R}} \mu(\mathrm{d}x) \int_{\mathbb{R}} \mu(\mathrm{d}y)\, b\left(\sigma_K\sqrt{K}x + z, \sigma_K\sqrt{K}y + z\right)(\mathbf{C})^K_{x,y},$$
$$(\mathbf{B})^K := \int_{\mathbb{R}} \mu(\mathrm{d}x) \int_{\mathbb{R}} \theta\left(\sigma_K\sqrt{K}x + z\right) m\left(z + \sigma_K\sqrt{K}x, h\right)(\mathbf{C})^K_{x,x+h/\sqrt{K}}\,\mathrm{d}h$$

where $(\mathbf{C})^K_{x,y} := \left\langle f \circ \tau_{-\frac{y-x}{K}}, \left[\mu - \frac{\delta_x}{K} + \frac{\delta_y}{K}\right]^n \right\rangle - \langle f, \mu^n\rangle$. Note that, from Taylor's formula, we obtain that

$$(\mathbf{C})^K_{x,y} = \sum_{k=1}^n \binom{n}{k} \left\langle f, \mu^{n-k}\left[-\frac{\delta_x}{K} + \frac{\delta_y}{K}\right]^k \right\rangle - \frac{y-x}{K}\sum_{i=1}^n \left\langle \partial_i f, \left[\mu - \frac{\delta_x}{K} + \frac{\delta_y}{K}\right]^n \right\rangle$$
$$+ \frac{(y-x)^2}{2K^2}\sum_{i=1}^n\sum_{j=1}^n \left\langle \partial^2_{ij} f, \left[\mu - \frac{\delta_x}{K} + \frac{\delta_y}{K}\right]^n \right\rangle + O\left(\frac{|y-x|^3}{K^3}\right).$$

By abuse of notation, we do not indicate the orders of the products of measure $\mu^{n-k}\left[-\frac{\delta_x}{K} + \frac{\delta_y}{K}\right]^k$.

Step 1. Decomposition and study of $(\mathbf{A})^K$. Note that $(\mathbf{A})^K = \sum_{i=1}^5 (\mathbf{A})^K_i$ where

$$(\mathbf{A})^K_1 := b(z,z)\sum_{k=1}^n \binom{n}{k} \int_{\mathbb{R}} \mu(\mathrm{d}x) \int_{\mathbb{R}} \mu(\mathrm{d}y) \left\langle f, \mu^{n-k}\left[-\frac{\delta_x}{K} + \frac{\delta_y}{K}\right]^k \right\rangle,$$
$$(\mathbf{A})^K_2 := -\frac{b(z,z)}{K}\sum_{i=1}^n\sum_{k=0}^n \binom{n}{k} \int_{\mathbb{R}} \mu(\mathrm{d}x) \int_{\mathbb{R}} \mu(\mathrm{d}y)(y-x) \left\langle \partial_i f, \mu^{n-k}\left[-\frac{\delta_x}{K} + \frac{\delta_y}{K}\right]^k \right\rangle,$$
$$(\mathbf{A})^K_3 := \frac{b(z,z)}{2K^2}\sum_{i,j=1}^n\sum_{k=0}^n \binom{n}{k} \int_{\mathbb{R}} \mu(\mathrm{d}x) \int_{\mathbb{R}} \mu(\mathrm{d}y)(y-x)^2 \left\langle \partial^2_{ij} f, \mu^{n-k}\left[-\frac{\delta_x}{K} + \frac{\delta_y}{K}\right]^k \right\rangle,$$
$$(\mathbf{A})^K_4 := O\left(\frac{M_3(\mu)}{K^3}\right),$$
$$(\mathbf{A})^K_5 := \int_{\mathbb{R}} \mu(\mathrm{d}x) \int_{\mathbb{R}} \mu(\mathrm{d}y)\left[b\left(\sigma_K\sqrt{K}x + z, \sigma_K\sqrt{K}y + z\right) - b(z,z)\right](\mathbf{C})^K_{x,y}.$$

We first study precisely the $(\mathbf{A})^K_2$ term by explaining the case $k = 0$, $k = 1$ and $k \geqslant 2$. We will just give the result for the other $(\mathbf{A})^K_i$, $i \in \{1,3,5\}$: the approach remains the same. By the centered assumption of μ, note that for all $i \in \{1, \cdots, n\}$,

$$\int_{\mathbb{R}} \mu(\mathrm{d}x) \int_{\mathbb{R}} \mu(\mathrm{d}y)(y-x)\left\langle \partial_i f, \mu^n\right\rangle = 0.$$

Then, using again the centered assumption of μ, we obtain that

$$
\begin{aligned}
&\binom{n}{1}\sum_{i=1}^{n}\int_{\mathbb{R}}\mu(\mathrm{d}x)\int_{\mathbb{R}}\mu(\mathrm{d}y)(y-x)\left\langle\partial_i f,\mu^{n-1}\left[-\frac{\delta_x}{K}+\frac{\delta_y}{K}\right]\right\rangle\\
&=\frac{1}{K}\sum_{i=1}^{n}\sum_{j=1}^{n}\int_{\mathbb{R}}\cdots\int_{\mathbb{R}}\int_{\mathbb{R}}\mu(\mathrm{d}x)\int_{\mathbb{R}}\mu(\mathrm{d}y)\\
&\quad\times\Big[x\partial_i f\left(x_1,\cdots,x_{j-1},x,x_{j+1}\cdots,x_n\right)-x\partial_i f\left(x_1,\cdots,x_{j-1},y,x_{j+1}\cdots,x_n\right)\\
&\quad-y\partial_i f\left(x_1,\cdots,x_{j-1},x,x_{j+1}\cdots,x_n\right)+y\partial_i f\left(x_1,\cdots,x_{j-1},y,x_{j+1}\cdots,x_n\right)\Big]\prod_{\substack{\ell=1\\ \ell\neq j}}^{n}\mu(\mathrm{d}x_\ell)\\
&=\frac{2}{K}\sum_{i=1}^{n}\sum_{j=1}^{n}\int_{\mathbb{R}}\cdots\int_{\mathbb{R}}\int_{\mathbb{R}}\mu(\mathrm{d}x)x\partial_i f\left(x_1,\cdots,x_{j-1},x,x_{j+1},\cdots,x_n\right)\prod_{\substack{\ell=1\\ \ell\neq j}}^{n}\mu\left(\mathrm{d}x_\ell\right)\\
&=\frac{2}{K}\left\langle\left(\nabla f\cdot\mathbf{1}\right)\left(\bullet\cdot\mathbf{1}\right),\mu^n\right\rangle.
\end{aligned}
$$

As ∇f is bounded, we deduce that for all $k\geqslant 2$, for all $i\in\{1,\cdots,n\}$,

$$
\int_{\mathbb{R}}\mu(\mathrm{d}x)\int_{\mathbb{R}}\mu(\mathrm{d}y)(y-x)\left\langle\partial_i f,\mu^{n-k}\left[-\frac{\delta_x}{K}+\frac{\delta_y}{K}\right]^k\right\rangle=O\left(\frac{M_1(\mu)}{K^2}\right).
$$

Therefore, we deduce that

$$
(\mathbf{A})_2^K=-\frac{2b(z,z)}{K^2}\left\langle\left(\nabla f\cdot\mathbf{1}\right)\left(\bullet\cdot\mathbf{1}\right),\mu^n\right\rangle+O\left(\frac{M_1(\mu)}{K^3}\right).
$$

In similar way and using Assumptions **(A)** we have that

$$
\begin{aligned}
(\mathbf{A})_1^K&=\frac{b(z,z)}{K^2}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\left[\left\langle\Phi_{i,j}f,\mu^{n-1}\right\rangle-\left\langle f,\mu^n\right\rangle\right]+O\left(\frac{1}{K^3}\right),\\
(\mathbf{A})_3^K&=\frac{b(z,z)}{K^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left\langle K_{i,j}f,\mu^{n+1}\right\rangle+O\left(\frac{M_2\left(\mu\right)}{K^3}\right),\\
(\mathbf{A})_5^K&=O\left(\frac{\sigma_K\sqrt{K}}{K}\left[M_1\left(\mu\right)+M_2\left(\mu\right)\right]\right).
\end{aligned}
$$

Step 2. Decomposition and study of $(\mathbf{B})^K$. For all $\alpha,h\in\mathbb{R}$, we denote $\theta_m\left(\alpha,h\right):=$

$\theta(\alpha)\, m(\alpha, h)$. Note that $(\mathbf{B})^K = \sum_{i=1}^5 (\mathbf{B})_i^K$ where

$$
\begin{aligned}
(\mathbf{B})_1^K &:= \sum_{k=1}^n \binom{n}{k} \int_{\mathbb{R}} \mu(\mathrm{d}x) \int_{\mathbb{R}} \theta_m(z,h) \left\langle f, \mu^{n-k} \left[-\frac{\delta_x}{K} + \frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K} \right]^k \right\rangle \mathrm{d}h, \\
(\mathbf{B})_2^K &:= -\frac{1}{K^{\frac{3}{2}}} \sum_{i=1}^n \sum_{k=0}^n \binom{n}{k} \int_{\mathbb{R}} \mu(\mathrm{d}x) \int_{\mathbb{R}} \theta_m(z,h) h \left\langle \partial_i f, \mu^{n-k} \left[-\frac{\delta_x}{K} + \frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K} \right]^k \right\rangle \mathrm{d}h, \\
(\mathbf{B})_3^K &:= \frac{1}{2K^3} \sum_{i,j=1}^n \sum_{k=0}^n \binom{n}{k} \int_{\mathbb{R}} \mu(\mathrm{d}x) \int_{\mathbb{R}} \theta_m(z,h) h^2 \left\langle \partial_{ij}^2 f, \mu^{n-k} \left[-\frac{\delta_x}{K} + \frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K} \right]^k \right\rangle \mathrm{d}h, \\
(\mathbf{B})_4^K &:= O\left(\frac{1}{K^{\frac{9}{2}}} \int_{\mathbb{R}} \mu(\mathrm{d}x) \int_{\mathbb{R}} \theta_m(z,h) \left| h \right|^3 \mathrm{d}h \right), \\
(\mathbf{B})_5^K &:= \int_{\mathbb{R}} \mu(\mathrm{d}x) \int_{\mathbb{R}} \left[\theta_m\left(\sigma_K \sqrt{K} x + z, h \right) - \theta_m(z,h) \right] (\mathbf{C})_{x, x+h/\sqrt{K}}^K \mathrm{d}h.
\end{aligned}
$$

Note that, from Taylor's formula and Assumptions **(A)**, we have that

$$
\begin{aligned}
&\binom{n}{1} \int_{\mathbb{R}} \mu(\mathrm{d}x) \int_{\mathbb{R}} \theta_m(z,h) \left\langle f, \mu^{n-1} \left[-\frac{\delta_x}{K} + \frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K} \right] \right\rangle \mathrm{d}h \\
&\quad = \frac{1}{K} \sum_{i=1}^n \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{\mathbb{R}} \mu(\mathrm{d}x_i) \int_{\mathbb{R}} \theta_m(z,h) \left[\frac{h}{\sqrt{K}} \partial_i f\left(x_1, \cdots, x_i, \cdots, x_n\right) \right. \\
&\qquad \left. + \frac{h^2}{2K} \partial_{ii}^2 f\left(x_1, \cdots, x_i, \cdots, x_n\right) + O\left(\frac{|h|^3}{K^{\frac{3}{2}}} \right) \right] \prod_{\substack{\ell=1 \\ \ell \neq i}}^n \mu\left(\mathrm{d}x_\ell\right) \\
&\quad = \frac{\theta(z) m_2(z)}{K^2} \left\langle \frac{\Delta f}{2}, \mu^n \right\rangle + O\left(\frac{1}{K^{\frac{5}{2}}} \right).
\end{aligned}
$$

In similar way to Step 1, we obtain that

$$
(\mathbf{B})_1^K = \frac{\theta(z) m_2(z)}{K^2} \left\langle \frac{\Delta f}{2}, \mu^n \right\rangle + O\left(\frac{1}{K^{\frac{5}{2}}} \right), \qquad (\mathbf{B})_2^K = O\left(\frac{1}{K^3} \right),
$$

$$
(\mathbf{B})_3^K = O\left(\frac{1}{K^3} \right), \quad (\mathbf{B})_4^K = O\left(\frac{1}{K^{\frac{9}{2}}} \right) \quad \text{and} \quad (\mathbf{B})_5^K = O\left(\frac{\sigma_K \sqrt{K}}{K^{\frac{3}{2}}} M_1(\mu) \right).
$$

and using Assumptions (5) the announced result follows.

Step 3. Martingality. By classical arguments, since $M^{K, P_{f,n}}$ is bounded, we obtain that $M^{K, P_{f,n}}$ is a square integrable martingale started at 0.

4 Moments estimates

Lemmas 3.6 and 3.7 lead us to look for moment estimates. In Lemma 3.6 we have given the Doob semi-martingale decomposition of the second-order moment and note that involves the third-order moment in the error term. The presence of the higher order moment generates a

difficulty to obtain a fine control of the second order moment. This difficulty is overcome by introducing the stopping time $\check{\tau}^K$, given by (15). In Section 4.1, we give estimates of the moment of order 6 and some corollaries useful for Sections 5 and 6. In Section 4.2 we establish estimates in expectation and probability of the second order moment up to the stopping time $\check{\tau}^K$ and we also control the martingale bracket of the Doob decomposition given in Lemma 3.6. In Sections 4.3 and 4.4, we prove that the stopping time $\tau^K := \check{\tau}^K \wedge \widehat{\tau}^K$, given by (13), converges in probability to $+\infty$ when $K \to +\infty$ using coupling arguments between $M_2\left(\mu^K\right)$ and a biased random walk on $\mathbb{N}^\star$, reflected in 1, and large deviations estimates for this random walk.

4.1 Estimates of the moment of order 6

In this section, we establish a technical lemma useful for Lemma 4.2 below. The idea is to obtain an explicit upper bound of the moments of order 6.

Lemma 4.1. *Given any continuous function $y(t)$ satisfying*

$$y(t) \leqslant y(s) + \int_s^t F\left(y(r)\right) \mathrm{d}r, \qquad \forall s \leqslant t \tag{29}$$

for some Lipschitz function $F : \mathbb{R} \to \mathbb{R}$, then $y(s) \leqslant z(s)$ for all $s \geqslant 0$, where $z(t)$ is the unique solution of

$$z(t) = z(s) + \int_s^t F\left(z(r)\right) \mathrm{d}r, \qquad \forall s \leqslant t \tag{30}$$

such that $z(0) = y(0)$.

Proof. First, it is sufficient to prove that, for all $\varepsilon > 0$, $y(s) \leqslant z_\varepsilon(s)$ for all $s \geqslant 0$, where z_ε is the unique solution of

$$z_\varepsilon(t) = z_\varepsilon(s) + \int_s^t \left[F\left(z_\varepsilon(r)\right) + \varepsilon\right] \mathrm{d}r, \qquad \forall s \leqslant t \tag{31}$$

such that $z_\varepsilon(0) = y(0) + \varepsilon$. Indeed, it follows from (30) and (31) that, for all $t \geqslant 0$ and $\varepsilon > 0$

$$|z(t) - z_\varepsilon(t)| \leqslant \varepsilon(1+t) + C_{\mathrm{Lip}} \int_0^t |z(s) - z_\varepsilon(s)| \, \mathrm{d}s$$

where C_{Lip} if the Lipschitz constant of F. Gronwall's lemma then entails

$$|z(t) - z_\varepsilon(t)| \leqslant \varepsilon(1+t) \exp\left(C_{\mathrm{Lip}} t\right) \xrightarrow[\varepsilon \to 0]{} 0$$

so that $y(t) \leqslant z_\varepsilon(t)$, for all $\varepsilon > 0$ indeed implies $y(t) \leqslant z(t)$.

So, let $\varepsilon > 0$ be fixed and let us prove that $y(t) \leqslant z_\varepsilon(t)$ for all $t \geqslant 0$. Define

$$\vartheta_\varepsilon := \inf\left\{ t \geqslant 0 \,\middle|\, y(t) \geqslant z_\varepsilon(t) \right\}.$$

Our goal is to prove that $\vartheta_\varepsilon = +\infty$.

By continuity of y, $\vartheta_\varepsilon > 0$ and, if $\vartheta_\varepsilon < \infty$, $y\left(\vartheta_\varepsilon\right) = z_\varepsilon\left(\vartheta_\varepsilon\right)$. Assume by contradiction that $\vartheta_\varepsilon < \infty$. Then, using (29) and (31), for all $\delta \in (0, \vartheta_\varepsilon)$,

$$\frac{z_\varepsilon\left(\vartheta_\varepsilon\right) - y\left(\vartheta_\varepsilon\right) - \left(z_\varepsilon\left(\vartheta_\varepsilon - \delta\right) - y\left(\vartheta_\varepsilon - \delta\right)\right)}{\delta} \geqslant \varepsilon + \frac{1}{\delta} \int_{\vartheta_\varepsilon - \delta}^{\vartheta_\varepsilon} \left[F\left(z_\varepsilon(r)\right) - F(y(r))\right] \mathrm{d}r.$$

By continuity of z_ε and y, the last term of the right-hand side converges to 0 when $\delta \to 0$. Hence the left-hand side is positive for all δ small enough. Since $z_\varepsilon(\vartheta_\varepsilon) = y(\vartheta_\varepsilon)$, this implies that

$$z_\varepsilon(\vartheta_\varepsilon - \delta) - y(\vartheta_\varepsilon - \delta) < 0$$

for such δ. This contradicts the definition of ϑ_ε as the first time t such that $z_\varepsilon(t) \leqslant y(t)$. □

Lemma 4.2. *There exists $K_0 \in \mathbb{N}^\star$ large enough and two constants $C_1, C_2 > 0$ such that for all $K \geqslant K_0$ and $t \geqslant 0$,*

$$\mathbb{E}\left(M_6\left(\mu_t^K\right)\mathbb{1}_{t \leqslant \check{\tau}^K}\right) \leqslant C_1 \exp\left(-\frac{9\underline{b}}{28K^2\sigma_K^2}t\right) + C_2$$

Proof. Step 1. Pathwise inequality. Note that for all $t \geqslant 0$,

$$M_7\left(\mu_{t\wedge\check{\tau}^K}^K\right) \leqslant \mathrm{Diam}\left(\mathrm{Supp}\,\mu_{t\wedge\check{\tau}^K}^K\right) M_6\left(\mu_{t\wedge\check{\tau}^K}^K\right) \leqslant \frac{M_6\left(\mu_{t\wedge\check{\tau}^K}^K\right)}{\sigma_K K^{\frac{3+\varepsilon}{2}}}. \tag{32}$$

Introducing for all $t \geqslant 0$, $A_t := \frac{3}{4}M_6\left(\mu_t^K\right) + 7M_4\left(\mu_t^K\right)M_2\left(\mu_t^K\right) - 4\widetilde{M}_3^2\left(\mu_t^K\right) + \frac{3}{2}M_2^3\left(\mu_t^K\right)$, we obtain from Lemma 3.7, (32) and Assumptions **(A)** that there exists a constant $C > 0$ such that for all $t > s \geqslant 0$,

$$\begin{aligned}A_{t\wedge\check{\tau}^K} \leqslant A_{s\wedge\check{\tau}^K} - \frac{1}{K^2\sigma_K^2}\int_{s\wedge\check{\tau}^K}^{t\wedge\check{\tau}^K}\Bigg\{\left(\underline{b} - \frac{C}{K^{\frac{\varepsilon}{2}}}\right)\left[3M_6\left(\mu_r^K\right) + \frac{9}{2}M_4\left(\mu_r^K\right)M_2\left(\mu_r^K\right) + 6M_2^3\left(\mu_r^K\right)\right] \\ - \overline{\theta}\overline{m}_2\left[\frac{91}{4}M_4\left(\mu_r^K\right) + 42M_2^2\left(\mu_r^K\right)\right]\mathrm{d}r\Bigg\} + \mathrm{Mart}_{t\wedge\check{\tau}^K} - \mathrm{Mart}_{s\wedge\check{\tau}^K}\end{aligned}$$

where $(\mathrm{Mart}_t)_{t\geqslant 0}$ is a martingale. Note that, choosing K_0 large enough, for all $K \geqslant K_0$, $\underline{b} - CK^{-\varepsilon/2} > \underline{b}/2$. From the inequalities $M_2^2(\mu) \leqslant M_4(\mu)$ and $M_4(\mu) \leqslant M_6^{2/3}(\mu) \leqslant \alpha M_6(\mu) + 1/\alpha^2$ with $\alpha := 2\underline{b}/259\overline{\theta}\overline{m}_2$, we deduce that

$$\frac{3\underline{b}}{2}M_6(\mu) - \overline{\theta}\overline{m}_2\left[\frac{91}{4}M_4(\mu) + 42M_2^2(\mu)\right] \geqslant \frac{3\underline{b}}{2}M_6(\mu) - \frac{259\overline{\theta}\overline{m}_2}{4}M_4(\mu) \geqslant \underline{b}M_6(\mu) - \frac{259^3\overline{\theta}^3\overline{m}_2^3}{16\underline{b}^2}.$$

Thus, for all $t > s \geqslant 0$,

$$\begin{aligned}A_{t\wedge\check{\tau}^K} &\leqslant A_{s\wedge\check{\tau}^K} + \mathrm{Mart}_{t\wedge\check{\tau}^K} - \mathrm{Mart}_{s\wedge\check{\tau}^K} \\ &\quad - \frac{\underline{b}}{2K^2\sigma_K^2}\int_{s\wedge\check{\tau}^K}^{t\wedge\check{\tau}^K}\left[2M_6\left(\mu_r^K\right) + \frac{9}{2}M_4\left(\mu_r^K\right)M_2\left(\mu_r^K\right) + 6M_2^3\left(\mu_r^K\right) - \left(\frac{259\overline{\theta}\overline{m}_2}{2\underline{b}}\right)^3\right]\mathrm{d}r \\ &\leqslant A_{s\wedge\check{\tau}^K} - \frac{9\underline{b}}{28K^2\sigma_K^2}\int_{s\wedge\check{\tau}^K}^{t\wedge\check{\tau}^K}\left(A_r - \frac{14}{9}\left(\frac{259\overline{\theta}\overline{m}_2}{2\underline{b}}\right)^3\right)\mathrm{d}r + \mathrm{Mart}_{t\wedge\check{\tau}^K} - \mathrm{Mart}_{s\wedge\check{\tau}^K}\end{aligned}$$

Step 2. Conclusion. Let us consider the stochastic process $\left(N\left(\mu_t^K\right)\right)_{t\geqslant 0}$ defined by

$$N\left(\mu_t^K\right) := \begin{cases} A_t & \text{if} \quad s < t \leqslant \check{\tau}^K \\ \left(A_{\check{\tau}^K} - C_0\right)\exp\left(-\frac{3\underline{b}}{4K^2\sigma_K^2}\left(t - \check{\tau}^K\right)\right) + C_0 & \text{if} \quad s < \check{\tau}^K < t \\ 0 & \text{if} \quad \check{\tau}^K < s < t \end{cases}$$

where $C_0 := \frac{14}{9}\left(\frac{259\bar{\theta}\overline{m}_2}{2\underline{b}}\right)^3$. From Step 1, for all $t > s \geqslant 0$,

$$N\left(\mu_t^K\right) \leqslant N\left(\mu_s^K\right) - \frac{9\underline{b}}{28K^2\sigma_K^2}\int_s^t \left[N\left(\mu_r^K\right) - C_0\right] \mathrm{d}r + \mathrm{Mart}_t - \mathrm{Mart}_s$$

which entails

$$y(t) \leqslant y(s) - \frac{9\underline{b}}{28K^2\sigma_K^2}\int_s^t (y(r) - C_0)\,\mathrm{d}r$$

where $y(t) := \mathbb{E}\left(N\left(\mu_t^K\right)\right)$. From Lemma 4.1, we deduce that for all $t \geqslant 0$, $y(t) \leqslant z(t)$ where $z(t)$ is the unique solution of

$$z(t) = z(s) - \frac{9\underline{b}}{28K^2\sigma_K^2}\int_s^t (z(r) - C_0)\,\mathrm{d}r$$

such that $z(0) = A_0$, i.e. $z(t) = (z(0) - C_0)\exp\left(-\frac{9\underline{b}}{28K^2\sigma_K^2}t\right) + C_0$. From (7) then (4), note that

$$M_6\left(\mu_0^K\right) = \left\langle \mathrm{id}^6, \mu_0^K\right\rangle = \frac{1}{K^3\sigma_K^6}\left\langle (\mathrm{id} - x_0)^6, \nu_0^K\right\rangle \leqslant C_6^\star.$$

From Cauchy-Schwarz's inequality $\widetilde{M}_3^2(\mu) \leqslant M_4(\mu)M_2(\mu)$, note that $z(0) \leqslant 21C_6^\star/4$ and so, for all $t \geqslant 0$,

$$\begin{aligned}\mathbb{E}\left(M_6\left(\mu_t^K\right)\mathbb{1}_{t\leqslant\check{\tau}^K}\right) &\leqslant \frac{4}{3}\mathbb{E}\left(\left[\frac{3}{4}M_6\left(\mu_t^K\right) + 3M_4\left(\mu_t^K\right)M_2\left(\mu_t^K\right) + \frac{3}{2}M_2^3\left(\mu_t^K\right)\right]\mathbb{1}_{t\leqslant\check{\tau}^K}\right)\\ &\leqslant \frac{4}{3}z(t) \leqslant \frac{4}{3}\left(\frac{21}{4}(C_6^\star - C_0)\exp\left(-\frac{9\underline{b}}{28K^2\sigma_K^2}t\right) + C_0\right)\end{aligned}$$

and the announced result follows. □

The following lemma will be useful in the proof of Theorem 5.1.

Lemma 4.3. *For all $t \geqslant 0$,*

$$\sup_{K\in\mathbb{N}^\star}\mathbb{E}\left(\int_0^{t\wedge\check{\tau}^K} M_6\left(\mu_s^K\right)\mathrm{d}s\right) < \infty.$$

Proof. Let $t \geqslant 0$ be fixed. By Fubini's theorem and Lemma 4.2, there exists two constants $C_1, C_2 > 0$ such that

$$\mathbb{E}\left(\int_0^t M_6\left(\mu_s^K\right)\mathbb{1}_{s\wedge\check{\tau}^K}\mathrm{d}s\right) \leqslant C_1\int_0^t \exp\left(-\frac{9\underline{b}}{28K^2\sigma_K^2}s\right)\mathrm{d}s + C_2 t \leqslant (C_1 + C_2)t$$

which ends the proof. □

The last lemma has the following consequence, used in Sections 5 and 6.

Corollary 4.4. *For all $T > 0$, the family*

$$\left(\int_0^{t \wedge \check{\tau}^K} M_5 \left(\mu_s^K \right) \mathrm{d}s \right)_{t \in [0,T], K \in \mathbb{N}^\star}$$

is uniformly integrable.

Proof. Note that for all $K \in \mathbb{N}^\star$, for all $t \in [0,T]$, for all $A \in (0, +\infty)$,

$$\mathbb{1}_{\int_0^t M_5(\mu_s^K) \mathbb{1}_{s \leqslant \check{\tau}^K} \mathrm{d}s \geqslant A} \int_0^t M_5 \left(\mu_s^K \right) \mathbb{1}_{s \leqslant \check{\tau}^K} \mathrm{d}s \leqslant \frac{1}{\sqrt[5]{A}} \left(\int_0^t M_5 \left(\mu_s^K \right) \mathbb{1}_{s \leqslant \check{\tau}^K} \mathrm{d}s \right)^{6/5},$$

we deduce from Hölder'inequality that, for all $t \in [0,T]$, for all $A \in (0, +\infty)$,

$$\begin{aligned} \sup_{K \in \mathbb{N}^\star} \mathbb{E} \Bigg(\mathbb{1}_{\int_0^t M_5(\mu_s^K) \mathbb{1}_{s \leqslant \check{\tau}^K} \mathrm{d}s \geqslant A} \int_0^t M_5 \left(\mu_s^K \right) \mathbb{1}_{s \leqslant \check{\tau}^K} \mathrm{d}s \Bigg) & \\ \leqslant \frac{T}{\sqrt[5]{A}} \sup_{K \in \mathbb{N}^\star} \mathbb{E} \left(\int_0^t M_5^{6/5} \left(\mu_s^K \right) \mathbb{1}_{s \leqslant \check{\tau}^K} \mathrm{d}s \right) & \\ \leqslant \frac{T}{\sqrt[5]{A}} \sup_{K \in \mathbb{N}^\star} \mathbb{E} \left(\int_0^t M_6 \left(\mu_s^K \right) \mathbb{1}_{s \leqslant \check{\tau}^K} \mathrm{d}s \right). & \end{aligned}$$

From Lemma 4.3, we deduce that the right hand side of the previous inequality goes to 0 when $A \to +\infty$ and so the announced result. □

4.2 Some inequalities on the moment of order 2

Lemma 4.5. *There exists $K_0 \in \mathbb{N}^\star$ large enough such that for all $K \geqslant K_0$, the process $M_2\left(\mu^K\right)$ satisfies for all $t \geqslant 0$ the following inequalities:*

$$\begin{aligned} & M_2 \left(\mu_{s \wedge \check{\tau}^K}^K \right) - \frac{5\underline{b}}{2K^2\sigma_K^2} \int_{s \wedge \check{\tau}^K}^{t \wedge \check{\tau}^K} \left[M_2 \left(\mu_r^K \right) - \frac{2\theta m_2}{5\underline{b}} + \frac{2\overline{\theta}\overline{m}_2}{5\underline{b}K} \right] \mathrm{d}r + M_{t \wedge \check{\tau}^K}^{K, P_{\mathrm{id}^2,1}} - M_{s \wedge \check{\tau}^K}^{K, P_{\mathrm{id}^2,1}} \\ & \quad \leqslant M_2 \left(\mu_{t \wedge \check{\tau}^K}^K \right) \leqslant M_2 \left(\mu_{s \wedge \check{\tau}^K}^K \right) - \frac{\underline{b}}{K^2\sigma_K^2} \int_{s \wedge \check{\tau}^K}^{t \wedge \check{\tau}^K} \left[M_2 \left(\mu_r^K \right) - \frac{\overline{\theta}\overline{m}_2}{\underline{b}} \right] \mathrm{d}r + M_{t \wedge \check{\tau}^K}^{K, P_{\mathrm{id}^2,1}} - M_{s \wedge \check{\tau}^K}^{K, P_{\mathrm{id}^2,1}} \end{aligned} \tag{33}$$

where $M_{t \wedge \check{\tau}^K}^{K, P_{\mathrm{id}^2,1}}$ is defined in Proposition 3.5. *Moreover, for $t \geqslant 0$,*

$$\mathbb{E} \left(M_2 \left(\mu_t^K \right) \mathbb{1}_{t \leqslant \check{\tau}^K} \right) \leqslant \mathbb{E} \left(M_2 \left(\mu_0^K \right) - \frac{\overline{\theta}\overline{m}_2}{\underline{b}} \right) \exp \left(- \frac{\underline{b}}{K^2\sigma_K^2} t \right) + \frac{\overline{\theta}\overline{m}_2}{\underline{b}}.$$

Proof. Note that for all $t \geqslant 0$,

$$M_3 \left(\mu_{t \wedge \check{\tau}^K}^K \right) \leqslant \mathrm{Diam} \left(\mathrm{Supp}\, \mu_{t \wedge \check{\tau}^K}^K \right) M_2 \left(\mu_{t \wedge \check{\tau}^K}^K \right) \leqslant \frac{M_2 \left(\mu_{t \wedge \check{\tau}^K}^K \right)}{\sigma_K K^{\frac{3+\varepsilon}{2}}}. \tag{34}$$

From Lemma 3.6, (34) and Assumptions **(A)**, there exists two constants $C, \widetilde{C} > 0$ such that for all $t > s \geqslant 0$

$$M_2\left(\mu^K_{s\wedge\check{\tau}^K}\right) - \frac{1}{K^2\sigma_K^2}\int_{s\wedge\check{\tau}^K}^{t\wedge\check{\tau}^K}\left(\left[2\overline{b} + \frac{C}{K^{\frac{\varepsilon}{2}}}\right] M_2\left(\mu^K_r\right) - \underline{\theta m}_2 + \frac{\overline{\theta}\,\overline{m}_2}{K}\right)\mathrm{d}r + M^{K,P_{\mathrm{id}^2,1}}_{t\wedge\check{\tau}^K} - M^{K,P_{\mathrm{id}^2,1}}_{s\wedge\check{\tau}^K}$$
$$\leqslant M_2\left(\mu^K_{t\wedge\check{\tau}^K}\right)$$
$$\leqslant M_2\left(\mu^K_{s\wedge\check{\tau}^K}\right) - \frac{1}{K^2\sigma_K^2}\int_{s\wedge\check{\tau}^K}^{t\wedge\check{\tau}^K}\left(\left[2\underline{b} - \frac{\widetilde{C}}{K^{\frac{\varepsilon}{2}}}\right] M_2\left(\mu^K_r\right) - \overline{\theta}\,\overline{m}_2 + \frac{\theta m_2}{K}\right)\mathrm{d}r + M^{K,P_{\mathrm{id}^2,1}}_{t\wedge\check{\tau}^K} - M^{K,P_{\mathrm{id}^2,1}}_{s\wedge\check{\tau}^K}.$$

Noting that for all $K \geqslant K_0$, $2\underline{b} - \widetilde{C}K^{-\varepsilon/2} > \underline{b}$ and $2\overline{b} + CK^{-\varepsilon/2} < 5\overline{b}/2$, the announced first result follows. In similar way to Step 2 of the proof of Lemma 4.2, we obtain the second part of the announced result. □

Corollary 4.6. *For all $t \geqslant 0$, $\mathbb{E}\left(M_2\left(\mu^K_t\right)\mathbb{1}_{t\leqslant\check{\tau}^K}\right) \leqslant \max\left\{C^\star_2, \overline{\theta}\,\overline{m}_2/\underline{b}\right\}$.*

Proof. From (7) and (4), note that

$$M_2\left(\mu^K_0\right) = \left\langle \mathrm{id}^2, \mu^K_0\right\rangle \leqslant C^\star_2.$$

The announced result follows from Lemma 4.5. □

The goal of the following result is to bound the martingale bracket $\left\langle M^{K,P_{\mathrm{id}^2,1}}\right\rangle_{t\wedge\check{\tau}^K}$. To do this, we exploit the martingale approximation of the moment of order 6 in order to control the dominant term $M_4\left(\mu^K_s\right) - M_2^2\left(\mu^K_s\right)$ in (28).

Lemma 4.7. *There exists a constant $C > 0$ such that for all $K \in \mathbb{N}^\star$, for all $t \geqslant 0$ the martingale bracket $\left\langle M^{K,P_{\mathrm{id}^2,1}}\right\rangle_t$ satisfies*

$$\mathbb{E}\left(\left\langle M^{K,P_{\mathrm{id}^2,1}}\right\rangle_{t\wedge\check{\tau}^K}\right) \leqslant \frac{C}{K^2\sigma_K^2}\mathbb{E}\left(t\wedge\check{\tau}^K\right).$$

Proof. From Lemma 3.6 and Hölder's inequalities $M_2^2(\mu) \leqslant M_4(\mu) \leqslant M_6^{2/3} \leqslant 1 + M_6(\mu)$ and $M_5(\mu) \leqslant M_6(\mu)^{5/6} \leqslant 1 + M_6(\mu)$, we deduce that there exists a constant $C > 0$ such that for all $t \geqslant 0$

$$\left\langle M^{K,P_{\mathrm{id}^2,1}}\right\rangle_{t\wedge\check{\tau}^K} \leqslant \frac{C}{K^2\sigma_K^2}\int_0^{t\wedge\check{\tau}^K}\left(1 + M_6\left(\mu^K_s\right)\right)\left(1 + \frac{1}{K}\right)\mathrm{d}s.$$

The announced result follows from Fubini's theorem and Lemma 4.2. □

4.3 Convergence to $+\infty$ of the stopping time τ^K and support concentration property

The main result of this section is the following convergence result for the stopping time τ^K, defined by (13), when $K \to +\infty$.

Proposition 4.8. *Under* Assumption (5), *τ^K converges in probability to $+\infty$ when $K \to +\infty$.*

The proof is based on the next lemma proved in Section 4.4 below. Thanks to Proposition 4.8, we deduce the support concentration property of $\left(\nu^K_{t/K\sigma_K^2}\right)$ given by (6). Indeed, with probability which tends to 1 when $K \to +\infty$, $\mathrm{Diam}\left(\mu_t^K\right) \leqslant \frac{1}{\sigma_K K^{\frac{3+\varepsilon}{2}}}$ for all $t \in [0,T]$. Hence,

$$\mathrm{Diam}\left(\nu^K_{t/K\sigma_K^2}\right) = \sigma_K\sqrt{K}\mathrm{Diam}\left(\mu_t^K\right) \leqslant \frac{1}{K^{1+\frac{\varepsilon}{2}}}$$

and (6) follows.

Lemma 4.9. *Under* Assumption (5)*, for all* $T \geqslant 0$*,*

$$\lim_{K\to+\infty} \mathbb{P}\left(\widehat{\tau}^K \leqslant \check{\tau}^K \wedge T\right) = \lim_{K\to+\infty} \mathbb{P}\left(\sup_{0\leqslant t\leqslant T\wedge\check{\tau}^K} M_2\left(\mu_t^K\right) \geqslant K^\varepsilon\right) = 0.$$

Proof of Proposition 4.8. Note that for all $T \geqslant 0$,

$$\mathbb{P}\left(\check{\tau}^K < T \wedge \widehat{\tau}^K\right) \leqslant \mathbb{P}\left(\exists t < T \wedge \widehat{\tau}^K, \mathrm{Diam}\left(\mathrm{Supp}\,\mu_t^K\right) > \frac{1}{\sigma_K K^{\frac{3+\varepsilon}{2}}}\right).$$

As for all $t \geqslant 0$, for all $K \in \mathbb{N}^\star$, $\mathrm{Diam}\left(\mathrm{Supp}\,\mu_t^K\right)^2 \leqslant \left(2\max\left\{|x| \,\middle|\, x \in \mathrm{Supp}\,\mu_t^K\right\}\right)^2 \leqslant 4KM_2\left(\mu_t^K\right)$, we deduce that

$$\begin{aligned}\mathbb{P}\left(\exists t < T \wedge \widehat{\tau}^K, \mathrm{Diam}\left(\mathrm{Supp}\,\mu_t^K\right) > \frac{1}{\sigma_K K^{\frac{3+\varepsilon}{2}}}\right)& \\ \leqslant \mathbb{P}\left(\exists t < T\wedge, \widehat{\tau}^K M_2\left(\mu_t^K\right) > \frac{1}{4\sigma_K^2 K^{4+\varepsilon}}\right).&\end{aligned}$$

According to (5), we deduce that the following inequality $4\sigma_K^2 K^{4+\varepsilon} \leqslant 1/K^\varepsilon$ holds true when $K \to +\infty$ and $\sigma_K \to 0$. Hence, we deduce that

$$\mathbb{P}\left(\exists t < T \wedge \widehat{\tau}^K, M_2\left(\mu_t^K\right) > \frac{1}{4\sigma_K^2 K^{4+\varepsilon}}\right) \leqslant \mathbb{P}\left(\exists t < T \wedge \widehat{\tau}^K, M_2\left(\mu_t^K\right) > K^\varepsilon\right) = 0.$$

Now, for all $T > 0$, $\left\{\check{\tau}^K \geqslant T \wedge \widehat{\tau}^K\right\} = \left\{\check{\tau}^K \geqslant \widehat{\tau}^K\right\} \cup \left\{\check{\tau}^K \geqslant T\right\}$, and so

$$\begin{aligned}\mathbb{P}\left(\left\{\check{\tau}^K \geqslant T \wedge \widehat{\tau}^K\right\} \cap \left\{\tau^K < T\right\}\right) &\leqslant \mathbb{P}\left(\left[\left\{\widehat{\tau}^K < T\right\} \cap \left\{\widehat{\tau}^K \leqslant \check{\tau}^K\right\}\right] \cup \left\{\widehat{\tau}^K < T \leqslant \check{\tau}^K\right\}\right) \\ &\leqslant \mathbb{P}\left(\widehat{\tau}^K \leqslant T \wedge \check{\tau}^K\right).\end{aligned}$$

Note that for all $T > 0$,

$$\mathbb{P}\left(\tau^K < T\right) \leqslant \mathbb{P}\left(\check{\tau}^K < T \wedge \widehat{\tau}^K\right) + \mathbb{P}\left(\widehat{\tau}^K \leqslant T \wedge \check{\tau}^K\right).$$

Hence, from Lemma 4.9, we have for all $T \geqslant 0$, $\lim_{K\to+\infty} \mathbb{P}\left(\tau^K < T\right) = 0$ and the announced result follows. □

4.4 Proof of Lemma 4.9

Note that if $M_2\left(\mu_s^K\right)$ becomes larger than $\overline{\theta}\overline{m}_2/\underline{b}$, the drift term in the right-hand side of (33) is then negative preventing $M_2\left(\mu_s^K\right)$ to become excessively large unless the martingale $M^{K,P_{\mathrm{id}^2,1}}$ has an exceptional path with large increments on a small time interval. We then expect to have large deviations estimates and bounds on the time of exit from attractive domains for $M_2\left(\mu^K\right)$. However, we cannot consider establishing directly a large deviation principle on $M_2\left(\mu^K\right)$, in particular because of the slow-fast limit. The approach considered below is based on coupling arguments between $M_2\left(\mu^K\right)$ and a simpler process for which large deviations estimates are known.

Let us introduce $u_0 := 0$ and for all $\ell \in \mathbb{N}^\star$ the real number $u_\ell := 3^\ell K^{\frac{\varepsilon}{2}}$ and the interval $\mathcal{I}_\ell := [u_{\ell-1}, u_{\ell+1})$. We will look at the process $M_2\left(\mu^K\right)$ at successive exit times of $(\mathcal{I}_\ell)_{\ell\in\mathbb{N}^\star}$. We set $L_0^K := 1$ when $M_2\left(\mu_0^K\right) \in [0, 2u_1+1]$ and $L_0^K := \ell_0 \in \mathbb{N}^\star$ when $M_2\left(\mu_0^K\right) \in (2u_{\ell_0-1}+1, 2u_{\ell_0}+1]$. We also set $T_0^K := 0$ and by induction on $k \in \mathbb{N}^\star$, if $\left(L_i^K\right)_{1\leqslant i\leqslant k}$ and $\left(T_i^K\right)_{1\leqslant i\leqslant k}$ are constructed,

$$T_{k+1}^K := \inf\left\{t \geqslant T_k^K \,\middle|\, M_2\left(\mu_t^K\right) \notin \mathcal{I}_{L_k^K}\right\}$$

and

$$L_{k+1}^K := \min\left\{\ell \in \mathbb{N} \,\middle|\, M_2\left(\mu_{T_{k+1}^K}^K\right) \leqslant 2u_\ell + 1\right\}.$$

Note that from (4), for K large enough, $L_0^K = 1$ because $M_2\left(\mu_0^K\right) \leqslant C_2^\star$. Since $\left(\mu_t^K\right)_{t\geqslant 0}$ is a pure jump process, note that $\left(M_2\left(\mu_t^K\right)\right)_{t\geqslant 0}$ is also a pure jump process. The definition of u_ℓ is motivated by the fact that, for all $k \geqslant 0$, $L_{k+1}^K \leqslant L_k^K + 1$, which follows from the next lemma.

Lemma 4.10. *Let ϑ_1 be the first jump time of μ^K. Then, for K large enough*

$$M_2\left(\mu_{\vartheta_1}^K\right) \leqslant 2M_2\left(\mu_0^K\right) + 1.$$

In particular, for all $t > 0$, $\Delta M_2\left(\mu_t^K\right) := M_2\left(\mu_t^K\right) - M_2\left(\mu_{t^-}^K\right) \leqslant M_2\left(\mu_{t^-}^K\right) + 1$.

Proof. Let us consider $\mu_0^K := \frac{1}{K}\sum_{k=1}^K \delta_{x_k}$ and $y = (y_k)_{1\leqslant k\leqslant K}$ defined by

$$y = \begin{cases} \left(x_1, \cdots, x_{i-1}, x_i + \frac{H}{\sqrt{K}}, x_{i+1}, \cdots, x_K\right) & \text{on } E_1, \\ (x_1, \cdots, x_{j-1}, x_i, x_{j+1}, \cdots, x_K) & \text{on } E_2, \end{cases}$$

where H follows the mutation law $m(x_i, \mathrm{d}h)$, E_1 is the event "a mutation occurs at time ϑ_1" and E_2 is the event "a resampling between i and j occurs at time ϑ_1". Note that $\mu_{\vartheta_1}^K = \tau_{-\left\langle \mathrm{id}, \frac{1}{K}\sum_{k=1}^K \delta_{y_k}\right\rangle}\sharp\left(\frac{1}{K}\sum_{k=1}^K \delta_{y_k}\right)$. Hence, noting that for all $k \in \{1, \cdots, K\}$,

$$|x_k| \leqslant \sqrt{K}\left(1 + M_2\left(\mu_0^K\right)\right)$$

and using Assumption **(A2)**, for K large enough we have that

$$
\begin{aligned}
M_2\left(\mu_{\vartheta_1}^K\right) &\leqslant M_2\left(\frac{1}{K}\sum_{i=1}^K \delta_{y_i}\right) = \frac{1}{K}\sum_{i=1}^K y_i^2 \\
&\leqslant \begin{cases} \dfrac{1}{K}\displaystyle\sum_{k=1}^K x_k^2 + \dfrac{2A_m x_k}{K^{\frac{3}{2}}} + \dfrac{A_m^2}{K^2} & \text{on} \quad E_1, \\ \dfrac{1}{K}\displaystyle\sum_{k=1}^K x_k^2 + \dfrac{x_j^2 - x_i^2}{K} & \text{on} \quad E_2, \end{cases} \\
&\leqslant \begin{cases} M_2\left(\mu_0^K\right) + \dfrac{2A_m\left(1 + M_2\left(\mu_0^K\right)\right)}{K} + \dfrac{A_m^2}{K^2} & \text{on} \quad E_1, \\ 2M_2\left(\mu_0^K\right) & \text{on} \quad E_2, \end{cases} \\
&\leqslant 2M_2\left(\mu_0^K\right) + 1.
\end{aligned}
$$

□

Each of the previous steps will be called *transitions of* $M_2\left(\mu_t^K\right)$. The idea of the proof of Lemma 4.9 is to estimate the transition probabilities of the sequence $\left(L_k^K\right)_{k\in\mathbb{N}}$ in order to construct a coupling between $\left(L_k^K\right)_{k\in\mathbb{N}}$ and a biased random walk on $\mathbb{N}^\star$ reflected in 1.

At time $\widehat{\tau}^K$, $M_2\left(\mu_{\widehat{\tau}^K}^K\right) \in \mathcal{I}_{\left\lfloor \frac{\varepsilon}{2\log(3)}\log(K)\right\rfloor}$ where $\lfloor x \rfloor$ is the lower integer part of x. So, the number of transitions of $M_2\left(\mu_t^K\right)$ before $\widehat{\tau}^K$ is greater than or equal to $T_{k_0}^K$ where k_0 is the first integer such that $L_{k_0}^K = \left\lfloor \frac{\varepsilon}{2\log(3)}\log(K)\right\rfloor$.

By using estimates on the number of steps that a biased random walk takes to reach $\left\lfloor \frac{\varepsilon}{2\log(3)}\log(K)\right\rfloor$ and estimates on the durations between two transitions $T_{k+1}^K - T_k^K$, we will deduce a lower bound, exponential in K, on $\widehat{\tau}^K$. An additional difficulty comes from the fact that the previous coupling argument is only valid up to time $\check{\tau}^K$. Hence we will construct a coupling that takes into account the possibility that $\check{\tau}^K$ happens during each transition step. The proof is divided into four steps: in Section 4.4.1, we characterise the behaviour of the first transition step; in Section 4.4.2, the proposed coupling is constructed; in Section 4.4.3, we give estimates on the first exit time from an attractive domain for random walks; finally, we conclude in Section 4.4.4.

4.4.1 One-step transitions

In this section we look at only one transition: we suppose that $M_2\left(\mu_0^K\right) \in [2u_{\ell-1}+1, 2u_\ell+1)$ for $\ell \in \mathbb{N}^\star$ fixed and we look for bounds on the first transition probabilities of $M_2\left(\mu_t^K\right)$ and on time T_1^K. The main result is the following.

Lemma 4.11. *For all $\ell \in \mathbb{N}$, for all $\mu_0^K \in \mathcal{M}_{1,K}^{c,2}(\mathbb{R})$ such that*

$$M_2\left(\mu_0^K\right) \in [2u_{\ell-1}+1, 2u_\ell+1),$$

there exist a constant $C > 0$, $K_0 \in \mathbb{N}^\star$ such that for all $K \geqslant K_0$, we have

$$\mathbb{P}_{\mu_0^K}\left(\left[\left\{L_1^K = \ell - 1\right\} \cap \left\{\check{\tau}^K > T_1^K \geqslant K\sigma_K^2\right\}\right] \cup \left\{T_1^K \geqslant \check{\tau}^K\right\}\right) \geqslant \frac{1}{2} + \eta_K, \tag{35}$$

where $\eta_K := \frac{1}{2} - \frac{C}{K^{\varepsilon/2}}$.

The proof of the previous result is based on Corollary 4.12 which is obtained as a straightforward consequence of Lemma 4.5.

Corollary 4.12. *Let* $\ell \in \mathbb{N}^\star$ *be fixed. There exists* $K_0 \in \mathbb{N}^\star$ *large enough such that for all* $K \geqslant K_0$*, for all* $\mu_0^K \in \mathcal{M}_{1,K}^{c,2}(\mathbb{R})$ *such that* $M_2\left(\mu_0^K\right) \in [2u_{\ell-1}+1, 2u_\ell+1)$*, we have for all* $t \leqslant T_1^K$ *that*

$$2u_{\ell-1}+1-\frac{5\bar{b}u_{\ell+1}+\underline{\theta}m_2}{2K^2\sigma_K^2}t\wedge\check{\tau}^K+M_{t\wedge\check{\tau}^K}^{K,P_{\mathrm{id}^2,1}}$$
$$\leqslant M_2\left(\mu_{t\wedge\check{\tau}^K}^K\right)\leqslant 2u_\ell+1-\frac{\underline{b}u_{\ell-1}}{K^2\sigma_K^2}t\wedge\check{\tau}^K+M_{t\wedge\check{\tau}^K}^{K,P_{\mathrm{id}^2,1}}.$$

Proof of Lemma 4.11. By passing on the complementary of (35), it is equivalent to prove that

$$\mathbb{P}_{\mu_0^K}\left(\left[\left\{L_1^K=\ell+1\right\}\cup\left\{T_1^K<K\sigma_K^2\right\}\right]\cap\left\{T_1^K<\check{\tau}^K\right\}\right)<\frac{1}{2}-\eta_K.$$

As for all events A, B, $\mathbb{P}(A\cup B)=\mathbb{P}(A\cap B^c)+\mathbb{P}(B)$, we have

$$\begin{aligned}
&\mathbb{P}_{\mu_0^K}\left(\left[\left\{L_1^K=\ell+1\right\}\cup\left\{T_1^K<K\sigma_K^2\right\}\right]\cap\left\{T_1^K<\check{\tau}^K\right\}\right)\\
&\quad=\mathbb{P}_{\mu_0^K}\left(\left\{T_1^K<\check{\tau}^K\right\}\cap\left\{L_1^K=\ell+1\right\}\cap\left\{T_1^K\geqslant K\sigma_K^2\right\}\right)\\
&\qquad+\mathbb{P}_{\mu_0^K}\left(\left\{T_1^K<\check{\tau}^K\right\}\cap\left\{T_1^K<K\sigma_K^2\right\}\right)\\
&\quad\leqslant\mathbb{P}_{\mu_0^K}\left(\left\{T_1^K<\check{\tau}^K\right\}\cap\left\{L_1^K=\ell+1\right\}\right)+\mathbb{P}_{\mu_0^K}\left(\left\{T_1^K<\check{\tau}^K\right\}\cap\left\{T_1^K<K\sigma_K^2\right\}\right).
\end{aligned}\tag{36}$$

Step 1. Control of the first right-hand term of (36). Let us consider the martingale $\left(\mathrm{Mart}_t^{K,+}\right)_{t\geqslant 0}$ defined by

$$\mathrm{Mart}_t^{K,+}:=\begin{cases}M_t^{K,P_{\mathrm{id}^2,1}} & \text{if} \quad t\leqslant\check{\tau}^K\\ M_{\check{\tau}^K}^{K,P_{\mathrm{id}^2,1}}+\frac{1}{K\sigma_K}B_{t-\check{\tau}^K} & \text{if} \quad t>\check{\tau}^K\end{cases}$$

where $(B_t)_{t\geqslant 0}$ is a standard Brownian motion independent of $\left(\mu_t^K\right)_{t\geqslant 0}$. Note that from Lemma 4.7, there exists a constant $C>0$ such that for all $t\geqslant 0$ we have

$$\mathbb{E}\left(\left\langle\mathrm{Mart}^{K,+}\right\rangle_t\right)\leqslant\frac{C}{K^2\sigma_K^2}t.\tag{37}$$

Let us consider the process $\left(M_2^{K,+}(t)\right)_{t\geqslant 0}$ defined by

$$M_2^{K,+}(t):=\begin{cases}M_2\left(\mu_t^K\right) & \text{if} \quad t\leqslant\check{\tau}^K\\ M_2\left(\mu_{\check{\tau}^K}^K\right)-\frac{\underline{b}u_{\ell-1}}{K^2\sigma_K^2}\left(t-\check{\tau}^K\right)+\left(\mathrm{Mart}_t^{K,+}-\mathrm{Mart}_{\check{\tau}^K}^{K,+}\right) & \text{if} \quad t>\check{\tau}^K\end{cases}.$$

Note that from Corollary 4.12, for all $t\geqslant 0$,

$$2u_{\ell-1}+1-\frac{5\bar{b}u_{\ell+1}+\underline{\theta}m_2}{2K^2\sigma_K^2}t+\mathrm{Mart}_t^{K,+}\leqslant M_2^{K,+}(t)\leqslant 2u_\ell+1-\frac{\underline{b}u_{\ell-1}}{K^2\sigma_K^2}t+\mathrm{Mart}_t^{K,+}.\tag{38}$$

By Girsanov's theorem, the process $\left(E_t^{K,+}\right)_{t\geqslant 0}$ defined for all $t\geqslant 0$ by

$$E_t^{K,+} := \exp\left(\mathrm{Mart}_t^{K,+} - \frac{1}{2}\left\langle \mathrm{Mart}^{K,+}\right\rangle_t\right)$$

is a local martingale. Let $t_0^K := \frac{6K^2\sigma_K^2}{\underline{b}}$ be fixed and let us consider

$$T_1^{K,+} := \inf\left\{t\geqslant 0 \,\middle|\, M_2^{K,+}(t)\notin \mathcal{I}_\ell\right\},$$
$$L_1^{K,+} := \min\left\{\ell\in\mathbb{N} \,\middle|\, M_2^{K,+}\left(T_1^{K,+}\right)\leqslant 2u_\ell - 1\right\}.$$

From (38) and denoting $\Gamma := \left\{\left\langle \mathrm{Mart}^{K,+}\right\rangle_{T_1^{K,+}\wedge t_0^K} \leqslant \frac{2\bar{b}u_{\ell-1}}{K^2\sigma_K^2}T_1^{K,+}\wedge t_0^K\right\}$, we have

$$\begin{aligned}
1 &= \mathbb{E}_{\mu_0^K}\left(E^{K,+}_{T_1^{K,+}\wedge t_0^K}\right) \geqslant \mathbb{E}_{\mu_0^K}\left(E^{K,+}_{T_1^{K,+}\wedge t_0^K}\mathbb{1}_\Gamma\right)\\
&\geqslant \exp(-2u_\ell - 1)\,\mathbb{E}_{\mu_0^K}\left(\exp\left(M_2^{K,+}\left(T_1^{K,+}\wedge t_0^K\right) + \frac{\underline{b}u_{\ell-1}}{K^2\sigma_K^2}T_1^{K,+}\wedge t_0^K\right.\right.\\
&\qquad\qquad\left.\left. - \frac{1}{2}\left\langle \mathrm{Mart}^{K,+}\right\rangle_{T_1^{K,+}\wedge t_0^K}\right)\mathbb{1}_\Gamma\right)\\
&\geqslant \exp(-2u_\ell - 1)\,\mathbb{E}_{\mu_0^K}\left(\exp\left(M_2^{K,+}\left(T_1^{K,+}\wedge t_0^K\right)\right)\mathbb{1}_\Gamma\right)\\
&\geqslant \exp(-2u_\ell - 1)\left[\exp(u_{\ell+1})\,\mathbb{P}_{\mu_0^K}\left(\left\{L_1^{K,+} = \ell+1\right\}\cap\Gamma\cap\left\{T_1^{K,+}\leqslant t_0^K\right\}\right)\right.\\
&\qquad\qquad\left. + \mathbb{P}_{\mu_0^K}\left(\left\{L_1^{K,+}\leqslant \ell-1\right\}\cap\Gamma\cap\left\{T_1^{K,+}\leqslant t_0^K\right\}\right)\right].
\end{aligned}$$

As for all events A, B, C, $\mathbb{P}(A\cap B\cap C)\geqslant \mathbb{P}(A) - \mathbb{P}(A\cap B^c) - \mathbb{P}(A\cap C^c)$, and denoting $q := \mathbb{P}_{\mu_0^K}\left(L_1^{K,+} = \ell+1\right)$ we have that

$$\begin{aligned}
1 &\geqslant \exp(u_\ell - 1)\left[q - \mathbb{P}_{\mu_0^K}\left(\left\{L_1^{K,+} = \ell+1\right\}\cap\Gamma^c\right)\right.\\
&\qquad\left. - \mathbb{P}_{\mu_0^K}\left(\left\{L_1^{K,+} = \ell+1\right\}\cap\left\{T_1^{K,+} > t_0^K\right\}\right)\right]\\
&\quad + \exp(-2u_\ell - 1)\left[1 - q - \mathbb{P}_{\mu_0^K}\left(\left\{L_1^{K,+}\leqslant \ell-1\right\}\cap\Gamma^c\right)\right.\\
&\qquad\left. - \mathbb{P}_{\mu_0^K}\left(\left\{L_1^{K,+}\leqslant \ell-1\right\}\cap\left\{T_1^{K,+} > t_0^K\right\}\right)\right]\\
&\geqslant \left(\exp(u_\ell - 1) - \exp(-2u_\ell - 1)\right)q + \exp(-2u_\ell - 1)\\
&\qquad - \left(\exp(u_\ell - 1) + \exp(-2u_\ell - 1)\right)\left[\mathbb{P}_{\mu_0^K}(\Gamma^c) + \mathbb{P}_{\mu_0^K}\left(T_1^{K,+} > t_0^K\right)\right].
\end{aligned}$$

Then, considering $s_0^K := \frac{u_{\ell-1}K^2\sigma_K^2}{5\bar{b}u_{\ell+1} + \underline{\theta}\underline{m}_2} < t_0^K$ be fixed, we deduce that

$$\begin{aligned}
q &\leqslant \frac{1-\exp(-2u_\ell - 1)}{\exp(u_\ell - 1) - \exp(-2u_\ell - 1)}\\
&\qquad + \frac{\exp(u_\ell) + \exp(-2u_\ell)}{\exp(u_\ell) - \exp(-2u_\ell)}\left[\mathbb{P}_{\mu_0^K}(\Gamma^c) + \mathbb{P}_{\mu_0^K}\left(T_1^{K,+} > t_0^K\right)\right]\\
&\leqslant \exp(-u_\ell + 1)\times\frac{1-\exp(-2u_\ell - 1)}{1-\exp(-u_{\ell+1})} + \frac{\exp(u_\ell) + \exp(-2u_\ell)}{\exp(u_\ell) - \exp(-2u_\ell)}\\
&\qquad\times\left[\mathbb{P}_{\mu_0^K}\left(\Gamma^c\cap\left\{T_1^{K,+}\geqslant s_0^K\right\}\right) + \mathbb{P}_{\mu_0^K}\left(T_1^{K,+} < s_0^K\right) + \mathbb{P}_{\mu_0^K}\left(T_1^{K,+} > t_0^K\right)\right].
\end{aligned}$$

Denoting $\lceil a \rceil$ the upper integer part of $a \in \mathbb{R}$, we obtain on the one hand that there exists a constant $C_1 > 0$ such that

$$
\begin{aligned}
&\mathbb{P}_{\mu_0^K}\left(\Gamma^c \cap \left\{T_1^{K,+} \geqslant s_0^K\right\}\right) \\
&= \mathbb{P}_{\mu_0^K}\left(\left\{\left\langle \mathrm{Mart}^{K,+}\right\rangle_{T_1^{K,+}\wedge t_0^K} > \frac{2\underline{b}u_{\ell-1}}{K^2\sigma_K^2} T_1^{K,+} \wedge t_0^K\right\} \cap \left\{T_1^{K,+} \geqslant s_0^K\right\}\right) \\
&= \sum_{k=1}^{\left\lceil \frac{t_0^K}{s_0^K}\right\rceil - 1} \mathbb{P}_{\mu_0^K}\left(\left\{\left\langle \mathrm{Mart}^{K,+}\right\rangle_{T_1^{K,+}\wedge t_0^K} > \frac{2\underline{b}u_{\ell-1}}{K^2\sigma_K^2} T_1^{K,+} \wedge t_0^K\right\} \cap \left\{ks_0^K \leqslant T_1^{K,+} \leqslant (k+1)s_0^K\right\}\right) \\
&\qquad + \mathbb{P}_{\mu_0^K}\left(\left\{\left\langle \mathrm{Mart}^{K,+}\right\rangle_{T_1^{K,+}\wedge t_0^K} > \frac{2\underline{b}u_{\ell-1}}{K^2\sigma_K^2} T_1^{K,+} \wedge t_0^K\right\} \cap \left\{T_1^{K,+} \geqslant \lceil t_0^K/s_0^K\rceil s_0^K\right\}\right) \\
&\leqslant \sum_{k=1}^{\left\lceil \frac{t_0^K}{s_0^K}\right\rceil - 1} \mathbb{P}_{\mu_0^K}\left(\left\langle \mathrm{Mart}^{K,+}\right\rangle_{(k+1)s_0^K} > \frac{2\underline{b}u_{\ell-1}}{K^2\sigma_K^2} k s_0^K\right) + \mathbb{P}_{\mu_0^K}\left(\left\langle \mathrm{Mart}^{K,+}\right\rangle_{t_0^K} > \frac{2\underline{b}u_{\ell-1}}{K^2\sigma_K^2} t_0^K\right) \\
&\leqslant K^2\sigma_K^2 \left(\sum_{k=1}^{\lceil t_0^K/s_0^K\rceil - 1} \frac{\mathbb{E}\left(\left\langle \mathrm{Mart}^{K,+}\right\rangle_{(k+1)s_0^K}\right)}{2\underline{b}u_{\ell-1}ks_0^K} + \frac{\mathbb{E}\left(\left\langle \mathrm{Mart}^{K,+}\right\rangle_{t_0^K}\right)}{2\underline{b}u_{\ell-1}t_0^K}\right) \\
&\leqslant \frac{C_1}{K^{\frac{\varepsilon}{2}}}
\end{aligned}
$$

where we use Markov's inequality in the second inequality and (38) in the last inequality. On the other hand, we control $\mathbb{P}_{\mu_0^K}\left(T_1^{K,+} < s_0^K\right)$ as follows:

$$
\begin{aligned}
\mathbb{P}_{\mu_0^K}\left(T_1^{K,+} < s_0^K\right) \leqslant\ & \mathbb{P}_{\mu_0^K}\left(\left\{T_1^{K,+} < s_0^K\right\} \cap \left\{\sup_{0\leqslant t\leqslant s_0^K} M_2^{K,+}(t) \geqslant u_{\ell+1}\right\}\right) \\
& + \mathbb{P}_{\mu_0^K}\left(\left\{T_1^{K,+} < s_0^K\right\} \cap \left\{\inf_{0\leqslant t\leqslant s_0^K} M_2^{K,+}(t) < u_{\ell-1}\right\}\right).
\end{aligned}
$$

Let us consider $K_1 \in \mathbb{N}^\star$ independent of ℓ satisfying $K_1^{\varepsilon/2} \geqslant 2$. There exists a constant $C_1 > 0$

such that for all $K \geqslant K_1$, we have that

$$\begin{aligned}
&\mathbb{P}_{\mu_0^K}\left(\left\{T_1^{K,+} < s_0^K\right\} \cap \left\{\sup_{0\leqslant t\leqslant s_0^K} M_2^{K,+}(t) \geqslant u_{\ell+1}\right\}\right) \\
&\qquad \leqslant \mathbb{P}_{\mu_0^K}\left(\left\{T_1^{K,+} < s_0^K\right\} \cap \left\{\sup_{0\leqslant t\leqslant s_0^K} \left\{2u_\ell + 1 - \frac{\underline{b}u_{\ell-1}}{K^2\sigma_K^2}t + \mathrm{Mart}_t^{K,+}\right\} \geqslant u_{\ell+1}\right\}\right) \\
&\qquad \leqslant \mathbb{P}_{\mu_0^K}\left(\left\{T_1^{K,+} < s_0^K\right\} \cap \left\{\sup_{0\leqslant t\leqslant s_0^K} \mathrm{Mart}_t^{K,+} \geqslant u_\ell - 1\right\}\right) \\
&\qquad \leqslant \mathbb{P}_{\mu_0^K}\left(\sup_{0\leqslant t\leqslant s_0^K} \left|\mathrm{Mart}_t^{K,+}\right| \geqslant \frac{3^\ell K^{\frac{\varepsilon}{2}}}{2}\right) \\
&\qquad \leqslant 4\mathbb{E}_{\mu_0^K}\left(\left\langle \mathrm{Mart}^{K,+}\right\rangle_{s_0^K}\right) \times \frac{4}{3^{2\ell}K^\varepsilon} \\
&\qquad \leqslant \frac{C_1}{K^\varepsilon}
\end{aligned}$$

where we use (38) in the first inequality, Doob's maximal inequality in the fourth inequality and (37) in the last inequality. Now, let us consider $K_2 \in \mathbb{N}^\star$ independent of ℓ satisfying $K_2^{\varepsilon/2} \geqslant 4$. Using again (38), Doob's maximal inequality, (37), and the definition of s_0^K, we deduce that there exists a constant $C_2 > 0$ such that for all $K \geqslant K_2$, we have that

$$\begin{aligned}
&\mathbb{P}_{\mu_0^K}\left(\left\{T_1^{K,+} < s_0^K\right\} \cap \left\{\inf_{0\leqslant t\leqslant s_0^K} M_2^{K,+}(t) < u_{\ell-1}\right\}\right) \\
&\leqslant \mathbb{P}_{\mu_0^K}\left(\left\{T_1^{K,+} < s_0^K\right\} \cap \left\{\inf_{0\leqslant t\leqslant s_0^K} \left\{2u_{\ell-1} - 1 - \frac{5\bar{b}u_{\ell+1} + \theta m_2}{2K^2\sigma_K^2}t + \mathrm{Mart}_t^{K,+}\right\} < u_{\ell-1}\right\}\right) \\
&\leqslant \mathbb{P}_{\mu_0^K}\left(\left\{T_1^{K,+} < s_0^K\right\} \cap \left\{\inf_{0\leqslant t\leqslant s_0^K} \mathrm{Mart}_t^{K,+} < -\frac{u_{\ell-1}}{2} + 1\right\}\right) \\
&\leqslant \mathbb{P}_{\mu_0^K}\left(\sup_{0\leqslant t\leqslant s_0^K} \left|\mathrm{Mart}_t^{K,+}\right| > \frac{u_{\ell-1}}{2} - 1\right) \\
&\leqslant \frac{C_2}{K^\varepsilon}.
\end{aligned}$$

Hence, for all $K \geqslant \max\{K_1, K_2\}$, $\mathbb{P}_{\mu_0^K}\left(T_1^{K,+} < s_0^K\right) \leqslant \frac{C_1+C_2}{K^\varepsilon}$. Finally, from (38) and Markov's inequality and the definition of t_0^K, note that there exists a constant $C_3 > 0$ such that

$$\begin{aligned}
\mathbb{P}_{\mu_0^K}\left(T_1^{K,+} > t_0^K\right) &\leqslant \mathbb{P}_{\mu_0^K}\left(\left\{\sup_{0\leqslant t\leqslant t_0^K} M_2^{K,+}(t) < u_{\ell+1}\right\} \cap \left\{\inf_{0\leqslant t\leqslant t_0^K} M_2^{K,+}(t) \geqslant u_{\ell-1}\right\}\right) \\
&\leqslant \mathbb{P}_{\mu_0^K}\left(\mathrm{Mart}_{t_0^K}^{K,+} \geqslant u_{\ell-1} - 2u_\ell - 1 + \frac{\underline{b}u_{\ell-1}}{K^2\sigma_K^2}t_0^K\right) \\
&\leqslant \frac{\mathbb{E}_{\mu_0^K}\left(\left\langle \mathrm{Mart}^{K,+}\right\rangle_{t_0^K}\right)}{\left(u_{\ell-1}-1\right)^2} \\
&\leqslant \frac{C_3}{K^\varepsilon}.
\end{aligned}$$

Hence, for all $K \geqslant \max\{K_1, K_2\}$, we deduce that $q \leqslant \varepsilon_K$ where

$$\varepsilon_K := \exp(-u_\ell + 1) \times \frac{1 - \exp(-2u_\ell - 1)}{1 - \exp(-u_{\ell+1})} + \frac{\exp(u_\ell) + \exp(-2u_\ell)}{\exp(u_\ell) - \exp(-2u_\ell)} \times \frac{C_1}{K^{\frac{\varepsilon}{2}}}$$

and thus, for all $K \geqslant \max\{K_1, K_2\}$,

$$\mathbb{P}_{\mu_0^K}\left(\left\{T_1^K < \check{\tau}^K\right\} \cap \left\{L_1^K = \ell + 1\right\}\right) = \mathbb{P}_{\mu_0^K}\left(\left\{T_1^{K,+} < \check{\tau}^K\right\} \cap \left\{L_1^K = \ell + 1\right\}\right) \leqslant q \leqslant \varepsilon_K.$$

Step 2. Control of the second right-hand term of (36). Similarly, let us consider $K_3 \in \mathbb{N}^\star$ independent of ℓ satisfying $K_3^{\varepsilon/2} \geqslant 2$. Using Corollary 4.12, Doob's maximal inequality and Lemma 4.7 in similar way to Step 1, there exists a constant $C_4 > 0$ such that for all $K \geqslant K_3$ we have that

$$\begin{aligned}\mathbb{P}_{\mu_0^K}\left(\left\{T_1^K < \check{\tau}^K\right\} \cap \left\{\sup_{0 \leqslant t \leqslant K\sigma_K^2 \wedge \check{\tau}^K} M_2\left(\mu_t^K\right) \geqslant u_{\ell+1}\right\}\right) &\leqslant \frac{16\mathbb{E}_{\mu_0^K}\left(\left\langle M^{K,P_{\mathrm{id}^2,1}}\right\rangle_{K\sigma_K^2 \wedge \check{\tau}^K}\right)}{3^{2\ell}K^\varepsilon} \\ &\leqslant \frac{C_4}{K^{1+\varepsilon}}.\end{aligned}$$

Let us consider $K_4 \in \mathbb{N}^\star$ independent of ℓ satisfying the relation $u_{\ell-1} > 1 + \frac{5\bar{b}u_{\ell+1} + \theta m_2}{2K_4}$. As previously, we establish that there exists a constant $C_5 > 0$ such that for all $K \geqslant K_4$ we have that

$$\begin{aligned}\mathbb{P}_{\mu_0^K}\left(\left\{T_1^K < \check{\tau}^K\right\} \cap \left\{\inf_{0 \leqslant t \leqslant K\sigma_K^2 \wedge \check{\tau}^K} M_2\left(\mu_t^K\right) < u_{\ell-1}\right\}\right) &\leqslant \frac{4\mathbb{E}_{\mu_0^K}\left(\left\langle M^{K,P_{\mathrm{id}^2,1}}\right\rangle_{K\sigma_K^2 \wedge \check{\tau}^K}\right)}{\left(u_{\ell-1} - 1 - \frac{5\bar{b}u_{\ell+1} + \theta m_2}{2K}\right)^2} \\ &\leqslant \frac{C_5}{K^{1+\varepsilon}}.\end{aligned}$$

To conclude this proof, there exists a constant $C > 0$ such that for all $K \geqslant K_0$ where $K_0 := \max_{i \in \{1,\cdots,4\}} K_i$,

$$\mathbb{P}_{\mu_0^K}\left(\left[\left\{L_1^K = \ell + 1\right\} \cup \left\{T_1^K < K\sigma_K^2\right\}\right] \cap \left\{T_1^K < \check{\tau}^K\right\}\right) < \frac{1}{2} - \eta_K$$

for any $\eta_K \geqslant \frac{1}{2} - \epsilon_K - \frac{C_4 + C_5}{K^{1+\varepsilon}}$. A convenient choice is given by $\eta_K := \frac{1}{2} - \frac{C}{K^{\varepsilon/2}}$ for K large enough independent of ℓ which completes the proof. □

4.4.2 Construction of the coupling

The goal of this section is to construct a coupling between $\left(L_k^K\right)_{k \in \mathbb{N}}$ and a biased random walk on $\mathbb{N}^\star$ and reflected in 1. To do this, we will construct a sequence $\left(\xi_k^K\right)_{k \in \mathbb{N}}$ of i.i.d. random variables with values in $\{-1, 1\}$ as follows.

Step 1. First of all, thanks to Lemma 4.11 we can construct the random variable $\xi_0^K \in \{-1, 1\}$ such that

$$\left\{\xi_0^K = -1\right\} \subset \left[\left\{L_1^K = L_0^K - 1\right\} \cap \left\{K\sigma_K^2 \leqslant T_1^K < \check{\tau}^K\right\}\right] \cup \left\{T_1^K \geqslant \widehat{\tau}^K\right\}$$

and $\mathbb{P}\left(\xi_0^K = -1\right) = \frac{1}{2} + \eta_K$.

Step 2. Then, thanks to Lemma 4.11 again and after having applied the MARKOV property at time T_1^K, we can construct the random variable $\xi_1^K \in \{-1, 1\}$ such that, conditionally to $\mathcal{F}_{T_1^K} \vee \sigma\left(\xi_0^K\right)$,

$$\left\{\xi_1^K = -1\right\} \subset \left[\left\{L_2^K = L_1^K - 1\right\} \cap \left\{K\sigma_K^2 + T_1^K \leqslant T_2^K < \check{\tau}^K\right\}\right] \cup \left\{T_2^K \geqslant \widehat{\tau}^K\right\}$$

and $\mathbb{P}\left(\xi_1^K = -1 \,\middle|\, \mathcal{F}_{T_1^K} \vee \sigma\left(\xi_0^K\right)\right) = \frac{1}{2} + \eta_K$. Note that this implies in particular that ξ_1^K is independent of ξ_0^K.

Step 3. By induction on $k \in \mathbb{N}$, if $\left(\xi_i^K\right)_{0 \leqslant i \leqslant k}$ are constructed, then thanks to Lemma 4.11 again and after having applied the MARKOV property at time T_{k+1}^K, we can construct the random variable $\xi_{k+1}^K \in \{-1, 1\}$ such that, conditionally to $\mathcal{F}_{T_{k+1}^K} \vee \sigma\left(\xi_0^K, \cdots, \xi_k^K\right)$,

$$\left\{\xi_{k+1}^K = -1\right\} \subset \left[\left\{L_{k+2}^K = L_{k+1}^K - 1\right\} \cap \left\{K\sigma_K^2 + T_{k+1}^K \leqslant T_{k+2}^K < \check{\tau}^K\right\}\right] \cup \left\{T_{k+2}^K \geqslant \widehat{\tau}^K\right\}$$

and $\mathbb{P}\left(\xi_1^K = -1 \,\middle|\, \mathcal{F}_{T_{k+1}^K} \vee \sigma\left(\xi_0^K, \cdots, \xi_k^K\right)\right) = \frac{1}{2} + \eta_K$.

To conclude, we set

$$\forall K \in \mathbb{N}^\star, \quad \forall k \in \mathbb{N}, \qquad Z_{k+1}^K - Z_k^K := \xi_k^K \qquad \text{and} \qquad Z_0^K = L_0^K \tag{39}$$

and the sequence $\left(Z_k^K\right)_{k \in \mathbb{N}}$ satisfies by construction the following lemma:

Lemma 4.13. *The sequence $\left(Z_k^K\right)_{k \in \mathbb{N}}$, given by* (39) *is a biased simple random walk on $\mathbb{N}^\star$, reflected in* 1 *and for all $k \in \mathbb{N}$ as long as $T_{k+1}^K < \check{\tau}^K$, that*

(1) $Z_{k+1}^K - Z_k^K \geqslant L_{k+1}^K - L_k^K$

(2) $T_{k+1}^K - T_k^K \geqslant K\sigma_K^2$ when $Z_{k+1}^K - Z_k^K = -1$.

On the one hand, note that for all $k \in \mathbb{N}$, $L_k^K \leqslant Z_k^K$. On the other hand, note that Lemma 4.10 implies for all $k \in \mathbb{N}$, $L_{k+1}^K - L_k^K \in \{\cdots, -3, -2, -1, 1\}$ which justifies (1) of Lemma 4.13. In Figure 4.7, we illustrate the coupling between $\left(L_k^K\right)_{0 \leqslant k \leqslant 7}$ and $\left(Z_k^K\right)_{0 \leqslant k \leqslant 7}$. In this figure, we observe that $\xi_k^K = L_{k+1}^K - L_k^K$ for all $k \in \{0, 1, 2, 4, 5, 6\}$ and $\xi_3^K = -1 \geqslant -2 = L_4^K - L_3^K$. This illustrates that $Z_{k+1}^K - Z_k^K$ is always 1 when $L_{k+1}^K - L_k^K$ is equal to 1. Note that $Z_{k+1}^K - Z_k^K$ can be 1 when $L_{k+1}^K - L_k^K$ is equal to -1.

4.4.3 Estimates of exit from an attractive domain for random walks

Let us consider N^K the number of transitions before reaching $\left\lfloor \frac{\varepsilon}{2\log(3)} \log(K) \right\rfloor$ for the random walk $\left(Z_k^K\right)_{k \in \mathbb{N}}$.

Remark 4.14. *By* Lemma 4.13*, $N^K \leqslant k_0$ where k_0 is the first integer such that $L_{k_0}^K = \left\lfloor \frac{\varepsilon}{2\log(3)} \log(K) \right\rfloor$, so $\widehat{\tau}^K \geqslant T_{k_0}^K \geqslant T_{N^K}^K$.*

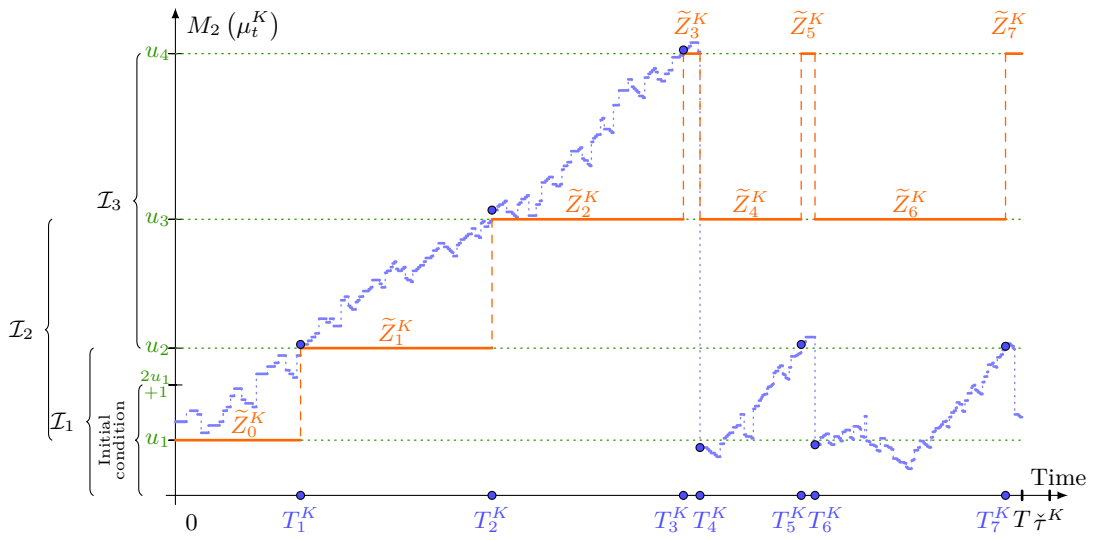

Figure 4.7: For K large enough, coupling between $\left(L_k^K\right)_{0\leqslant k\leqslant 7}$ and $\left(Z_k^K\right)_{0\leqslant k\leqslant 7}$ up to time T and before $\check{\tau}^K$ and where $M_2\left(\mu_0^K\right)\in[0,2u_1+1)$ so that $L_0^K=1$ and $\widetilde{Z}_k^K:=u_{Z_k^K}=3^{Z_k^K}K^{\varepsilon/2}$ for all $k\in\mathbb{N}$.

The following lemma gives an estimate on the problem of exit from a domain for Z^K. For all $k\in\mathbb{N}^\star$, we denote by $\mathbf{P}_k$ the law of the Markov chain Z^K given $Z_0^K=L_0^K=k$.

Lemma 4.15. *We have,*

$$\lim_{K\to+\infty}\mathbf{P}_1\left(N^K\geqslant\exp\left(\frac{\varepsilon^2}{16\log(3)}\log^2(K)\right)\right)=1.$$

Proof. For all $k\in\mathbb{N}^\star$, let us consider the stopping time $\mathcal{T}_k^K:=\inf\left\{n\in\mathbb{N}\,\middle|\,Z_n^K=k\right\}$ and we set $v_k^K:=\mathbf{P}_k\left(\mathcal{T}_1^K<\mathcal{T}^K_{\left\lfloor\frac{\varepsilon}{2\log(3)}\log(K)\right\rfloor}\right)$. Note that $v_1^K=1$. By a classical approach and setting $r_K:=\frac{1+2\eta_K}{1-2\eta_K}>1$, we prove that

$$\begin{aligned}v_k^K&=\frac{1-r_K^{-\left(\left\lfloor\frac{\varepsilon}{2\log(3)}\log(K)\right\rfloor-k\right)}}{1-r_K^{-\left(\left\lfloor\frac{\varepsilon}{2\log(3)}\log(K)\right\rfloor-1\right)}}\\&=1-\left(r_K^{k-1}-1\right)\exp\left(-\left(\left\lfloor\frac{\varepsilon}{2\log(3)}\log(K)\right\rfloor-1\right)\log(r_K)\right).\end{aligned}$$

Note that N^K is greater than the number of transitions from 2 to 1 that occur before $\mathcal{T}^K_{\left\lfloor\frac{\varepsilon}{2\log(3)}\log(K)\right\rfloor}$. Hence, under $\mathbf{P}_1$, $N^K\geqslant X$ where X is random variable with geometric law of parameter $q_K:=(r_K-1)\exp\left(-\left(\left\lfloor\frac{\varepsilon}{2\log(3)}\log(K)\right\rfloor-1\right)\log(r_K)\right)$. Since $r_K\sim_{K\to+\infty}\frac{K^{\varepsilon/2}}{C}$, with $C>0$ given by Lemma 4.11, $q_K\to0$ when $K\to+\infty$ and using that for all $x>0$ small enough,

$\log(1-x) \geqslant -2x$ we deduce that for all K large enough,

$$\begin{aligned}
\mathbf{P}_1 &\left(X \geqslant \exp\left(\frac{\left\lfloor \frac{\varepsilon}{2\log(3)} \log(K) \right\rfloor - 1}{2} \log\left(r_K\right) \right) \right) \\
&= \exp\left(\exp\left(\frac{\left\lfloor \frac{\varepsilon}{2\log(3)} \log(K) \right\rfloor - 1}{2} \log\left(r_K\right) \right) \log\left(1-q_K\right) \right) \\
&\geqslant \exp\left(-2\left(r_K - 1\right) \exp\left(-\frac{\left\lfloor \frac{\varepsilon}{2\log(3)} \log(K) \right\rfloor - 1}{2} \log\left(r_K\right) \right) \right).
\end{aligned}$$

Using again that $r_K \sim_{K\to+\infty} \frac{K^{\varepsilon/2}}{C}$, we deduce that for K large enough,

$$\frac{\left\lfloor \frac{\varepsilon}{2\log(3)} \log(K) \right\rfloor - 1}{2} \log\left(r_K\right) \geqslant \frac{\varepsilon^2}{16\log(3)} \log^2\left(K\right)$$

and the announced result follows. □

It is in the next Corollary 4.16 that we see the importance of Assumption (5) where σ_K must not be too small.

Corollary 4.16. *For all $T > 0$,*

$$\lim_{K\to+\infty} \mathbb{P}\left(\frac{K\sigma_K^2}{2} \left(N^K - \frac{\varepsilon}{2\log(3)} \log(K) \right) < T \right) = 0.$$

Proof. From Lemma 4.15 and Assumption (5) we deduce in a straightfoward manner that

$$\lim_{K\to+\infty} \mathbf{P}_1\left(\frac{K\sigma_K^2}{2} \left(N^K - \frac{\varepsilon}{2\log(3)} \log(K) \right) < T \right) = 0.$$

As L_0^K depends on μ_0^K note that $L_0^K = Z_0^K$, is function of the initial condition μ_0^K. However, by Assumption (4) and Markov's inequality, we have that

$$\mathbb{P}\left(L_0^K > 1 \right) \leqslant \mathbb{P}\left(M_2\left(\mu_0^K\right) \geqslant K^{\frac{\varepsilon}{2}} \right) \leqslant \frac{C_2^\star}{K^\varepsilon}$$

which tends to 0 when $K \to +\infty$ and the announced result follows. □

4.4.4 Conclusion

We denote respectively by N_+^K and N_-^K the number of upward transitions and downward transitions for $\left(Z_k^K\right)_{k\in\mathbb{N}}$ before reaching $\left\lfloor \frac{\varepsilon}{2\log(3)} \log(K) \right\rfloor$. Let $T \geqslant 0$ be fixed. Note that

$$\begin{aligned}
\mathbb{P}\left(\widehat{\tau}^K \leqslant \check{\tau}^K \wedge T \right) \leqslant\ & \mathbb{P}\left(\left\{ \widehat{\tau}^K \leqslant \check{\tau}^K \wedge T \right\} \cap \left\{ \check{\tau}^K > T_{N^K}^K \right\} \right) \\
& + \mathbb{P}\left(\left\{ \widehat{\tau}^K \leqslant \check{\tau}^K \wedge T \right\} \cap \left\{ \check{\tau}^K \leqslant T_{N^K}^K \right\} \right).
\end{aligned} \tag{40}$$

Step 1. Control of the first right-hand term of (40). As $Z_0^K = 1$ with probability converging to 1 and $Z_{N^K}^K = \left\lfloor \frac{\varepsilon}{2\log(3)} \log(K) \right\rfloor$, we have $N_+^K - N_-^K \leqslant \left\lfloor \frac{\varepsilon}{2\log(3)} \log(K) \right\rfloor$ and $N_+^K + N_-^K = N^K$. Thus,

$$N_-^K \geqslant \frac{N^K - \left\lfloor \frac{\varepsilon}{2\log(3)} \log(K) \right\rfloor}{2}.$$

Hence, if $\check{\tau}^K > T_{N^K}^K$, then by Lemma 4.13 (2),

$$\widehat{\tau}^K \geqslant T_{N^K}^K \geqslant \sum_{i=1}^{N^K} \left(T_i^K - T_{i-1}^K \right) \geqslant N_-^K K \sigma_K^2 \geqslant \frac{K\sigma_K^2}{2} \left(N^K - \frac{\varepsilon}{2\log(3)} \log(K) \right).$$

Therefore,

$$\mathbb{P}\left(\left\{ \widehat{\tau}^K \leqslant \check{\tau}^K \wedge T \right\} \cap \left\{ \check{\tau}^K > T_{N^K}^K \right\} \right) \leqslant \mathbb{P}\left(\frac{K\sigma_K^2}{2} \left(N^K - \frac{\varepsilon}{2\log(3)} \log(K) \right) \leqslant T \right)$$

which tends to 0 when $K \to +\infty$ according to Corollary 4.16.

Step 2. Control of the second right-hand term of (40). If $\check{\tau}^K \leqslant T_{N^K}^K$, let us consider $\widetilde{N}^K$ the last index k such that $T_k^K < \check{\tau}^K$. We have $\widetilde{N}^K < N^K$ and so $\widetilde{N}^K + 1 \leqslant \widetilde{N}^K$ which implies, thanks to Lemma 4.10, that for all $t \leqslant T_{\widetilde{N}^K+1}$ and K large enough,

$$M_2\left(\mu_t^K\right) \leqslant 2u_{\lfloor \frac{\varepsilon}{2\log(3)} \log(K) \rfloor} + 1 \leqslant 2K^{\frac{\varepsilon}{2}} + 1 < K^{\varepsilon}.$$

As the jump times of ν^K are isolated, we deduce that $\widehat{\tau}^K > T_{\widetilde{N}^K+1}^K$. Now, by definition of $\widetilde{N}^K$, we have $T_{\widetilde{N}^K+1}^K \geqslant \check{\tau}^K$ and thus $\widehat{\tau}^K > \check{\tau}^K$. Hence, we obtain that

$$\mathbb{P}\left(\left\{ \widehat{\tau}^K \leqslant \check{\tau}^K \wedge T \right\} \cap \left\{ \check{\tau}^K \leqslant T_{N^K}^K \right\} \right) = \mathbb{P}(\varnothing) = 0.$$

Therefore, $\lim_{K\to+\infty} \mathbb{P}\left(\widehat{\tau}^K \leqslant \check{\tau}^K \wedge T\right) = 0$ which concludes the proof of Lemma 4.9.

5 Tightness on the torus

The main result of this section is given by Theorem 5.1. This is a stochastic averaging result inspired by Kurtz [57, Theorem 2.1] establishing a tightness result, in the torus case, of the sequence of laws of $\left(\left(z^K, \Gamma^K\right)\right)_{K\in\mathbb{N}^\star}$ where Γ^K is the occupation measure of the fast component μ^K. We will use criteria proposed by Ethier-Kurtz in [40, Theorems 3.9.1 and 3.9.4]. In Sections 6 and 7 we identify the limit and we prove its uniqueness.

Let introduce the torus $\mathbb{T}_R := [x_0 - 2R, x_0 + 2R]$ of length $4R$ with $R > 0$ fixed and x_0 is the value of the mean trait of ν_0^K as in Theorem 2.1. We define $\mathcal{M}_1(\mathbb{T}_R)$ the set of probability measure on $\mathbb{T}_R$. We denote by $\mathcal{M}_m(\mathcal{M}_1(\mathbb{R}))$ the set of measures Γ on $\mathbb{R}_+ \times \mathcal{M}_1(\mathbb{R})$ such that for all $t \geqslant 0$, $\Gamma([0,t] \times \mathcal{M}_1(\mathbb{R})) = t$. For any $t \geqslant 0$, we denote by $\mathcal{M}_m^t(\mathcal{M}_1(\mathbb{R}))$ the set of measures $\Gamma \in \mathcal{M}_m(\mathcal{M}_1(\mathbb{R}))$ restricted to $[0,t] \times \mathcal{M}_1(\mathbb{R})$. For all $T > 0$, we denote by $\mathbb{D}([0,T], \mathbb{T}_R)$ the space of càd-làg functions on $[0,T]$ with values in $\mathbb{T}_R$.

Let $b_R \in \mathscr{C}_b^2\left(\mathbb{T}_R^2, \mathbb{R}\right)$ and $\theta_R \in \mathscr{C}_b^2\left(\mathbb{T}_R, \mathbb{R}\right)$ be two functions satisfying $b_R = b$ on $[x_0 - R, x_0 + R]^2$ and $\theta_R = \theta$ on $[x_0 - R, x_0 + R]$. Let $T > 0$ be fixed. We define on $\mathbb{T}_R$, in similar way as in Section 2, the processes $\nu^{K,R} := \left(\nu_t^{K,R}\right)_{t\in[0,T]}$, $z^{K,R} := \left(z_t^{K,R}\right)_{t\in[0,T]}$ and $\mu^{K,R} = \left(\mu_t^{K,R}\right)_{t\in[0,T]}$ as follows

$$\nu_t^{K,R} := \frac{1}{K}\sum_{i=1}^{K} \delta_{x_i^R(t)}, \quad z_t^{K,R} := \left\langle \mathrm{id}, \nu_{t/K\sigma_K^2}^{K,R} \right\rangle \quad \text{and} \quad \mu_t^{K,R} := \left(h_{\frac{1}{\sigma_K\sqrt{K}}} \circ \tau_{-z_t^K}\right) \sharp\, \nu_{t/K\sigma_K^2}^{K}$$

where for all $i \in \{1, \cdots, K\}$, $\nu_t^{K,R}$ is defined on $\mathcal{M}_1\left(\mathbb{T}_R\right)$ from b_R and θ_R as ν_t^K was defined from b and θ and $z_t^{K,R} \in \mathbb{T}_R$. Note that $\mu_t^{K,R}$ takes values in $\mathcal{M}_1\left(\mathbb{T}_R^K\right)$ where $\mathbb{T}_R^K$ is the torus corresponding to the interval $\left[\frac{1}{\sigma_K\sqrt{K}}\left(x_0 - 2R - z_t^K\right), \frac{1}{\sigma_K\sqrt{K}}\left(x_0 + 2R - z_t^K\right)\right]$. However, we will identify in the sequel $\mathcal{M}_1\left(\mathbb{T}_R^K\right)$ as a subset of $\mathcal{M}_1(\mathbb{R})$ using the natural embedding of $\mathbb{T}_R^K$ in $\mathbb{R}$. So, $\mu_t^{K,R} \in \mathcal{M}_1(\mathbb{R})$. We define also the stopping times

$$\widehat{\tau}^{K,R} := \inf\left\{ t \geqslant 0 \,\middle|\, M_2\left(\mu_t^{K,R}\right) \geqslant K^{\varepsilon} \right\},$$
$$\check{\tau}^{K,R} := \inf\left\{ t \geqslant 0 \,\middle|\, \mathrm{Diam}\left(\mathrm{Supp}\, \nu_t^{K,R}\right) > \frac{1}{\sigma_K K^{\frac{3+\varepsilon}{2}}} \right\},$$

and $\tau^{K,R} := \check{\tau}^{K,R} \wedge \widehat{\tau}^{K,R}$. We define on $\mathcal{M}_m^T\left(\mathcal{M}_1\left(\mathbb{R}\right)\right)$ the sequence of random measures $\left(\Gamma^{K,R}\right)_{K\in\mathbb{N}^\star}$ as follows

$$\Gamma^{K,R}\left(\mathrm{d}t, \mathrm{d}\mu\right) = \delta_{\mu_t^{K,R}}\left(\mathrm{d}\mu\right)\mathrm{d}t.$$

In the sequel, we study the limit as K tends to $+\infty$ of the sequence of laws of

$$\left(\left(z^{K,R}, \Gamma^{K,R}\right)\right)_{K\in\mathbb{N}^\star}$$

in $\mathcal{M}_1\left(\mathbb{D}\left([0,T], \mathbb{T}_R\right) \times \mathcal{M}_m^T\left(\mathcal{M}_1\left(\mathbb{R}\right)\right)\right)$.

Theorem 5.1. *Let $T, R > 0$. The sequence of laws of $\left(\left(z^{K,R}, \Gamma^{K,R}\right)\right)_{K\in\mathbb{N}^\star}$ is tight in the set of probability measures on $\mathbb{D}\left([0,T], \mathbb{T}_R\right) \times \mathcal{M}_m^T\left(\mathcal{M}_1\left(\mathbb{R}\right)\right)$ and for any limiting value $\mathbb{Q}$ of this sequence, the canonical process $\left(\zeta^R, \Gamma^R\right)$ on $\mathbb{D}\left([0,T], \mathbb{T}_R\right) \times \mathcal{M}_m^T\left(\mathcal{M}_1\left(\mathbb{R}\right)\right)$ satisfies that for all $f \in \mathscr{C}_b^2\left(\mathbb{T}_R, \mathbb{T}_R\right)$,*

$$N_t^{f,R} := f\left(\zeta_t^R\right) - f\left(\zeta_0^R\right) - \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\mathrm{SLOW}} f\left(\zeta_s^R, \mu\right) \Gamma^R\left(\mathrm{d}s, \mathrm{d}\mu\right) = 0 \tag{41}$$

$\mathbb{Q}-$*a.s. Moreover, for all $t \in [0,T]$,*

$$\mathbb{E}_{\mathbb{Q}}\left(\int_0^t \int_{\mathcal{M}_1(\mathbb{R})} M_5\left(\mu\right) \Gamma^R\left(\mathrm{d}s, \mathrm{d}\mu\right)\right) < \infty, \tag{42}$$

so that the definition (41) *makes sense.*

This result is based on Propositions 3.1 and 4.8, Lemma 4.3. The proof is divided in eight steps. In Step 1, we establish the tightness of the family of laws of the stopped slow component $\left(z_{\bullet\wedge\tau^{K,R}}^{K,R}\right)_{K\in\mathbb{N}^\star}$. In Step 2, we establish a compact containment condition for the stopped

fast component $\left(\mu^{K,R}_{\bullet\wedge\tau^{K,R}}\right)_{K\in\mathbb{N}^\star}$. In Step 3, we prove the tightness of the family of laws of the occupation measure. In Step 4, we deduce the tightness of the family of laws of the couple (slow, occupation measure fast). In Step 5, we prove uniform integrability results for a family $\left(N^{f,K,R}_{t\wedge\tau^{K,R}}\right)_{t\in[0,T],K\in\mathbb{N}^\star}$ constructed from $z^{K,R}$ and $\Gamma^{K,R}$ similarly as $N^{f,R}_t$ in (41). Step 6 is devoted to establish the convergence in distribution of $\left(N^{f,K,R}_{\bullet\wedge\tau^{K,R}}\right)_{K\in\mathbb{N}^\star}$ to $N^{f,R}$. In Steps 7 and 8, we prove that the limit $N^{f,R}$ is null $\mathbb{Q}-$a.s.

The main modification of Kurtz's setting of [57, Theorem 2.1] is that we have to work with stopped times and need to be careful with moment estimates and uniform integrability properties. This led us to rewrite the proof.

Proof of Theorem 5.1. Step 1. Tightness of the family of laws of $\left(z^{K,R}_{\bullet\wedge\tau^{K,R}}\right)_{K\in\mathbb{N}^\star}$ *on* $\mathbb{D}\left([0,T],\mathbb{T}_R\right)$. Let $f\in\mathscr{C}^2_b(\mathbb{T}_R,\mathbb{T}_R)$. For all $K\in\mathbb{N}^\star$ and for all $t\in[0,T]$, let us consider the two processes $Y^{K,R}_t:=f\left(z^{K,R}_{t\wedge\tau^{K,R}}\right)$ and $Z^{K,R}_t$ defined by the relation

$$Y^{K,R}_t=Y^{K,R}_0+\int_0^t Z^{K,R}_s\mathrm{d}s+M^{f,K,R}_{t\wedge\tau^{K,R}}$$

where $M^{f,K,R}_{t\wedge\tau^{K,R}}$ is the martingale given by (19) in the torus case. Note that, for all $s>\tau^{K,R}$, $Z^{K,R}_s=0$. From Proposition 3.2, (34) and Lemma 4.3, note that there exists a constant $C>0$ such that

$$\begin{aligned}
&\sup_{K\in\mathbb{N}^\star}\mathbb{E}\left[\left(\int_0^T\left|Z^{K,R}_t\right|^2\mathrm{d}t\right)^{\frac{1}{2}}\right]\\
&\qquad\leqslant\sup_{K\in\mathbb{N}^\star}\mathbb{E}\left[1+\int_0^T\left|Z^{K,R}_t\right|^2\mathrm{d}t\right]\\
&\qquad\leqslant C\left\{1+\sup_{K\in\mathbb{N}^\star}\mathbb{E}\left[\int_0^{T\wedge\tau^{K,R}}\left(M_2^2\left(\mu^{K,R}_t\right)+\frac{1}{K^2}\left[1+M_2^2\left(\mu^{K,R}_t\right)\right]\right)\mathrm{d}t\right]\right\}\\
&\qquad<\infty.
\end{aligned}$$

Hence, from [40, Theorem 3.9.4], the family of laws of $\left(Y^{K,R}\right)_{K\in\mathbb{N}^\star}$, on $\mathbb{D}\left([0,T],\mathbb{T}_R\right)$, is tight. Let us observe that the compact containment condition is satisfied by the stopped *slow* component $\left(z^{K,R}_{\bullet\wedge\tau^{K,R}}\right)_{K\in\mathbb{N}^\star}$ since $\mathbb{T}_R$ is compact and

$$\forall T>0,\quad\inf_{K\in\mathbb{N}^\star}\mathbb{P}\left(\sup_{0\leqslant t\leqslant T}z^{K,R}_{t\wedge\tau^{K,R}}\in\mathbb{T}_R\right)=1. \tag{43}$$

As $\mathscr{C}^2_b\left(\mathbb{T}_R,\mathbb{T}_R\right)$ is a dense subset of $\mathscr{C}^0_b\left(\mathbb{T}_R,\mathbb{T}_R\right)$ in the topology of the uniform norm, we deduce from (43) and [40, Theorem 3.9.1] that the family of laws of $\left(z^{K,R}_{\bullet\wedge\tau^{K,R}}\right)_{K\in\mathbb{N}^\star}$ is tight on $\mathbb{D}\left([0,T],\mathbb{T}_R\right)$.

Step 2. Compact containment condition.

Lemma 5.2. *Let $R>0$. For all $T\geqslant 0$, the family of laws of the marginal random variables of the stopped* fast *process $\left(\mu^{K,R}_{t\wedge\tau^{K,R}}\right)_{t\in[0,T],K\in\mathbb{N}^\star}$ is tight on $\mathcal{M}_1\left(\mathbb{R}\right)$ i.e.*

$$\forall\eta>0,\ \exists\, D^{R,T}_\eta\subset\mathcal{M}_1\left(\mathbb{R}\right)\ \text{compact},\ \forall t\in[0,T]\,,\ \forall K\in\mathbb{N}^\star,\quad\mathbb{P}\left(\mu^{K,R}_{t\wedge\tau^{K,R}}\in D^{R,T}_\eta\right)\geqslant 1-\eta.$$

Proof. Let $\eta > 0$ be fixed. From Proposition 4.8, there exists $K_0 \in \mathbb{N}^\star$ large enough such that for all $K \geqslant K_0$, for all $t \in [0,T]$, $\mathbb{P}\left(t > \tau^{K,R}\right) \leqslant \frac{\eta}{2}$. We consider the $\mathbb{R}-$valued sequence $(a_{q,\eta})_{q\in\mathbb{N}}$ satisfying for all $q \in \mathbb{N}, a_{q,\eta} > 1$, and $\sum_{q\in\mathbb{N}^\star} \frac{q}{a_{q,\eta}^2} < \frac{\eta}{2\overline{M}_2}$ where $\overline{M}_2 := \max\left\{C_2^\star, \overline{\theta}\overline{m}_2/\underline{b}\right\}$ is a uniform upper bound of $\mathbb{E}\left(M_2\left(\mu^{K,R}_{t\wedge\tau^{K,R}}\right)\right)$ given by Corollary 4.6. Let $(\kappa_{q,\eta})_{q\in\mathbb{N}}$ be a sequence of compact intervals on $\mathbb{R}$, increasing for inclusion, of the form $[-a_{q,\eta}, a_{q,\eta}]$. Let $t \in [0,T]$. For all $q \in \mathbb{N}^\star$, and $K \geqslant K_0$, note that

$$M_2\left(\mu_t^{K,R}\right)\mathbb{1}_{t\leqslant\tau^{K,R}} \geqslant \mathbb{1}_{t\leqslant\tau^{K,R}}\int_{\kappa_{q,\eta}^c} x^2\mu_t^{K,R}(\mathrm{d}x) \geqslant a_{q,\eta}^2\mu^{K,R}_{t\wedge\tau^{K,R}}\left(\kappa_{q,\eta}^c\right)\mathbb{1}_{t\leqslant\tau^{K,R}},$$

and we deduce that $\left\{\mu^{K,R}_{t\wedge\tau^{K,R}}\left(\kappa_{q,\eta}^c\right) > \frac{1}{q}\right\} \cap \left\{t \leqslant \tau^{K,R}\right\} \subset \left\{M_2\left(\mu^{K,R}_{t\wedge\tau^{K,R}}\right) > \frac{a_{q,\eta}^2}{q}\right\} \cap \left\{t \leqslant \tau^{K,R}\right\}$. Hence,

$$\begin{aligned}
&\mathbb{P}\left(\left\{\exists q \in \mathbb{N}^\star,\ \mu^{K,R}_{t\wedge\tau^{K,R}}\left(\kappa_{q,\eta}^c\right) > \frac{1}{q}\right\}\right)\\
&\qquad\leqslant \mathbb{P}\left(\tau^{K,R} < t\right) + \sum_{q\in\mathbb{N}^\star}\mathbb{P}\left(\left\{\mu^{K,R}_{t\wedge\tau^{K,R}}\left(\kappa_{q,\eta}^c\right) > \frac{1}{q}\right\} \cap \left\{t \leqslant \tau^{K,R}\right\}\right)\\
&\qquad\leqslant \mathbb{P}\left(\tau^{K,R} < t\right) + \sum_{q\in\mathbb{N}^\star}\mathbb{P}\left(M_2\left(\mu_t^{K,R}\right)\mathbb{1}_{t\leqslant\tau^{K,R}} > \frac{a_{q,\eta}^2}{q}\right)\\
&\qquad\leqslant \mathbb{P}\left(\tau^{K,R} < t\right) + \sum_{q\in\mathbb{N}^\star}\frac{\mathbb{E}\left(M_2\left(\mu_t^{K,R}\right)\mathbb{1}_{t\leqslant\tau^{K,R}}\right)}{\frac{a_{q,\eta}^2}{q}}\\
&\qquad\leqslant \eta.
\end{aligned}$$

Therefore, we have proved that

$$\forall\eta > 0,\ \forall t \in [0,T],\ \forall K \geqslant K_0, \quad \mathbb{P}\left(\mu^{K,R}_{t\wedge\tau^{K,R}} \in \mathcal{K}_\eta\right) \geqslant 1-\eta,$$

where $\mathcal{K}_\eta := \left\{\mu \in \mathcal{M}_1(\mathbb{R}) \,\middle|\, \forall q \in \mathbb{N}^\star, \mu\left(\kappa_{q,\eta}^c\right) \leqslant \frac{1}{q}\right\}$ which is compact by Prohorov's theorem. □

Step 3. Tightness of the family of laws of $\left(\Gamma^{K,R}\right)_{K\in\mathbb{N}^\star}$ on $\mathcal{M}_m^T\left(\mathcal{M}_1\left(\mathbb{R}\right)\right)$. Let $\eta > 0$ be fixed. From Proposition 4.8, there exists $K_0 \in \mathbb{N}^\star$ such that for all $K \geqslant K_0$, for all $t \in [0,T]$, $\mathbb{P}\left(t \geqslant \tau^{K,R}\right) \leqslant \frac{\eta}{2}$. Consider $D_{\frac{\eta}{2}}^{R,T}$ the compact set in Lemma 5.2. It follows that for all $t \in [0,T]$, $K \geqslant K_0$,

$$\begin{aligned}
\mathbb{E}\left(\Gamma^{K,R}\left([0,t]\times D_{\frac{\eta}{2}}^{R,T}\right)\right) &\geqslant \mathbb{E}\left(\Gamma^{K,R}\left([0,t]\times D_{\frac{\eta}{2}}^{R,T}\right)\mathbb{1}_{t<\tau^{K,R}}\right)\\
&= \int_0^t \mathbb{P}\left(\left\{\mu^{K,R}_{s\wedge\tau^{K,R}} \in D_{\frac{\eta}{2}}^{R,T}\right\} \cap \left\{t < \tau^{K,R}\right\}\right)\mathrm{d}s\\
&\geqslant \int_0^t \mathbb{P}\left(\mu^{K,R}_{s\wedge\tau^{K,R}} \in D_{\frac{\eta}{2}}^{R,T}\right)\mathrm{d}s - \int_0^t \mathbb{P}\left(t \geqslant \tau^{K,R}\right)\mathrm{d}s\\
&\geqslant t\left(1-\eta\right).
\end{aligned}$$

Therefore, the tightness of the family of laws of $\left(\Gamma^{K,R}\right)_{K\in\mathbb{N}^\star}$ follows from [57, Lemma 1.3].

Step 4. Tightness of the family of laws of $\left(\left(z^{K,R},\Gamma^{K,R}\right)\right)_{K\in\mathbb{N}^\star}$ *on* $\mathbb{D}\left([0,T],\mathbb{T}_R\right)\times\mathcal{M}_m^T\left(\mathcal{M}_1\left(\mathbb{R}\right)\right)$. From Steps 1 and 3 and Prohorov's theorem, we deduce that the family of laws of the couple $\left(\left(z^{K,R}_{\bullet\wedge\tau^{K,R}},\Gamma^{K,R}\right)\right)_{K\in\mathbb{N}^\star}$ is relatively compact in $\mathcal{M}_1\left(\mathbb{D}\left([0,T],\mathbb{T}_R\right)\times\mathcal{M}_m^T\left(\mathcal{M}_1\left(\mathbb{R}\right)\right)\right)$. Hence, there exists a probability measure $\mathbb{Q}$ on the canonical space $\mathbb{D}\left([0,T],\mathbb{T}_R\right)\times\mathcal{M}_m^T\left(\mathcal{M}_1\left(\mathbb{R}\right)\right)$ and an increasing function $n:\mathbb{N}^\star\to\mathbb{N}^\star$ such that the subsequence of laws of $\left(\left(z^{n(K),R}_{\bullet\wedge\tau^{n(K),R}},\Gamma^{n(K),R}\right)\right)_{K\in\mathbb{N}^\star}$ converges weakly to the limiting value $\mathbb{Q}$ when $K\to+\infty$. Thanks to Proposition 4.8, we deduce that the family of laws of $\left(\left(z^{n(K),R},\Gamma^{n(K),R}\right)\right)_{K\in\mathbb{N}^\star}$ converges weakly to $\mathbb{Q}$ when $K\to+\infty$ and therefore that the family of laws of $\left(\left(z^{K,R},\Gamma^{K,R}\right)\right)_{K\in\mathbb{N}^\star}$ is relatively compact, thus tight on $\mathbb{D}\left([0,T],\mathbb{T}_R\right)\times\mathcal{M}_m^T\left(\mathcal{M}_1\left(\mathbb{R}\right)\right)$ by Prohorov's theorem.

Step 5. Uniform integrability. For all $f\in\mathscr{C}_b^2\left(\mathbb{T}_R,\mathbb{T}_R\right)$, $K\in\mathbb{N}^\star$, let us consider $\left(N_t^{f,K,R}\right)_{t\in[0,T]}$ the stochastic process defined by

$$N_t^{f,K,R}:=f\left(z_t^{K,R}\right)-f\left(z_0^{K,R}\right)-\int_0^t\int_{\mathcal{M}_1(\mathbb{R})}\mathcal{L}_{\mathrm{SLOW}}f\left(z_s^{K,R},\mu\right)\Gamma^{K,R}\left(\mathrm{d}s,\mathrm{d}\mu\right).$$

From (9), we have for all $f\in\mathscr{C}_b^2\left(\mathbb{T}_R,\mathbb{T}_R\right)$, there exists a constant $C>0$ such that

$$\forall t\in[0,T],\ \forall K\in\mathbb{N}^\star,\quad\left|N^{f,K,R}_{t\wedge\tau^{K,R}}\right|\leqslant C\left(1+\int_0^t M_2\left(\mu_s^{K,R}\right)\mathbb{1}_{s\leqslant\tau^{K,R}}\mathrm{d}s\right),$$

Hence, the uniform integrability of $\left(N^{f,K,R}_{\bullet\wedge\tau^{K,R}}\right)_{K\in\mathbb{N}^\star}$ follows from Corollary 4.4.

Step 6. Proof of (42) and almost sure convergence of $\left(\widetilde{N}^{f,K,R}_{\bullet\wedge\tau^{K,R}}\right)_{K\in\mathbb{N}^\star}$ *to* $\widetilde{N}^{f,R}$. From Step 4, Proposition 4.8 and Skorohod's representation theorem, there exists an increasing function $\bar{n}:\mathbb{N}^\star\to\mathbb{N}^\star$ and a probability space on which we define, the random variable $\widetilde{\tau}^{\bar{n}(K),R}$, the families $\left(\widetilde{z}^{\bar{n}(K),R}\right)_{K\in\mathbb{N}^\star}$ and $\left(\widetilde{\Gamma}^{\bar{n}(K),R}\right)_{K\in\mathbb{N}^\star}$ and $\widetilde{\zeta}^R,\widetilde{\Gamma}^R$ copies of $\left(z^{\bar{n}(K),R}\right)_{K\in\mathbb{N}^\star}$, $\left(\Gamma^{\bar{n}(K),R}\right)_{K\in\mathbb{N}^\star}$, ζ^R, Γ^R under $\mathbb{Q}$ such that the sequence $\left(\left(\widetilde{z}^{\bar{n}(K),R},\widetilde{\Gamma}^{\bar{n}(K),R},\widetilde{\tau}^{\bar{n}(K),R}\right)\right)_{K\in\mathbb{N}^\star}$ converges a.s. to $\left(\widetilde{\zeta}^R,\widetilde{\Gamma}^R,+\infty\right)$ when $K\to+\infty$. Note that for all $t\in[0,T]$,

$$\widetilde{\Gamma}^{\bar{n}(K),R}\left(\left[0,t\wedge\widetilde{\tau}^{\bar{n}(K),R}\right]\times\mathcal{M}_1\left(\mathbb{R}\right)\right)=t\wedge\widetilde{\tau}^{\bar{n}(K),R}\xrightarrow[K\longrightarrow+\infty]{\text{a.s.}}t=\widetilde{\Gamma}^R\left([0,t]\times\mathcal{M}_1\left(\mathbb{R}\right)\right).$$

From [57, Lemma 1.5 (b)(c)(d)] and Corollary 4.4, we have

$$\int_0^{t\wedge\widetilde{\tau}^{K,R}}\int_{\mathcal{M}_1(\mathbb{R})}M_5\left(\mu\right)\widetilde{\Gamma}^{K,R}\left(\mathrm{d}s,\mathrm{d}\mu\right)\xrightarrow[K\longrightarrow+\infty]{\text{a.s.}}\int_0^t\int_{\mathcal{M}_1(\mathbb{R})}M_5\left(\mu\right)\widetilde{\Gamma}^R\left(\mathrm{d}s,\mathrm{d}\mu\right)$$

and (42) follows. From [57, Lemma 1.5 (b)(c)(d)], we also deduce that for all $f\in\mathscr{C}_b^2\left(\mathbb{T}_R,\mathbb{T}_R\right)$, for all $t\in[0,T]$, a.s.

$$\begin{aligned}\lim_{K\to+\infty}\int_0^{t\wedge\widetilde{\tau}^{\bar{n}(K),R}}&\int_{\mathcal{M}_1(\mathbb{R})}\mathcal{L}_{\mathrm{SLOW}}f\left(\widetilde{z}_s^{\bar{n}(K),R},\mu\right)\widetilde{\Gamma}^{\bar{n}(K),R}\left(\mathrm{d}s,\mathrm{d}\mu\right)\\&=\int_0^t\int_{\mathcal{M}_1(\mathbb{R})}\mathcal{L}_{\mathrm{SLOW}}f\left(\widetilde{\zeta}^R,\mu\right)\widetilde{\Gamma}^R\left(\mathrm{d}s,\mathrm{d}\mu\right)\end{aligned}$$

and thus the sequence $\left(\widetilde{N}^{f,\bar{n}(K),R}_{\bullet\wedge\widetilde{\tau}^{\bar{n}(K),R}}\right)_{K\in\mathbb{N}^\star}$ converges a.s. to the process $\left(\widetilde{N}^{f,R}_t\right)_{t\in[0,T]}$ where $\widetilde{N}^{f,K,R}_t$ and $\widetilde{N}^{f,R}_t$ are respectively defined for all $K\in\mathbb{N}^\star$ by

$$\widetilde{N}^{f,K,R}_t := f\left(\widetilde{z}^{K,R}_t\right) - f\left(\widetilde{z}^{K,R}_0\right) - \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\mathrm{SLOW}} f\left(\widetilde{z}^{K,R}_t,\mu\right)\widetilde{\Gamma}^{K,R}\left(\mathrm{d}s,\mathrm{d}\mu\right), \qquad t\leqslant T\wedge\widetilde{\tau}^{K,R}$$

$$\widetilde{N}^{f,R}_t := f\left(\widetilde{\zeta}^{R}_t\right) - f\left(\widetilde{\zeta}^{R}_0\right) - \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\mathrm{SLOW}} f\left(\widetilde{\zeta}^{R}_t,\mu\right)\widetilde{\Gamma}^{R}\left(\mathrm{d}s,\mathrm{d}\mu\right), \quad t\leqslant T.$$

Step 7. $N^{f,R}$ is a martingale. Let us consider the filtration $\left(\widetilde{\mathcal{F}}^{K,R}_t\right)_{t\in[0,T]}$ defined by $\widetilde{\mathcal{F}}^{K,R}_t := \sigma\left(\widetilde{z}^{K,R}_s, \widetilde{\Gamma}^{K,R}\left([0,s]\times H\right)\middle| s\leqslant t, H\in\mathcal{B}\left(\mathcal{M}_1\left(\mathbb{R}\right)\right)\right)$, $(f_i)_{1\leqslant i\leqslant q}$, $q\in\mathbb{N}^\star$ bounded Lipschitz functions from $\mathbb{T}_R$ to $\mathbb{T}_R$ and $0<t_1\leqslant\cdots\leqslant t_q\leqslant s<t$. Let us denote for all $K\in\mathbb{N}^\star$, for all $t\leqslant\widetilde{\tau}^{K,R}\wedge T$, $\widetilde{M}^{f,K,R}_t := \widetilde{N}^{f,K,R}_t + \widetilde{\mathcal{E}}^{f,K,R}_t$ where $\widetilde{M}^{f,K,R}$ is constructed from $\widetilde{z}^{K,R}$ and $\widetilde{\mu}^{K,R}$ as in (19) and $\widetilde{\mathcal{E}}^{f,K,R}_t$ is an error term. Note that $\widetilde{M}^{f,K,R}_t$ is a $\left(\widetilde{\mathcal{F}}^{K,R}_t\right)_{t\in[0,T]}-$martingale as in Proposition 3.1. Hence,

$$\mathbb{E}\left(\widetilde{M}^{f,\bar{n}(K),R}_{t\wedge\widetilde{\tau}^{\bar{n}(K),R}}\prod_{i=1}^q f_i\left(\widetilde{M}^{f,\bar{n}(K),R}_{t_i\wedge\widetilde{\tau}^{\bar{n}(K),R}}\right)\right) = \mathbb{E}\left(\widetilde{M}^{f,\bar{n}(K),R}_{s\wedge\widetilde{\tau}^{\bar{n}(K),R}}\prod_{i=1}^q f_i\left(\widetilde{M}^{f,\bar{n}(K),R}_{t_i\wedge\widetilde{\tau}^{\bar{n}(K),R}}\right)\right).$$

From Proposition 3.2 and (34), $\widetilde{\mathcal{E}}^{f,K,R}_{t\wedge\widetilde{\tau}^{K,R}} = \frac{1}{K}O\left(\int_0^t\left[1+M_2\left(\widetilde{\mu}^{K,R}_s\right)\right]\mathbb{1}_{s\leqslant\widetilde{\tau}^{K,R}}\mathrm{d}s\right)$ and then satisfy the condition $\lim_{K\to+\infty}\mathbb{E}\left(\sup_{0\leqslant t\leqslant T}\left|\mathcal{E}^{f,K,R}_{t\wedge\widetilde{\tau}^{K,R}}\right|\right)=0$. As for all $i\in\{1,\cdots,q\}$, f_i is Lipschitz, there exists a constant $C>0$ such that

$$\begin{aligned}\mathbb{E}\left(\left|f_i\left(\widetilde{M}^{f,\bar{n}(K),R}_{t_i\wedge\widetilde{\tau}^{\bar{n}(K),R}} - \mathcal{E}^{f,\bar{n}(K),R}_{t_i\wedge\tau^{\bar{n}(K),R}}\right) - f_i\left(\widetilde{M}^{f,\bar{n}(K),R}_{t_i\wedge\widetilde{\tau}^{\bar{n}(K),R}}\right)\right|\right)&\\ \leqslant \frac{C}{\bar{n}\left(K\right)}\mathbb{E}\left(\int_0^{t_i}\left[1+M_2\left(\mu^{\bar{n}(K),R}_r\right)\right]\mathbb{1}_{r\leqslant\widetilde{\tau}^{\bar{n}(K),R}}\mathrm{d}r\right)&\end{aligned}$$

where the term of the right hand side of the previous inequality goes to 0 when $K\to+\infty$ according to Lemma 4.3. We deduce that for all $u\in\{s,t\}$,

$$\lim_{K\to+\infty}\mathbb{E}\left(\widetilde{M}^{f,\bar{n}(K),R}_{u\wedge\widetilde{\tau}^{\bar{n}(K),R}}\prod_{i=1}^q f_i\left(\widetilde{M}^{f,\bar{n}(K),R}_{t_i\wedge\widetilde{\tau}^{\bar{n}(K),R}}\right)\right) = \lim_{K\to+\infty}\mathbb{E}\left(\widetilde{N}^{f,\bar{n}(K),R}_{u\wedge\widetilde{\tau}^{\bar{n}(K),R}}\prod_{i=1}^q f_i\left(\widetilde{N}^{f,\bar{n}(K),R}_{t_i\wedge\widetilde{\tau}^{\bar{n}(K),R}}\right)\right).$$

From Steps 5 and 6, we deduce that

$$\mathbb{E}\left(\widetilde{N}^{f,R}_t\prod_{i=1}^q f_i\left(\widetilde{N}^{f,R}_{t_i}\right)\right) = \mathbb{E}\left(\widetilde{N}^{f,R}_s\prod_{i=1}^q f_i\left(\widetilde{N}^{f,R}_{t_i}\right)\right).$$

Hence,

$$\mathbb{E}\left(N^{f,R}_t\prod_{i=1}^q f_i\left(N^{f,R}_{t_i}\right)\right) = \mathbb{E}\left(N^{f,R}_s\prod_{i=1}^q f_i\left(N^{f,R}_{t_i}\right)\right).$$

Since the last property is true for all $q \in \mathbb{N}^\star$, $t_1 \leqslant \cdots \leqslant t_q \leqslant s < t$ and for all bounded LIPSCHITZ functions $(f_i)_{1\leqslant i\leqslant q}$, the monotone class theorem ensures us that

$$\mathbb{E}\left(N_t^{f,R}\,\middle|\,\sigma\left(N_u^{f,R}\,\middle|\,u\leqslant s\right)\right) = N_s^{f,R}.$$

Hence the announced result.

Step 8. Nullity of $N^{f,R}$. On the one hand, from ITÃ´'s formula [80, Theorem 32 of Chapter II], for all $f \in \mathscr{C}_b^2(\mathbb{T}_R, \mathbb{T}_R)$ and $t \in [0,T]$

$$\begin{aligned} f^2\left(\zeta_t^R\right) = f^2\left(\zeta_0^R\right) &+ 2\int_0^t \int_{\mathcal{M}_1(\mathbb{R})} f\left(\zeta_s^R\right) \mathcal{L}_{\mathrm{SLOW}} f\left(\zeta_s^R, \mu\right) \Gamma^R(\mathrm{d}s, \mathrm{d}\mu) \\ &+ 2\int_0^t f\left(\zeta_s^R\right) \mathrm{d}N_s^{f,R} + \left\langle N^{f,R}\right\rangle_t + \sum_{0<s\leqslant t}\left(f\left(\zeta_s^R\right) - f\left(\zeta_{s^-}^R\right)\right)^2. \end{aligned} \tag{44}$$

On the other hand, applying (41) with $f^2 \in \mathscr{C}_b^2(\mathbb{T}_R, \mathbb{T}_R)$, we obtain that for all $t \in [0,T]$

$$f^2\left(\zeta_t^R\right) = f^2\left(\zeta_0^R\right) + \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\mathrm{SLOW}} f^2\left(\zeta_s^R, \mu\right) \Gamma^R(\mathrm{d}s, \mathrm{d}\mu) + N_t^{f^2,R}. \tag{45}$$

Comparing (44) and (45) leads for all $t \in [0,T]$ to

$$N_t^{f^2,R} - 2\int_0^t f\left(\zeta_s^R\right) \mathrm{d}N_s^{f,R} = \left\langle N^{f,R}\right\rangle_t + \sum_{0<s\leqslant t}\left(f\left(\zeta_s^R\right) - f\left(\zeta_{s^-}^R\right)\right)^2$$

and thus by [48, Theorem 4.1] that $\mathbb{Q}-$a.s. for all $t \in [0,T]$,

$$\left\langle N^{f,R}\right\rangle_t = -\sum_{0<s\leqslant t}\left(f\left(\zeta_s^R\right) - f\left(\zeta_{s^-}^R\right)\right)^2 \leqslant 0$$

so that $\left\langle N^{f,R}\right\rangle_t = 0$. Hence, $\mathbb{Q}-$a.s. $N^{f,R} = 0$ which completes the proof. □

6 Characterisation of the occupation measure limit on the torus

Consider a probability measure $\mathbb{Q}$ on $\mathbb{D}([0,T], \mathbb{T}_R) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$ and the canonical process $\left(\zeta^R, \Gamma^R\right)$ as in Theorem 5.1. The following lemma gives us a desintegration result of the occupation measure Γ^R that we characterise below.

Lemma 6.1. *Let $T \geqslant 0$ be fixed. With the notations of* Theorem 5.1*, there exists a random probability measure-valued process $\left(\gamma_t^R\right)_{t\in[0,T]}$ that is predictable in (ω, t) and such that for all bounded measurable function $\psi : [0,T] \times \mathcal{M}_1(\mathbb{R}) \to \mathbb{T}_R$,*

$$\int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \psi(s,\mu) \Gamma^R(\mathrm{d}s, \mathrm{d}\mu) = \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \psi(s,\mu) \gamma_s^R(\mathrm{d}\mu)\mathrm{d}s. \tag{46}$$

Proof. The desintegration result of Γ^R follows directly from [57, Lemma 1.4]. □

Corollary 6.2. *Let $T \geqslant 0$ be fixed. With the notations of* Theorem 5.1*, we have that $\zeta^R \in \mathscr{C}^0([0,T],\mathbb{T}_R)$ is differentiable of derivative in $L^1(\mathbb{R})$ $\mathbb{Q}-a.s.$*

Proof. Applying (41) with $f = \mathrm{id} \in \mathscr{C}_b^1(\mathbb{T}_R,\mathbb{T}_R)$, we deduce from Theorem 5.1, (9) and (46), $\mathbb{Q}-$a.s., for all $t \in [0,T]$

$$\zeta_t^R = \zeta_0^R + \int_0^t \left(\int_{\mathcal{M}_1(\mathbb{R})} M_2(\mu)\, \gamma_s^R(\mathrm{d}\mu) \right) \partial_1 \mathrm{Fit}\left(\zeta_s^R, \zeta_s^R\right) \mathrm{d}s.$$

From Assumptions **(A)** and (42), the integrand of the previous time integral is in $L^1(\mathbb{R})$ $\mathbb{Q}-$a.s. Hence, the announced result follows from the fundamental theorem of calculus. □

We now want to characterise the limiting value $\Gamma^R(\mathrm{d}t,\mathrm{d}\mu) = \gamma_t^R(\mathrm{d}\mu)\,\mathrm{d}t$ under $\mathbb{Q}$.

Proposition 6.3. *With the notations of* Theorem 5.1*, for a.e. $t \in [0,T]$, $\mathbb{Q}-a.s.$, $\gamma_t^R = \pi^{\lambda(\zeta^R)}$ where π^λ is the unique invariant probability measure of the centered* Fleming-Viot *process with resampling rate λ (see* [18, Section 4]*).*

It is here that we exploit ergodicity properties for the fast limit component. The proof of Proposition 6.3, given in Section 6.1 is based on the following technical lemma giving a characterisation of π^λ and proved in Section 6.2. To state this lemma, let us first recall from (22) the definition of polynomials in μ:

$$P_{f,n}(\mu) := \langle f, \mu^n \rangle := \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, \cdots, x_n)\, \mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_n)$$

with $n \in \mathbb{N}^\star$, $\mu \in \mathcal{M}_1^c(\mathbb{R})$, $f \in \mathscr{C}_b^3(\mathbb{R}^n,\mathbb{R})$. For all $n \in \mathbb{N}^\star$, for any function $f : \mathbb{R}^n \to \mathbb{R}$ whose second derivatives exist, we denote by $\mathrm{Hess}(f) := \left(\partial_{ij}^2 f\right)_{1 \leqslant i,j \leqslant n}$ the Hessian matrix of f. For all $n \in \mathbb{N}^\star$, let us denote by $\mathscr{C}_{\|\cdot\|}^2(\mathbb{R}^n,\mathbb{R})$ the set

$$\left\{ f \in \mathscr{C}^2(\mathbb{R}^n,\mathbb{R}) \,\middle|\, \exists\, C > 0,\ \forall x \in \mathbb{R}^n,\ |f(x)| + \|\nabla f(x)\|_\infty + \|\mathrm{Hess}(f)(x)\|_\infty \leqslant C\left(1 + \|x\|_\infty^2\right) \right\}.$$

From (26), we can see that if $f \in \mathscr{C}_b^4(\mathbb{R}^n,\mathbb{R})$, $\mathcal{L}_{\mathrm{FVc}}^\lambda P_{f,n}(\mu)$ is a polynomial in μ of the form $P_{J,n+1}(\mu)$ for some function $J \in \mathscr{C}_{\|\cdot\|}^2\left(\mathbb{R}^{n+1},\mathbb{R}\right)$. We recall from [18, Proposition 2.11] that if $\mu \in \mathcal{M}_1^{c,5}(\mathbb{R})$, then $\sup_{0 \leqslant t \leqslant T} \mathbb{E}_\mu\left(M_5(X_t)\right) < \infty$ where $(X_t)_{t \geqslant 0}$ denotes the centered Fleming-Viot process with resampling rate λ. Since, by (26), for all $\mu \in \mathcal{M}_1^{c,4}(\mathbb{R})$,

$$|\mathcal{L}_{\mathrm{FVc}} P_{J,n+1}(\mu)| \leqslant C\left(1 + M_4(\mu)\right),$$

for some constant $C > 0$, we can apply the martingale problem (27) to the function $P_{J,n+1}(\mu)$ using classical localisation techniques to obtain that the process $\left(\widehat{M}_t^{P_{J,n+1}}\right)_{t \geqslant 0}$ defined by

$$\widehat{M}_t^{P_{J,n+1}} := P_{J,n+1}(X_t) - P_{J,n+1}(X_0) - \int_0^t \mathcal{L}_{\mathrm{FVc}}^\lambda P_{J,n+1}(X_s)\,\mathrm{d}s \tag{47}$$

is a $\mathbb{P}_\mu-$martingale for all $\mu \in \mathcal{M}_1^{c,5}(\mathbb{R})$.

Lemma 6.4. *Let $\lambda \in \mathbb{R}$ be fixed. If $\gamma \in \mathcal{M}_1(\mathcal{M}_1(\mathbb{R}))$ satisfies $\int_{\mathcal{M}_1(\mathbb{R})} M_4(\mu)\,\gamma(\mathrm{d}\mu) < \infty$ and*

$$\forall n \in \mathbb{N}^\star,\ \forall f \in \mathscr{C}_{\|\cdot\|}^2(\mathbb{R}^n,\mathbb{R}),\quad \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\mathrm{FVc}}^\lambda P_{f,n}(\mu)\,\gamma(\mathrm{d}\mu) = 0,$$

then, $\gamma = \pi^\lambda$.

6.1 Proof of Proposition 6.3

Let us define for all $\ell \in \mathbb{N}$, $\mathscr{C}^{\ell}_K(\mathbb{R}^n,\mathbb{R})$ the space of real functions of class $\mathscr{C}^{\ell}(\mathbb{R}^n,\mathbb{R})$ with compact support. From Lemma 3.4 and (13), for all $t \in [0,T]$, for all $n \in \mathbb{N}^\star$ we have for all $f \in \mathscr{C}^3_K(\mathbb{R}^n,\mathbb{R})$ that

$$K^2\sigma_K^2 M^{K,P_{f,n}}_{t\wedge\tau^{K,R}} = \sum_{i=1}^{3} (\mathbf{A})_i^{K,R}\left(t\wedge\tau^{K,R}\right)$$

where for all $t \leqslant \tau^{K,R}\wedge T$

$$\begin{aligned}
(\mathbf{A})_1^{K,R}(t) &:= K^2\sigma_K^2\left(P_{f,n}\left(\mu_t^{K,R}\right) - P_{f,n}\left(\mu_0^{K,R}\right)\right),\\
(\mathbf{A})_2^{K,R}(t) &:= -\int_0^t\int_{\mathcal{M}_1(\mathbb{R})} \theta\left(z_s^{K,R}\right) m_2\left(z_s^{K,R}\right) \mathcal{L}_{\mathrm{FVc}}^{\lambda\left(z_s^{K,R}\right)} P_{f,n}(\mu)\,\Gamma^{K,R}(\mathrm{d}s,\mathrm{d}\mu),\\
(\mathbf{A})_3^{K,R}(t) &:= O\left(\frac{1}{\sqrt{K}} + \sigma_K K^{\frac{3}{2}+\varepsilon} + \int_0^t \frac{M_3\left(\mu_s^{K,R}\right)}{K}\,\mathrm{d}s\right),
\end{aligned}$$

is a martingale. Since $\sigma_K K \to 0$ by Assumption (5), $(\mathbf{A})_1^{K,R}\left(t\wedge\tau^{K,R}\right) \to 0$ when $K \to +\infty$. From Corollary 4.4, the sequence $\left((\mathbf{A})_2^{K,R}\left(t\wedge\tau^{K}\right)\right)_{t\in[0,T],K\in\mathbb{N}^\star}$ is uniformly integrable and converges in law, when $K\to+\infty$, to

$$M_t^{P_{f,n}} := -\int_0^t \theta\left(\zeta_s^R\right) m_2\left(\zeta_s^R\right)\int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\mathrm{FVc}}^{\lambda\left(\zeta_s^R\right)} P_{f,n}(\mu)\,\Gamma^R(\mathrm{d}s,\mathrm{d}\mu).$$

Note that, from Cauchy-Schwarz's inequality and Lemma 4.3, there exists a constant $C>0$ such that

$$\mathbb{E}\left(\sup_{0\leqslant t\leqslant T}\int_0^{t\wedge\tau^{K,R}} \frac{M_3\left(\mu_s^{K,R}\right)}{K}\mathrm{d}s\right)^2 \leqslant \frac{T}{K^2}\mathbb{E}\left(\int_0^{T\wedge\tau^{K,R}} M_6\left(\mu_s^{K,R}\right)\mathrm{d}s\right) \leqslant \frac{CT^2}{K^2}.$$

Hence, $\lim_{K\to+\infty}\sup_{0\leqslant t\leqslant T}(\mathbf{A})_3^{K,R}(t\wedge\tau^{K,R}) = 0$ in $L^2(\mathbb{R})$. In particular, we deduce that a subsequence of $\left(\sup_{0\leqslant t\leqslant T}(\mathbf{A})_3^{K,R}(t\wedge\tau^{K,R})\right)_{K\in\mathbb{N}^\star}$ converges almost surely to 0 and the family $\left(\sup_{0\leqslant t\leqslant T}(\mathbf{A})_3^{K,R}(t\wedge\tau^{K,R})\right)_{K\in\mathbb{N}^\star}$ is uniformly integrable along this subsequence [84, Theorem 13.7]. Using the same method based on Skorohod's representation theorem as in the proof of Theorem 5.1, we deduce that the process $\left(M_t^{P_{f,n}}\right)_{t\geqslant 0}$ is a $\mathbb{Q}-$martingale. As it is also a continuous and finite variation process by Lemma 6.1, it must hence be $\mathbb{Q}-$a.s. null [48, Theorem 4.1]. Hence, using Lemma 6.1 again, we have proved that

$$\forall f\in\mathscr{C}^3_K(\mathbb{R}^n,\mathbb{R}),\ \mathbb{Q}-\text{a.s.},\ \mathrm{d}t-\text{a.e.},\quad \int_{\mathcal{M}_1(\mathbb{R})}\mathcal{L}_{\mathrm{FVc}}^{\lambda\left(\zeta_t^R\right)} P_{f,n}(\mu)\,\gamma_t^R(\mathrm{d}\mu) = 0.$$

The space $\mathscr{C}^2_K(\mathbb{R}^n,\mathbb{R})$ equipped with the norm

$$\|f\|_{W_0^{2,\infty}} = \|f\|_\infty + \|\nabla f\|_\infty + \|\mathrm{Hess}(f)\|_\infty$$

is separable. So, we can choose a dense countable family $\mathcal{B}\subset\mathscr{C}^3_K(\mathbb{R}^n,\mathbb{R})$ such that

$$\forall f\in\mathscr{C}^2_K(\mathbb{R}^n,\mathbb{R}),\ \exists\,(f_q)_{q\in\mathbb{N}^\star}\in\mathcal{B}^{\mathbb{N}^\star},\quad f_q \xrightarrow[q\longrightarrow+\infty]{\|\cdot\|_{W_0^{2,\infty}}} f.$$

Then,

$$\mathbb{Q}-\text{a.s.},\ \mathrm{d}t-\text{a.e.},\ \forall q\in\mathbb{N}^\star,\quad \int_{\mathcal{M}_1(\mathbb{R})}\mathcal{L}_{\mathrm{FVc}}^{\lambda\left(\zeta_t^R\right)}P_{f_q,n}\left(\mu\right)\gamma_t^R\left(\mathrm{d}\mu\right)=0.$$

From (42) and Lemma 6.1, we have that $\int_{\mathcal{M}_1(\mathbb{R})}M_4\left(\mu\right)\gamma_t^R(\mathrm{d}\mu)<\infty$. As for all $q\in\mathbb{N}^\star$,

$$\mathcal{L}_{\mathrm{FVc}}^{\lambda\left(\zeta_t^R\right)}P_{f_q,n}\left(\mu\right)\leqslant C_1\left(1+M_2\left(\mu\right)\right)\left\|f_q\right\|_{W_0^{2,\infty}}.$$

for some constant $C_1>0$, we obtain by the dominated convergence theorem that

$$\mathbb{Q}-\text{a.s.},\ \mathrm{d}t-\text{a.e.},\ \forall f\in\mathscr{C}_K^2\left(\mathbb{R}^n,\mathbb{R}\right),\quad \int_{\mathcal{M}_1(\mathbb{R})}\mathcal{L}_{\mathrm{FVc}}^{\lambda\left(\zeta_t^R\right)}P_{f,n}\left(\mu\right)\gamma_t^R\left(\mathrm{d}\mu\right)=0. \tag{48}$$

Let us consider $f\in\mathscr{C}_{\|\cdot\|}^2\left(\mathbb{R}^n,\mathbb{R}\right)$ and for all $q\in\mathbb{N}^\star$, $x\mapsto\chi_q(x)=\exp\left(-\frac{1}{q^2-\|x\|_\infty^2}\right)\mathbb{1}_{\|x\|_\infty<q}$ of class $\mathscr{C}^\infty\left(\mathbb{R}^n,\mathbb{R}\right)$ with compact support. Then, for all $q\in\mathbb{N}^\star$, $f\chi_q\in\mathscr{C}_K^2\left(\mathbb{R}^n,\mathbb{R}\right)$. Noting that, $\left|\int_{\mathcal{M}_1(\mathbb{R})}\mathcal{L}_{\mathrm{FVc}}^{\lambda\left(\zeta_t^R\right)}P_{f\chi_q,n}\left(\mu\right)\gamma_t^R\left(\mathrm{d}\mu\right)\right|$ is dominated by $C_2\int_{\mathcal{M}_1(\mathbb{R})}\left(1+M_4\left(\mu\right)\right)\gamma_t^R(\mathrm{d}\mu)$ for some constant $C_2>0$, and since $\int_{\mathcal{M}_1(\mathbb{R})}M_4\left(\mu\right)\gamma_t^R(\mathrm{d}\mu)<\infty$, we deduce from (48) applied to $f\chi_q$, by the dominated convergence theorem, that

$$\mathbb{Q}-\text{a.s.},\ \mathrm{d}t-\text{a.e.},\ \forall f\in\mathscr{C}_{\|\cdot\|}^2\left(\mathbb{R}^n,\mathbb{R}\right),\quad \int_{\mathcal{M}_1(\mathbb{R})}\mathcal{L}_{\mathrm{FVc}}^{\lambda\left(\zeta_t^R\right)}P_{f,n}\left(\mu\right)\gamma_t^R\left(\mathrm{d}\mu\right)=0.$$

Therefore, it follows from Lemma 6.4 that $\mathbb{Q}-$a.s, $\mathrm{d}t-$a.e., $\gamma_t^R\left(\mathrm{d}\mu\right)=\pi^{\lambda\left(\zeta_t^R\right)}\left(\mathrm{d}\mu\right)$ which concludes the proof.

6.2 Proof of Lemma 6.4

In Section 6.2.1 we extend some duality results for the centered Fleming-Viot process, obtained in [18], which are be useful to prove Lemma 6.4 in Section 6.2.2.

6.2.1 Extension of the duality result for the centered Fleming-Viot process

Let us recall that the dual process $(\xi_t)_{t\geqslant0}$ of the centered Fleming-Viot process $(X_t)_{t\geqslant0}$ with resampling rate λ, on the state space $\bigcup_{n\in\mathbb{N}^\star}\mathscr{C}^2\left(\mathbb{R}^n,\mathbb{R}\right)$, obtained in [18, Section 3.2] is defined as below.

Let us consider $M:=(M(t))_{t\geqslant0}$ a Markov's birth and death process in $\mathbb{N}$ whose transition rates $q_{i,j}$ from i to j are given by:

$$\textbf{(1)}\ \ q_{n,n+1}=\lambda n^2\qquad \textbf{(2)}\ \ q_{n,n-1}=\lambda n(n-1)\qquad \textbf{(3)}\ \ q_{i,j}=0\ \text{otherwise.}$$

For all $M(0)\in\mathbb{N}^\star$, $\xi_0\in\mathscr{C}_b^2\left(\mathbb{R}^{M(0)},\mathbb{R}\right)$ and $\lambda>0$, we define

$$\xi_t:=T_\lambda^{(M(\tau_n))}\left(t-\tau_n\right)\Lambda_nT_\lambda^{(M(\tau_{n-1}))}\left(\tau_n-\tau_{n-1}\right)\Lambda_{n-1}\cdots\Lambda_1T_\lambda^{(M(0))}\left(\tau_1\right)\xi_0,\qquad \tau_n\leqslant t<\tau_{n+1},\ n\in\mathbb{N}, \tag{49}$$

where $(\tau_n)_{n\in\mathbb{N}}$ is the sequence of jump times of M with $\tau_0 = 0$, $\left(T_\lambda^{(n)}(t)\right)_{t\geqslant 0}$ is the semi-group of operator associated to the generator $B_\lambda^{(n)}$ given by (23) and where $(\Lambda_n)_{n\in\mathbb{N}}$ is a sequence of random operators. These are conditionally independent given M and satisfy for all $k\in\mathbb{N}$, $n\geqslant 1$ and $1\leqslant i\neq j\leqslant n$,

$$\mathbb{P}\left(\Lambda_k = \Phi_{i,j}\,\middle|\,\left\{M\left(\tau_k^-\right) = n, M\left(\tau_k\right) = n-1\right\}\right) = \frac{1}{n(n-1)}$$

and for all $n\geqslant 1$ and $1\leqslant i,j\leqslant n$,

$$\mathbb{P}\left(\Lambda_k = K_{i,j}\,\middle|\,\left\{M\left(\tau_k^-\right) = n, M\left(\tau_k\right) = n+1\right\}\right) = \frac{1}{n^2},$$

where $\Phi_{i,j}$ and $K_{i,j}$ are respectively defined in (24) and (25). Note that if $M(0) = 1$, the dual process can only jump from ξ_0 to $K_{ij}\xi_0$. As for all $t\geqslant 0$, $\xi_t\in\mathscr{C}^2\left(\mathbb{R}^{M(t)},\mathbb{R}\right)$, $K_{i,j}\xi_t$ is well defined. Moreover, the dual process $(\xi_t)_{t\geqslant 0}$ with initial condition ξ_0 is constructed on the same probability space and independently of the centered Fleming-Viot process $(X_t)_{t\geqslant 0}$ with resampling rate λ and initial condition $\mu\in\mathcal{M}_1^{c,2}(\mathbb{R})$. We shall denote by $\mathbb{P}_{(\mu,\xi_0)}$, the law of the couple $((X_t,\xi_t))_{t\geqslant 0}$ on this probability space.

We denote by S_t the number of jumps of the process M on $[0,t]$. We start with a result giving bounds on the dual process, which is an extension of similar estimates obtained in [18, Lemma 6.4].

Lemma 6.5. *For all $\xi_0\in\mathscr{C}_b^2\left(\mathbb{R}^{M(0)},\mathbb{R}\right)$ there exists $C_0:\bigcup_{k\in\mathbb{N}^\star}(0,+\infty)^{k-1}\times\{k\}\to\mathbb{R}_+$, locally bounded, such that*

$$C_0(\tau_1,1) := \|\xi_0\|_\infty + \frac{3}{M(0)}\|\mathrm{Hess}\,(\xi_0)\|_\infty$$

and, for all $(t_j)_{j\in\mathbb{N}}\in(0,+\infty)^{\mathbb{N}}$, $k\mapsto C_0\left((t_i)_{0\leqslant i\leqslant k-1},k\right)$ is non-decreasing and satisfying

$$\begin{aligned}&\forall k\in\mathbb{N}^\star,\ \forall t\leqslant\tau_k,\ \forall x\in\mathbb{R}^{M(t)},\\ &\qquad\qquad |\xi_t(x)|\leqslant C_0\left((\tau_{i+1}-\tau_i)_{0\leqslant i\leqslant k-1},k\right)\left(1+\|x\|_\infty^{2S_t}\right).\end{aligned}\tag{50}$$

Proof. By induction on $k\in\mathbb{N}^\star$, we prove the property

$$(\mathcal{Q}_k):\quad \forall x\in\mathbb{R}^{M(\tau_k)},\quad |\xi_{\tau_k}(x)|\leqslant C_0((\tau_{i+1}-\tau_i)_{0\leqslant i\leqslant k-1},k)\left(1+\|x\|_\infty^{2k}\right).$$

Step 1. Initial case: computation of $C_0(\tau_1,1)$. Let us recall some notations of [18, Theorem 6.1]. For all $n\in\mathbb{N}^\star$, we denote by $\mathbf{1}\in\mathbb{R}^n$, the vector whose coordinates are all 1. For all $n\in\mathbb{N}^\star$, $t\geqslant 0$, $x\in\mathbb{R}^n$ and $\lambda>0$, let us consider $g_{t,x}^\lambda$ the density of the Gaussian distribution $\mathcal{N}^{(n)}\left(m_{t,x}^\lambda,\Sigma_t^\lambda\right)$ where $m_{t,x}^\lambda := x - \frac{(1-\exp(-2\lambda nt))}{n}(x\cdot\mathbf{1})\mathbf{1}$ and $\Sigma_t^\lambda := P\sigma_t^\lambda P^{-1}$ with

$$\sigma_t^\lambda := \begin{pmatrix} \frac{1-\exp(-4\lambda nt)}{4\lambda n} & 0 & \dots & \dots & 0\\ 0 & t & 0 & \dots & 0\\ \vdots & 0 & \ddots & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & 0\\ 0 & 0 & \dots & 0 & t\end{pmatrix} \quad\text{and}\quad P := \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{2}} & \cdots & \cdots & \frac{1}{\sqrt{n(n-1)}}\\ \vdots & -\sqrt{\frac{1}{2}} & \ddots & \cdots & \frac{1}{\sqrt{n(n-1)}}\\ \vdots & 0 & \ddots & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & \frac{1}{\sqrt{n(n-1)}}\\ \frac{1}{\sqrt{n}} & 0 & \cdots & 0 & -\sqrt{\frac{n-1}{n}}\end{pmatrix}.$$

Note that for all $i,j \in \{1,\cdots,n\}$, $\partial_{x_i} m^{\lambda}_{t,x} = \epsilon_i - \frac{1-\exp(-2\lambda n t)}{n}\mathbf{1}$ and $\partial^2_{x_i x_j} m^{\lambda}_{t,x} = 0$ where $(\epsilon_1,\cdots,\epsilon_n)$ is the canonical basis of $\mathbb{R}^n$. For all $\lambda > 0$, the key identity for the sequel is

$$\forall f \in L^{\infty}(\mathbb{R}^n),\ \forall t \geqslant 0,\ \forall x \in \mathbb{R}^n, \quad T^{(M(0))}_{\lambda}(t) f(x) = \left(f * g^{\lambda}_{t,0}\right)\left(m^{\lambda}_{t,x}\right) \tag{51}$$

where $*$ stands for the convolution product (see (46) in [18, Theorem 6.1]).

On the one hand, note that for all $t < \tau_1$, $M(t) = M(0)$. In this case, for all $x \in \mathbb{R}^{M(0)}$, $\lambda > 0$, $\xi_t(x) = T^{(M(0))}_{\lambda}(t)\xi_0(x)$ and so, from [18, Theomem 6.1], $|\xi_t(x)| \leqslant \|\xi_0\|_{\infty}$. On the other hand, at time τ_1, we make a partition of cases according to whether the dual process loses or gains a variable. Let $i,j \in \{1,\cdots,M(0)\}$ be fixed.

- *Case* $\Lambda_1 = \Phi_{i,j}$. In this case, $M(\tau_1) = M(0) - 1$ and we deduce from (49) that

$$\forall x \in \mathbb{R}^{M(0)-1}, \qquad \xi_{\tau_1}(x) = \Phi_{i,j} T^{(M(0))}_{\lambda}(\tau_1)\,\xi_0(x).$$

  By (24), we deduce that for all $x \in \mathbb{R}^{M(0)-1}$, $|\xi_{\tau_1}(x)| \leqslant \|\xi_0\|_{\infty}$.

- *Case* $\Lambda_1 = K_{i,j}$. In this case, $M(\tau_1) = M(0) + 1$ and we deduce from (49) and (25) then from [18, Theorem 6.1 **(3)**] and properties of the convolution product that for all $x \in \mathbb{R}^{M(0)+1}$,

$$\begin{aligned}
\xi_{\tau_1}(x) &= K_{i,j} T^{(M(0))}_{\lambda}(\tau_1)\,\xi_0(x) \\
&= \partial_{x_i x_j} T^{(M(0))}_{\lambda}(\tau_1)\,\xi_0(\widetilde{x})\, x^2_{M(0)+1} \\
&= \left(\partial_{x_j} m^{\lambda}_{\tau_1,\widetilde{x}}\right)^t \left[\left(\mathrm{Hess}\,(\xi_0) * g^{\lambda}_{\tau_1,0}\right)\left(m^{\lambda}_{\tau_1,\widetilde{x}}\right)\partial_{x_i} m^{\lambda}_{t,\widetilde{x}}\right] x^2_{M(0)+1}
\end{aligned}$$

  where $\widetilde{x} = \left(x_1,\cdots,x_{M(0)}\right) \in \mathbb{R}^{M(0)}$. As,

$$\left(\mathrm{Hess}\,(\xi_0) * g^{\lambda}_{\tau_1,0}\right)\left(m^{\lambda}_{\tau_1,\widetilde{x}}\right) = \left(\int_{\mathbb{R}^{M(0)}} \partial^2_{ij}\xi_0(u) g^{\lambda}_{\tau_1,0}\left(m^{\lambda}_{\tau_1,\widetilde{x}} - u\right)\mathrm{d}u\right)_{1\leqslant i,j\leqslant M(0)}$$

  we obtain that $\left\|\left(\mathrm{Hess}\,(\xi_0) * g^{\lambda}_{\tau_1,0}\right)\left(m^{\lambda}_{\tau_1,\widetilde{x}}\right)\right\|_{\infty} \leqslant \left\|\partial^2_{ij}\xi_0\right\|_{\infty}$. Noting that

$$\left|\left(\partial_{x_j} m^{\lambda}_{\tau_1,\widetilde{x}}\right)^t \partial_{x_i} m^{\lambda}_{t,\widetilde{x}}\right| \leqslant \frac{3}{M(0)},$$

  we deduce that for all $x \in \mathbb{R}^{M(0)+1}$,

$$|\xi_{\tau_1}(x)| \leqslant \frac{3}{M(0)} \sup_{1\leqslant i,j\leqslant M(0)} \left\|\partial^2_{ij}\xi_0\right\|_{\infty} x^2_{M(0)+1} \leqslant \frac{3}{M(0)} \|\mathrm{Hess}(\xi_0)\|_{\infty} \|x\|^2_{\infty}$$

  and $(\mathcal{Q}_1)$ follows.

Step 2. Inductive Step. We assume that, for $k \in \mathbb{N}\setminus\{0,1\}$, $(\mathcal{Q}_{k-1})$ is satisfied and prove that $(\mathcal{Q}_k)$ is also. We make again a partition of cases according to whether the dual process loses or gains a variable. Let $i,j \in \{1,\cdots,M(\tau_{k-1})\}$ be fixed.

- *Case $\Lambda_k = \Phi_{i,j}$ at the k^{th} jump.* In this case, $M(\tau_k) = M(\tau_{k-1}) - 1$ and we deduce from (49) that for all $x \in \mathbb{R}^{M(\tau_{k-1})-1}$, $\xi_{\tau_k}(x) = \Phi_{i,j} T_\lambda^{(M(\tau_{k-1}))}(\tau_k - \tau_{k-1})\xi_{\tau_{k-1}}(x)$. By using (24) and $(\mathcal{Q}_{k-1})$, we deduce from [18, Corollary 6.2 **(1)**] that for all $x \in \mathbb{R}^{M(\tau_{k-1})-1}$,

$$\begin{aligned}&\left|\Phi_{i,j} T_\lambda^{(M(\tau_{k-1}))}\left(\tau_k - \tau_{k-1}\right)\xi_{\tau_{k-1}}(x)\right| \\ &\qquad \leqslant C_2\left(\tau_k - \tau_{k-1}, M\left(\tau_{k-1}\right)\right) C_0\left(\left(\tau_{i+1} - \tau_i\right)_{0\leqslant i\leqslant k-1}, k\right)\left(1 + \|x\|_\infty^{2(k-1)}\right),\end{aligned}$$

where C_2C_0 is locally bounded. As $t \mapsto C_2\left(t, M\left(\tau_{k-1}\right)\right)$ is non-decreasing, we deduce $(\mathcal{Q}_k)$ in that case.

- *Case $\Lambda_k = K_{i,j}$ at the k^{th} jump.* In this case, $M(\tau_k) = M(\tau_{k-1}) + 1$. From (49), (25) and [18, Theorem 6.1 **(3)**], we have for all $x \in \mathbb{R}^{M(\tau_{k-1})+1}$,

$$\begin{aligned}\left|\xi_{\tau_k}(x)\right| &= \left|K_{i,j} T_\lambda^{(M(\tau_{k-1})+1)}\left(\tau_k - \tau_{k-1}\right)\xi_{\tau_{k-1}}(x)\right| \\ &= \left|\left(\partial_{x_j} m^\lambda_{\tau_k - \tau_{k-1}, \widetilde{x}}\right)^t \left[\left(\xi_{\tau_{k-1}} * \operatorname{Hess}\left(g^\lambda_{\tau_k - \tau_{k-1}, 0}\right)\right)(m^\lambda_{\tau_k - \tau_{k-1}, \widetilde{x}})\partial_{x_i} m^\lambda_{\tau_k - \tau_{k-1}, \widetilde{x}}\right]\right| \\ &\qquad \times x^2_{M(\tau_{k-1})+1},\end{aligned}$$

where $\widetilde{x} = \left(x_1, \cdots, x_{M(\tau_{k-1})}\right)^t \in \mathbb{R}^{M(\tau_{k-1})}$. From $(\mathcal{Q}_{k-1})$ and [18, Corollary 6.2 **(2)**], we deduce that

$$\begin{aligned}&\left|K_{i,j} T_\lambda^{(M(\tau_{k-1})+1)}\left(\tau_k - \tau_{k-1}\right)\xi_{\tau_{k-1}}(x)\right| \\ &\qquad \leqslant C_3\left(\tau_k - \tau_{k-1}, M\left(\tau_{k-1}\right)\right) C_0\left(\left(\tau_{i+1} - \tau_i\right)_{0\leqslant i\leqslant k-1}, k\right)\left(1 + \|x\|_\infty^{2k}\right),\end{aligned}$$

where C_3C_0 is locally bounded and $(\mathcal{Q}_k)$ follows in that case.

We conclude by the principle of induction.

Step 3. Proof of (50) for $t < \tau_k$. Note that from (49), for all $k \in \mathbb{N}$ and $t \in (\tau_k, \tau_{k+1})$,

$$\forall x \in \mathbb{R}^{M(t)}, \qquad \xi_t(x) = T_\lambda^{(M(\tau_k))}\left(t - \tau_k\right)\xi_{\tau_k}(x),$$

so the announced result follows from [18, Corollary 6.2 **(1)**]. □

Let us consider for all $k \in \mathbb{N}^\star$, $\ell, m \in \mathbb{N}$ the stopping times

$$\begin{aligned}\vartheta_{k,\ell} &:= \inf\left\{t \geqslant 0 \;\middle|\; S_t \geqslant k \quad \text{or} \quad \exists s \in [0,t],\ \left\langle \xi_s, X_{t-s}^{M(s)}\right\rangle \geqslant \ell\right\}, \\ \vartheta'_{k,m} &:= \inf\left\{t \geqslant 0 \;\middle|\; S_t \geqslant k \quad \text{or} \quad C_0\left(\left(\tau_{i+1} - \tau_i\right)_{0\leqslant i\leqslant k-1}, k\right) \geqslant m\right\}, \\ \vartheta'_{k,\ell,m} &:= \inf\left\{t \geqslant 0 \;\middle|\; S_t \geqslant k \quad \text{or} \quad \exists s \in [0,t],\ \left\langle \xi_s, X_{t-s}^{M(s)}\right\rangle \geqslant \ell\right. \\ &\qquad\qquad \left. \text{or} \quad C_0\left(\left(\tau_{i+1} - \tau_i\right)_{0\leqslant i\leqslant k-1}, k\right) \geqslant m\right\}.\end{aligned}$$

As $(M(t),\xi_t)_{t\geqslant 0}$ is independent of $(X_t)_{t\geqslant 0}$, note that for all $k\in\mathbb{N}^\star$, $m\in\mathbb{N}$, $\vartheta'_{k,m}$ is independent of $(X_t)_{t\geqslant 0}$. Recall that for all $t\geqslant 0$, $\xi_t\in\mathscr{C}^2\left(\mathbb{R}^{M(t)},\mathbb{R}\right)$. For all $n\geqslant M(t)$, for all $x\in\mathbb{R}^n$, we denote by $\xi_t^{(n)}(x)=\xi_t^{(n)}(x_1,\cdots,x_n):=\xi_t\left(x_1,\cdots,x_{M(t)}\right)$ so that for all $\mu\in\mathcal{M}_1^{c,2}(\mathbb{R})$,

$$\left\langle \xi_t^{(n)},\mu^n\right\rangle=\left\langle \xi_t,\mu^{M(t)}\right\rangle.$$

The next lemma is an extension of the duality identity proved in [18].

Lemma 6.6. *Given $(X_t)_{t\geqslant 0}$, $(\xi_t)_{t\geqslant 0}$ as above with $X_0:=\mu\in\mathcal{M}_1^{c,2k}(\mathbb{R})$, $k\in\mathbb{N}^\star$ and $\xi_0\in\mathscr{C}_b^2\left(\mathbb{R}^{M(0)},\mathbb{R}\right)$, we have that $\mathbb{P}_{(\mu,\xi_0)}-$a.s., for all $t\geqslant 0$, for all $m\in\mathbb{N}$,*

$$\mathbb{E}_{(\mu,\xi_0)}\left(\left\langle \xi_0,X_{t\wedge\vartheta'_{k,m}}^{M(0)}\right\rangle\right)=\left\langle \mathbb{E}_{\xi_0}\left(\xi_{t\wedge\vartheta'_{k,m}}^{(M(0)+k)}\exp\left(\lambda\int_0^{t\wedge\vartheta'_{k,m}}M^2(u)\mathrm{d}u\right)\right),\mu^{M(0)+k}\right\rangle. \tag{52}$$

Compared to the duality identity of [18], the key point of this lemma is that the right-hand side of the last equation, is a polynomial in μ.

Proof. Let us consider $t\geqslant 0$, $k\in\mathbb{N}^\star$, $\ell,m\in\mathbb{N}$. As $\vartheta'_{k,\ell,m}\leqslant\vartheta_{k,\ell}$, the weak duality identity of [18, Theorem 3.4] implies that:

$$\begin{aligned}
&\mathbb{E}_{(\mu,\xi_0)}\left(\left\langle \xi_0,X_{t\wedge\vartheta'_{k,\ell,m}}^{M(0)}\right\rangle\right)\\
&\qquad=\mathbb{E}_{(\mu,\xi_0)}\left(\left\langle \xi_{t\wedge\vartheta'_{k,\ell,m}},\mu^{M\left(t\wedge\vartheta'_{k,\ell,m}\right)}\right\rangle\exp\left(\lambda\int_0^{t\wedge\vartheta'_{k,\ell,m}}M^2(u)\mathrm{d}u\right)\right)\\
&\qquad=\mathbb{E}_{(\mu,\xi_0)}\left(\left\langle \xi_{t\wedge\vartheta'_{k,\ell,m}}^{(M(0)+k)},\mu^{M(0)+k}\right\rangle\exp\left(\lambda\int_0^{t\wedge\vartheta'_{k,\ell,m}}M^2(u)\mathrm{d}u\right)\right).
\end{aligned}$$

Now, from Lemma 6.5 and the definition of $\vartheta'_{k,\ell,m}$, note that for all $x\in\mathbb{R}^{M(0)+k}$,

$$\begin{aligned}
&\mathbb{E}_{(\mu,\xi_0)}\left(\xi_{t\wedge\vartheta'_{k,\ell,m}}^{(M(0)+k)}(x)\exp\left(\lambda\int_0^{t\wedge\vartheta'_{k,\ell,m}}M^2(u)\mathrm{d}u\right)\right)\\
&\qquad\qquad\leqslant m\left(1+\|x\|_\infty^{2k}\right)\exp\left(\lambda t\left(M(0)+k\right)^2\right).
\end{aligned}\tag{53}$$

Since $\mu\in\mathcal{M}_1^{c,2k}(\mathbb{R})$, we deduce from Fubini's theorem that

$$\begin{aligned}
&\mathbb{E}_{(\mu,\xi_0)}\left(\left\langle \xi_0,X_{t\wedge\vartheta'_{k,\ell,m}}^{M(0)}\right\rangle\right)\\
&\qquad=\left\langle \mathbb{E}_{(\mu,\xi_0)}\left(\xi_{t\wedge\vartheta'_{k,\ell,m}}^{(M(0)+k)}\exp\left(\lambda\int_0^{t\wedge\vartheta'_{k,\ell,m}}M^2(u)\mathrm{d}u\right)\right),\mu^{M(0)+k}\right\rangle.
\end{aligned}$$

On the one hand, from [18, Section 6.2.1], $\lim_{\ell\to+\infty}\vartheta'_{k,\ell,m}=\vartheta'_{k,m}$ $\mathbb{P}_{(\mu,\xi_0)}-$a.s., and since $(X_t)_{t\geqslant 0}$ has continous paths for the topology of weak convergence, we have $\mathbb{P}_{(\mu,\xi_0)}-$a.s.

$$\lim_{\ell\to+\infty}\left\langle \xi_0,X_{t\wedge\vartheta'_{k,\ell,m}}^{M(0)}\right\rangle=\left\langle \xi_0,X_{t\wedge\vartheta'_{k,m}}^{M(0)}\right\rangle.$$

Therefore, we deduce from the LEBESGUE dominated convergence theorem, that for all $\xi_0 \in \mathscr{C}_b^2\left(\mathbb{R}^{M(0)}, \mathbb{R}\right)$,

$$\lim_{\ell \to +\infty} \mathbb{E}_{(\mu,\xi_0)}\left(\left\langle \xi_0, X^{M(0)}_{t\wedge\vartheta'_{k,\ell,m}}\right\rangle\right) = \mathbb{E}_{(\mu,\xi_0)}\left(\left\langle \xi_0, X^{M(0)}_{t\wedge\vartheta'_{k,m}}\right\rangle\right).$$

On the other hand, Lemma 6.5 and the dominated convergence theorem imply that for all $x \in \mathbb{R}^{M(0)+k}$,

$$\begin{aligned}\lim_{\ell \to +\infty} \mathbb{E}_{(\mu,\xi_0)}&\left(\xi^{(M(0)+k)}_{t\wedge\vartheta'_{k,\ell,m}}(x)\exp\left(\lambda\int_0^{t\wedge\vartheta'_{k,\ell,m}} M^2(u)\mathrm{d}u\right)\right)\\ &= \mathbb{E}_{(\mu,\xi_0)}\left(\xi^{(M(0)+k)}_{t\wedge\vartheta'_{k,m}}(x)\exp\left(\lambda\int_0^{t\wedge\vartheta'_{k,m}} M^2(u)\mathrm{d}u\right)\right)\end{aligned}$$

and the limit satisfies the inequality (53). As $\mu \in \mathcal{M}_1^{c,2k}(\mathbb{R})$, we deduce again from the dominated convergence theorem that

$$\begin{aligned}\lim_{\ell \to +\infty} \Big\langle \mathbb{E}_{(\mu,\xi_0)}&\left(\xi^{(M(0)+k)}_{t\wedge\vartheta'_{k,\ell,m}}\exp\left(\lambda\int_0^{t\wedge\vartheta'_{k,\ell,m}} M^2(u)\mathrm{d}u\right)\right), \mu^{M(0)+k}\Big\rangle\\ &= \left\langle \mathbb{E}_{(\mu,\xi_0)}\left(\xi^{(M(0)+k)}_{t\wedge\vartheta'_{k,m}}\exp\left(\lambda\int_0^{t\wedge\vartheta'_{k,m}} M^2(u)\mathrm{d}u\right)\right), \mu^{M(0)+k}\right\rangle\end{aligned}$$

and the announced result follows from the fact that $\vartheta'_{k,m}$ is independent of $(X_t)_{t\geqslant 0}$, which implies

$$\begin{aligned}\mathbb{E}_{(\mu,\xi_0)}&\left(\xi^{(M(0)+k)}_{t\wedge\vartheta'_{k,m}}(x)\exp\left(\lambda\int_0^{t\wedge\vartheta'_{k,m}} M^2(u)\mathrm{d}u\right)\right)\\ &= \mathbb{E}_{\xi_0}\left(\xi^{(M(0)+k)}_{t\wedge\vartheta'_{k,m}}(x)\exp\left(\lambda\int_0^{t\wedge\vartheta'_{k,m}} M^2(u)\mathrm{d}u\right)\right).\end{aligned}$$

□

From Lemma 6.5, we deduce that we can choose $m \in \mathbb{N}$ large enough such that $\vartheta'_{1,m} = \tau_1$. In this case, the function inside the brackets in the left-hand side of (52) is defined, for all $t \geqslant 0$ and for all $x \in \mathbb{R}^{M(0)+1}$, by

$$\psi_t(x) := \mathbb{E}_{(\mu,\xi_0)}\left(\xi^{(M(0)+1)}_{t\wedge\tau_1}(x)\exp\left(\lambda\int_0^{t\wedge\tau_1} M^2(u)\mathrm{d}u\right)\right).$$

For all $t \geqslant 0$, $n \in \mathbb{N}^\star$, $f \in \mathscr{C}_b^2\left(\mathbb{R}^n, \mathbb{R}\right)$ and $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ let us denote

$$P_{\mathrm{FVc}}(t)P_{f,n}\left(\mu\right) := \mathbb{E}_\mu\left(P_{f,n}\left(X_t\right)\right)$$

the semi-group of the centered FLEMING-VIOT process $(X_t)_{t\geqslant 0}$ with resampling rate λ.

The following result is the main result of this section. It gives an extension of Lemma 6.5 to a function which appears naturally in the proof of Lemma 6.4.

Proposition 6.7. *For all $t \geqslant 0$, for all $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$, for all $\xi_0 \in \mathscr{C}_b^2\left(\mathbb{R}^{M(0)}, \mathbb{R}\right)$*

$$P_{\mathrm{FVc}}(t) P_{\xi_0, M(0)}(\mu) = P_{V_t, M(0)+1}(\mu) := \left\langle V_t, \mu^{M(0)+1} \right\rangle$$

where

$$V_t := \exp(\alpha_0 t)\, \psi_t - \alpha_0 \int_0^t \exp(\alpha_0 s)\, \psi_s \mathrm{d}s \tag{54}$$

and $\alpha_0 := \lambda M(0)\left[2M(0) - 1\right]$. In addition, for all $t \geqslant 0$, there exist a constant $C_t > 0$ such that for all $x \in \mathbb{R}^{M(0)+1}$,

$$|V_t(x)| \leqslant C_t \left(1 + \|x\|_\infty^2\right).$$

Proof. Let $t \geqslant 0$, $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$, $\xi_0 \in \mathscr{C}_b^2\left(\mathbb{R}^{M(0)}, \mathbb{R}\right)$ be fixed. Note that from Lemma 6.5, there exists a constant $\widetilde{C}_t$, depending only on ξ_0 and t such that for all $x \in \mathbb{R}^{M(0)+1}$,

$$|\psi_t(x)| \leqslant C_0(\tau_1, 1)\left(1 + \|x\|_\infty^2\right) \mathbb{E}_{(\mu,\xi_0)}\left(\exp\left(\lambda t M^2(0)\right)\right) \leqslant \widetilde{C}_t\left(1 + \|x\|_\infty^2\right). \tag{55}$$

Now, from Lemma 6.6 and the fact that τ_1 is independent of $(X_t)_{t \geqslant 0}$, we obtain that

$$\left\langle \psi_t, \mu^{M(0)+1} \right\rangle = \mathbb{E}_{(\mu,\xi_0)}\left(P_{\xi_0, M(0)}\left(X_{t \wedge \tau_1}\right)\right) = \mathbb{E}_{\xi_0}\left(P_{\mathrm{FVc}}(t \wedge \tau_1) P_{\xi_0, M(0)}(\mu)\right).$$

As τ_1 follows an exponential law of parameter α_0 we have that

$$\begin{aligned} \left\langle \psi_t, \mu^{M(0)+1} \right\rangle = {} & P_{\mathrm{FVc}}(t) P_{\xi_0, M(0)}(\mu) \exp(-\alpha_0 t) \\ & + \alpha_0 \int_0^t P_{\mathrm{FVc}}(s) P_{\xi_0, M(0)}(\mu) \exp(-\alpha_0 s)\, \mathrm{d}s. \end{aligned} \tag{56}$$

Thanks to (55) and since μ has its moment of order 2 finite, the Fubini theorem ensures us that

$$\left\langle \int_0^t \exp(\alpha_0 s)\, \psi_s \mathrm{d}s, \mu^{M(0)+1} \right\rangle = \int_0^t \exp(\alpha_0 s) \left\langle \psi_s, \mu^{M(0)+1} \right\rangle \mathrm{d}s.$$

From (56) we deduce that

$$\begin{aligned} & \left\langle \int_0^t \exp(\alpha_0 s)\, \psi_s \mathrm{d}s, \mu^{M(0)+1} \right\rangle \\ & = \alpha_0 \int_0^t \exp(\alpha_0 s) \int_0^s \exp(-\alpha_0 u)\, P_{\mathrm{FVc}}(u) P_{\xi_0, M(0)} \mathrm{d}u \mathrm{d}s + \int_0^t P_{\mathrm{FVc}}(s) P_{\xi_0, M(0)}(\mu) \mathrm{d}s \\ & = \exp(\alpha_0 t) \int_0^t \exp(-\alpha_0 u)\, P_{\mathrm{FVc}}(u) P_{\xi_0, M(0)}(\mu) \mathrm{d}u. \end{aligned} \tag{57}$$

Then, the first announced result follows from (56) and (57) and the second one from (55). □

6.2.2 Proof of Lemma 6.4

The key point of this proof is to establish that for all $\xi_0 \in \mathscr{C}_b^4\left(\mathbb{R}^{M(0)}, \mathbb{R}\right)$, $t \geqslant 0$ and $\mu \in \mathcal{M}_1^{c,4}(\mathbb{R})$,

$$\begin{aligned} P_{\mathrm{FVc}}(t) P_{\xi_0, M(0)}(\mu) &= \left\langle \xi_0, \mu^{M(0)} \right\rangle + \int_0^t \mathcal{L}_{\mathrm{FVc}}^{\lambda} P_{\mathrm{FVc}}(s) P_{\xi_0, M(0)}(\mu) \mathrm{d}s \\ &= P_{\xi_0, M(0)}(\mu) + \int_0^t \mathcal{L}_{\mathrm{FVc}}^{\lambda} P_{V_s, M(0)+1}(\mu) \mathrm{d}s. \end{aligned} \tag{58}$$

where we use Proposition 6.7 in the second equality. To do this, we will first prove that for all $t \geqslant 0$, there exists a constant $C_t > 0$ such that for all $x \in \mathbb{R}^{M(0)+1}$, $\left\|\mathrm{Hess}\left(V_t\right)(x)\right\|_{\infty} \leqslant C_t\left(1 + \|x\|_{\infty}^2\right)$, so that $V_t \in \mathscr{C}_{\|\cdot\|}^2\left(\mathbb{R}^{M(0)+1}, \mathbb{R}\right)$.

For all $n \in \mathbb{N}^{\star}$, for any function $f : \mathbb{R}^n \to \mathbb{R}$, $k \in \mathbb{N}$ times differentiable, we denote by $D^k f$ the differential of order k of f. Let $t \geqslant 0$, $\lambda > 0$, $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ and $\xi_0 \in \mathscr{C}_b^4\left(\mathbb{R}^{M(0)}, \mathbb{R}\right)$ be fixed.

Step 1. Preliminary bounds. Let $x \in \mathbb{R}^{M(0)}$. It follows from (51) that

$$\begin{aligned} &\partial_{x_i x_j x_k x_\ell}^4 T^{(M(0))}(t) \xi_0(x) \\ &\quad = \sum_{m,p,q,r=1}^{M(0)} \partial_{x_i}\left(m_{t,x}^{\lambda}\right)_m \partial_{x_j}\left(m_{t,x}^{\lambda}\right)_p \partial_{x_k}\left(m_{t,x}^{\lambda}\right)_q \partial_{x_\ell}\left(m_{t,x}^{\lambda}\right)_r \left(g_{t,0}^{\lambda} * \partial_{y_m y_p y_q y_r}^4 \xi_0\right)\left(m_{t,x}^{\lambda}\right) \end{aligned}$$

where $\left(m_{t,x}^{\lambda}\right)_p$ designates the p^{th} component of $m_{t,x}^{\lambda}$. Then, for all $k \in \{0, \cdots, 4\}$, there exists a constant $C_1 > 0$ independent of x and t such that

$$\left\|D^k T^{(M(0))}(t) \xi_0(x)\right\|_{\infty} \leqslant C_1\left(1 + \left\|D^k \xi_0\right\|_{\infty}\right).$$

From (49), (24) and (25), we deduce that there exists a constant $C_2 > 0$ independent of t such that for all $x \in \mathbb{R}^{M(0)+1}$, for all $k \in \{0, 1, 2\}$,

$$\begin{aligned} \left\|D^k \xi_{t \wedge \tau_1}^{(M(0)+1)}(x)\right\|_{\infty} &\leqslant C_1\left(1 + \left\|D^k \xi_0\right\|_{\infty}\right) + \sum_{i,j=1}^{n} \left\|D^k K_{i,j} T_{\lambda}^{(M(0))}\left(\tau_1\right) \xi_0(x)\right\|_{\infty} \\ &\quad + \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left\|D^k \Phi_{i,j} T_{\lambda}^{(M(0))}\left(\tau_1\right) \xi_0\left(x_1, \cdots, x_{M(0)-1}\right)\right\|_{\infty} \\ &\leqslant C_2\left(1 + \left\|D^k \xi_0\right\|_{\infty} + \left\|D^{k+2} \xi_0\right\|_{\infty} x_{M(0)+1}^2\right). \end{aligned}$$

By the theorem of differentiation under the integral sign, for all $k \in \{0, 1, 2\}$ and $x \in \mathbb{R}^{M(0)+1}$, we deduce that

$$\begin{aligned} \left\|D^k \psi_t(x)\right\|_{\infty} &\leqslant \exp\left(\lambda M(0)^2 t\right) \mathbb{E}_{(\mu, \xi_0)}\left(\left\|D^k \xi_{t \wedge \tau_1}^{(M(0)+1)}(x)\right\|_{\infty}\right) \\ &\leqslant C_2 \exp\left(\lambda M(0)^2 t\right)\left(1 + \|x\|_{\infty}^2\right). \end{aligned}$$

Then, from the definition of V_t in (54) we deduce that there exists a constant C_t such that for all $k \in \{0, 1, 2\}$, for all $x \in \mathbb{R}^{M(0)+1}$, $\left\|D^k V_t(x)\right\|_{\infty} \leqslant C_t\left(1 + \|x\|_{\infty}^2\right)$, so that $V_t \in \mathscr{C}_{\|\cdot\|}^2\left(\mathbb{R}^{M(0)+1}, \mathbb{R}\right)$.

Step 2. Proof of the key point (58). Thanks to (27), for all $t \geqslant 0$,

$$P_{\mathrm{FVc}}(t)P_{\xi_0,M(0)}(\mu) = P_{\xi_0,M(0)}(\mu) + \mathbb{E}_\mu\left(\int_0^t \mathcal{L}^\lambda_{\mathrm{FVc}} P_{\xi_0,M(0)}\left(X_s\right)\mathrm{d}s\right).$$

From (26), there exists a constant $C_3 > 0$ such that for all $t \geqslant 0$, $\mathbb{E}_\mu\left(\left|\mathcal{L}^\lambda_{\mathrm{FVc}} P_{\xi_0,M(0)}\left(X_t\right)\right|\right) \leqslant C_3\left(1 + \mathbb{E}_\mu\left(M_2\left(X_t\right)\right)\right)$ which is finite from [18, Proposition 2.11] since $\mu \in \mathcal{M}_1^{c,4}(\mathbb{R})$. Hence, from Fubini's theorem we obtain that

$$P_{\mathrm{FVc}}(t)P_{\xi_0,M(0)}(\mu) = P_{\xi_0,M(0)}(\mu) + \int_0^t \mathbb{E}_\mu\left(\mathcal{L}^\lambda_{\mathrm{FVc}} P_{\xi_0,M(0)}\left(X_s\right)\right)\mathrm{d}s. \tag{59}$$

As $\mathcal{L}^\lambda_{\mathrm{FVc}} P_{\xi_0,M(0)}(\mu)$ is a polynomial in μ and since

$$\left|\mathcal{L}^\lambda_{\mathrm{FVc}}\left(\mathcal{L}^\lambda_{\mathrm{FVc}} P_{\xi_0,M(0)}(\mu)\right)\right| \leqslant C_4\left(1 + M_4\left(\mu\right)\right)$$

for some constant $C_4 > 0$, we deduce as in (47) that

$$\begin{aligned}\mathbb{E}_\mu\left(\mathcal{L}^\lambda_{\mathrm{FVc}} P_{\xi_0,M(0)}\left(X_t\right)\right) - \mathbb{E}_\mu\left(\mathcal{L}^\lambda_{\mathrm{FVc}} P_{\xi_0,M(0)}\left(X_0\right)\right)& \\ = \mathbb{E}_\mu\left(\int_0^t \mathcal{L}^\lambda_{\mathrm{FVc}}\left(\mathcal{L}^\lambda_{\mathrm{FVc}} P_{\xi_0,M(0)}\right)\left(X_s\right)\mathrm{d}s\right).&\end{aligned}$$

In particular, $t \mapsto \mathbb{E}_\mu\left(\mathcal{L}_{\mathrm{FVc}} P_{\xi_0,M(0)}\left(X_t\right)\right)$ is continuous, so

$$\begin{aligned}\mathbb{E}_\mu\left(\mathcal{L}^\lambda_{\mathrm{FVc}} P_{\xi_0,M(0)}\left(X_t\right)\right) &= \lim_{h\to 0}\frac{1}{h}\mathbb{E}_\mu\left(\int_t^{t+h} \mathcal{L}_{\mathrm{FVc}} P_{\xi_0,M(0)}\left(X_s\right)\mathrm{d}s\right)\\ &= \lim_{h\to 0}\frac{1}{h}\mathbb{E}_\mu\left(P_{\xi_0,M(0)}\left(X_{t+h}\right) - P_{\xi_0,M(0)}\left(X_t\right)\right)\\ &= \lim_{h\to 0}\frac{1}{h}\mathbb{E}_\mu\left(P_{\mathrm{FVc}}\left(t\right)P_{\xi_0,M(0)}\left(X_h\right)\right) - P_{\mathrm{FVc}}\left(t\right)P_{\xi_0,M(0)}(\mu)\\ &= \lim_{h\to 0}\frac{1}{h}\mathbb{E}_\mu\left(P_{V_t,M(0)+1}\left(X_h\right) - P_{V_t,M(0)+1}\left(\mu\right)\right)\end{aligned}$$

where we used Markov's property in the third equality and Proposition 6.7 in the last equality. From Step 1, we deduce that there exists a constant $C_5 > 0$ such that

$$\left|\mathcal{L}^\lambda_{\mathrm{FVc}} P_{V_t,M(0)+1}(\mu)\right| \leqslant C_5\left(1 + M_4\left(\mu\right)\right). \tag{60}$$

Thus, from the martingale problem (47) and the continuity of $t \mapsto \mathbb{E}_\mu\left(\mathcal{L}_{\mathrm{FVc}} P_{V_t,M(0)}\left(\mu\right)\right)$, we have that

$$\begin{aligned}\lim_{h\to 0}\frac{1}{h}\mathbb{E}_\mu\left(P_{V_t,M(0)+1}\left(X_h\right) - P_{V_t,M(0)+1}\left(\mu\right)\right) &= \lim_{h\to 0}\frac{1}{h}\mathbb{E}_\mu\left(\int_0^h \mathcal{L}^\lambda_{\mathrm{FVc}} P_{V_t,M(0)+1}\left(X_s\right)\mathrm{d}s\right)\\ &= \mathcal{L}^\lambda_{\mathrm{FVc}} P_{V_t,M(0)+1}\left(\mu\right).\end{aligned}$$

Therefore, we have proved that for all $t \geqslant 0$, for all $\mu \in \mathcal{M}_1^{c,4}(\mathbb{R})$,

$$\mathbb{E}_\mu\left(\mathcal{L}^\lambda_{\mathrm{FVc}} P_{\xi_0,M(0)}\left(X_t\right)\right) = \mathcal{L}^\lambda_{\mathrm{FVc}} P_{V_t,M(0)+1}(\mu),$$

and so, by (59), we obtain (58).

Step 3. Conclusion. From (58) and (60) and since $\int_{\mathcal{M}_1(\mathbb{R})} M_4(\mu)\,\gamma(\mathrm{d}\mu) < \infty$, we deduce that

$$\begin{aligned}
&\int_{\mathcal{M}_1(\mathbb{R})} P_{\mathrm{FVc}}(t) P_{\xi_0,M(0)}(\mu)\gamma(\mathrm{d}\mu) \\
&\qquad = \int_{\mathcal{M}_1(\mathbb{R})} P_{\xi_0,M(0)}(\mu)\gamma(\mathrm{d}\mu) + \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}^{\lambda}_{\mathrm{FVc}} P_{V_s,M(0)+1}(\mu)\gamma(\mathrm{d}\mu)\mathrm{d}s \\
&\qquad = \int_{\mathcal{M}_1(\mathbb{R})} P_{\xi_0,M(0)}(\mu)\gamma(\mathrm{d}\mu),
\end{aligned}$$

where we used the assumption of Lemma 6.4 in the last equality.

As the set of test functions $\left\{ P_{f,n} \,\middle|\, f \in \mathscr{C}_b^4(\mathbb{R}^n,\mathbb{R}), n \in \mathbb{N}^\star \right\}$ is $\mathcal{M}_1(\mathcal{M}_1(\mathbb{R}))-$ convergence determining [26, Lemma 2.1.2], it is $\mathcal{M}_1(\mathcal{M}_1(\mathbb{R}))-$separating [40, Chapter 3, Section 4, p.112], so we have for any bounded continuous function ϕ from $\mathcal{M}_1(\mathbb{R})$ to $\mathbb{R}$ that

$$\int_{\mathcal{M}_1(\mathbb{R})} P_{\mathrm{FVc}}(t)\phi(\mu)\gamma(\mathrm{d}\mu) = \int_{\mathcal{M}_1(\mathbb{R})} \phi(\mu)\gamma(\mathrm{d}\mu).$$

Hence γ is an invariant probability measure for the centered Fleming-Viot process with resampling rate λ. Now, from [18, Theorem 4.1], π^λ is its unique invariant probability measure which ends the proof of Lemma 6.4.

7 Characterisation of the limiting values of the slow component

Combining the results of Sections 5 and 6, we have proved that for all $R > 0$, the sequence of laws of $\left(\left(z^{K,R}, \Gamma^{K,R}\right)\right)_{K\in\mathbb{N}^\star}$ is tight in $\mathcal{M}_1\left(\mathbb{D}\left([0,T],\mathbb{T}_R\right) \times \mathcal{M}_m^T\left(\mathcal{M}_1(\mathbb{R})\right)\right)$ and for any limiting value $\mathbb{Q}$ of this sequence, the canonical process $\left(\zeta^R, \Gamma^R\right)$ on $\mathscr{C}^0\left([0,T],\mathbb{T}_R\right) \times \mathcal{M}_m^T\left(\mathcal{M}_1(\mathbb{R})\right)$ satisfies for all $f \in \mathscr{C}_b^1\left(\mathbb{T}_R,\mathbb{T}_R\right)$, for all $t \in [0,T]$, $\mathbb{Q}-$a.s.,

$$f\left(\zeta_t^R\right) = f\left(\zeta_0^R\right) + \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\mathrm{SLOW}} f\left(\zeta_s^R, \mu\right) \pi^{\lambda\left(\zeta_s^R\right)}(\mathrm{d}\mu)\,\mathrm{d}s. \tag{61}$$

From now on, we will establish in Section 7.1 with Lemma 7.1 that the sequence of laws of $\left(z^{K,R}\right)_{K\in\mathbb{N}^\star}$ converges weakly in $\mathbb{D}\left([0,T],\mathbb{T}_R\right)$ to the solution of an ODE in the torus. Finally, Section 7.2 allows us to get away from the torus and to prove Theorem 2.1.

7.1 Convergence of the slow component on the torus

Lemma 7.1. *The sequence $\left(z^{K,R}\right)_{K\in\mathbb{N}^\star}$ converges in law in $\mathbb{D}\left([0,T],\mathbb{T}_R\right)$ to the unique solution of*

$$\forall t \in [0,T], \quad z_t = x_0 + \frac{1}{2}\int_0^t \partial_1 \mathrm{Fit}\left(z_s, z_s\right)\beta\left(z_s\right) m_2\left(z_s\right)\mathrm{d}s \tag{62}$$

on the torus $\mathbb{T}_R$.

Proof. For any limiting value $\mathbb{Q}$ of the sequence of laws of $\left(z^{K,R}\right)_{K\in\mathbb{N}^\star}$, from (9) and applying (61) with $f=\mathrm{id}\in\mathscr{C}_b^2\left(\mathbb{T}_R,\mathbb{T}_R\right)$, we deduce that $\mathbb{Q}-$a.s.,

$$\forall t\in[0,T],\quad \zeta_t^R=\zeta_0^R+\int_0^t\left(\int_{\mathcal{M}_1(\mathbb{R})}M_2\left(\mu\right)\pi^{\lambda\left(\zeta_s^R\right)}(\mathrm{d}\mu)\right)\partial_1\mathrm{Fit}\left(\zeta_s^R,\zeta_s^R\right)\mathrm{d}s.$$

From [18, Corollary 4.16], for all $s\in[0,t]$,

$$\int_{\mathcal{M}_1(\mathbb{R})}M_2\left(\mu\right)\pi^{\lambda\left(\zeta_s^R\right)}\left(\mathrm{d}\mu\right)=\frac{1}{2\lambda\left(\zeta_s^R\right)}=\frac{\beta\left(\zeta_s^R\right)m_2\left(\zeta_s^R\right)}{2},$$

so ζ^R is solution of (62) $\mathbb{Q}-$a.s. By uniqueness of the solution of (62), the announced result follows. □

7.2 End of the proof of Theorem 2.1

Let $T>0$ be fixed and recall that x_0 is the mean trait value of ν_0^K. From Lemma 7.1 and since the ODE (62) is non-explosive, we can choose $R>0$ large enough such that

$$\lim_{K\to+\infty}\mathbb{P}\left(z_t^{K,R}\in\left[x_0-\frac{R}{2},x_0+\frac{R}{2}\right],\forall t\in[0,T]\right)=1.$$

From Proposition 4.8 and (7), we have that,

$$\lim_{K\to+\infty}\mathbb{P}\left(\mathrm{Diam}\left(\mathrm{Supp}\,\nu_t^{K,R}\right)\leqslant\frac{1}{K^{1+\varepsilon/2}},\forall t\in[0,T]\right)=1.$$

Hence,

$$\lim_{K\to+\infty}\mathbb{P}\left(\mathrm{Supp}\,\nu_{t/K\sigma_K^2}^{K,R}\subset\left[x_0-R,x_0+R\right],\forall t\in[0,T]\right)=1.$$

Now, on the event $\left\{\mathrm{Supp}\,\nu_{t/K\sigma_K^2}^{K,R}\subset\left[x_0-R,x_0+R\right],\forall t\in[0,T]\right\}$, $\nu_{t/K\sigma_K^2}^{K,R}=\nu_{t/K\sigma_K^2}^{K}$ for all $t\in[0,T]$, identifying $x\in\mathbb{T}_R$ with its unique representant in $[x_0-2R,x_0+2R]$. In particular,

$$\forall t\in[0,T],\quad z_t^{K,R}=z_t^K.$$

Theorem 2.1 then follows from Lemma 7.1.

Funding. This work was partially funded by the Chair "ModÃ©lisation MathÃ©matique et BiodiversitÃ©" of VEOLIA-Ecole Polytechnique-MNHN-F.X. N.C. was partially 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.

Acknowledgments. We thank the two anonymous referees for their useful comments.

References

[1] C.B. Anfinsen. Principles that Govern the Folding of Protein Chains. *Science*, 181(4096):223–230, 1973.

[2] M. Baar, A. Bovier, and N. Champagnat. From stochastic, individual-based models to the canonical equation of adaptive dynamics in one step. *The Annals of Applied Probability*, 27(2):1093–1170, 2017.

[3] K. Ball, T. G. Kurtz, L. Popovic, and G. Rempala. Asymptotic analysis of multiscale approximations to reaction networks. *The Annals of Applied Probability*, 16(4), 2006.

[4] G. Ballif, F. Clément, and R. Yvinec. Averaging of a Stochastic Slow-Fast Model for Population Dynamics: Application to the Development of Ovarian Follicles. *SIAM Journal on Applied Mathematics*, 82(1):359–380, 2022.

[5] V. Bansaye and S. Méléard. *Stochastic models for structured populations - Scaling limits and long time behavior*. Springer International Publishing, Cham, 2015.

[6] T. Benoist, C. Bernardin, R. Chetrite, R. Chhaibi, J. Najnudel, and C. Pellegrini. Emergence of Jumps in Quantum Trajectories via Homogenization. *Communications in Mathematical Physics*, 387(3):1821–1867, 2021.

[7] A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures*. American Mathematical Society, Providence, R.I, 2011.

[8] S. Billiard, R. Ferrière, S. Méléard, and V. C. Tran. Stochastic dynamics of adaptive trait and neutral marker driven by eco-evolutionary feedbacks. *Journal of Mathematical Biology*, 71(5):1211–1242, 2015.

[9] P. Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Statistics, Second edition. John Wiley & Sons, Inc., Hoboken, NJ, USA, 1999.

[10] G. Blankenship and G. C. Papanicolaou. Stability and Control of Stochastic Systems with Wide-band Noise Disturbances. I. *SIAM Journal on Applied Mathematics*, 34(3):437–476, 1978.

[11] J. Blath, T. Paul, and A. Tóbiás. A stochastic adaptive dynamics model for bacterial populations with mutation, dormancy and transfer. *ALEA, Latin American Journal of Probability and Mathematical Statistics*, 20(1):313, 2023.

[12] B. Bolker and S. W. Pacala. Using Moment Equations to Understand Stochastically Driven Spatial Pattern Formation in Ecological Systems. *Theoretical Population Biology*, 52(3):179–197, 1997.

[13] B. Bolker and S. W. Pacala. Spatial Moment Equations for Plant Competition: Understanding Spatial Strategies and the Advantages of Short Dispersal. *The American Naturalist*, 153(6):575–602, 1999.

[14] C. Bonnet and S. Méléard. Large fluctuations in multi-scale modeling for rest hematopoiesis. *Journal of Mathematical Biology*, 82(6):58, 2021.

[15] A. Bovier, L. Coquille, and C. Smadi. Crossing a fitness valley as a metastable transition in a stochastic population model. *The Annals of Applied Probability*, 29(6), 2019.

[16] N. Champagnat. A microscopic interpretation for adaptive dynamics trait substitution sequence models. *Stochastic Processes and their Applications*, 116(8):1127–1160, 2006.

[17] N. Champagnat, R. Ferrière, and S. Méléard. Unifying evolutionary dynamics: From individual stochastic processes to macroscopic models. *Theoretical Population Biology*, 69(3):297–321, 2006.

[18] N. Champagnat and V. Hass. Existence, uniqueness and ergodicity for the centered Fleming-Viot process. *Stochastic Processes and their Applications*, 166:104219, 2023.

[19] N. Champagnat, P.-E. Jabin, and S. Méléard. Adaptation in a stochastic multi-resources chemostat model. *Journal de Mathématiques Pures et Appliquées*, 101(6):755–788, 2014.

[20] N. Champagnat and S. Méléard. Polymorphic evolution sequence and evolutionary branching. *Probability Theory and Related Fields*, 151(1-2):45–94, 2011.

[21] N. Champagnat, S. Méléard, and V.C. Tran. Stochastic analysis of emergence of evolutionary cyclic behavior in population dynamics with transfer. *The Annals of Applied Probability*, 31(4), 2021.

[22] B. Charlesworth. *Evolution in Age-Structured Populations*. Cambridge University Press, 2 edition, 1994.

[23] L. Coquille, A. Kraut, and C. Smadi. Stochastic individual-based models with power law mutation rate on a general finite trait space. *Electronic Journal of Probability*, 26, 2021.

[24] M. Costa. *A piecewise deterministic model for a prey-predator community. The Annals of Applied Probability*, 26(6):3491–3530, 2016.

[25] C. Costantini and T. G. Kurtz. Diffusion approximation for transport processes with general reflection boundary conditions. *Mathematical Models and Methods in Applied Sciences*, 16(05):717–762, 2006.

[26] D.A. Dawson. *Measure-Valued Markov processes*. P.-L. Hennequin (editor) École d'Été de probabilités de Saint-Flour XXI. Lecture Notes in Mathematics, Vol. 1541 : 1-260, 1993.

[27] D.A. Dawson and K. Hochberg. Wandering random measures in the Fleming-Viot model. *The Annals of Probability*, 10(3):554–580, 1982.

[28] D. L. DeAngelis. *Individual-Based Models and Approaches in Ecology: Populations, Communities and Ecosystems*. Chapman and Hall/CRC, 1 edition, 2018.

[29] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[30] U. Dieckmann and R. Law. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of Mathematical Biology*, 34(5-6):579–612, 1996.

[31] U. Dieckmann and R. Law. Relaxation projections and the method of moments. pages 412–455. Cambridge University Press, Cambridge, 2000.

[32] O. Diekmann, P.-E. Jabin, S. Mischler, and B. Perthame. The dynamics of adaptation: An illuminating example and a Hamilton-Jacobi approach. *Theoretical Population Biology*, 67(4):257–271, 2005.

[33] M. Doebeli and U. Dieckmann. Speciation along environmental gradients. *Nature*, 421(6920):259–264, 2003.

[34] R. Durrett and J. Mayberry. Traveling waves of selective sweeps. *The Annals of Applied Probability*, 21(2), 2011.

[35] B. Ernande, U. Dieckmann, and M. Heino. Adaptive changes in harvested populations: plasticity and evolution of age and size at maturation. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1537):415–423, 2004.

[36] M. Esser and A. Kraut. A general multi-scale description of metastable adaptive motion across fitness valleys. 2021. Publisher: arXiv Version Number: 1.

[37] A. Etheridge. *An Introduction to Superprocesses*, volume 20 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2000.

[38] A. Etheridge. *Some Mathematical Models from Population Genetics*, volume 2012 of *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[39] S. N. Ethier and T.G. Kurtz. Fleming-Viot Processes in Population Genetics. *SIAM Journal on Control and Optimization*, 31(2):345–386, 1993.

[40] S.N. Ethier and T.G. Kurtz. *Markov Processes*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA, 1986.

[41] R. Ferriere, J. L. Bronstein, S. Rinaldi, R. Law, and M. Gauduchon. Cheating and the evolutionary stability of mutualisms. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269(1493):773–780, 2002.

[42] R. Ferrière and V. C. Tran. Stochastic and deterministic models for age-structured populations with genetically variable traits. *ESAIM: Proceedings*, 27:289–310, 2009.

[43] W. Fleming and M. Viot. Some measure-valued Markov processes in population genetics theory. *Indiana University Mathematics Journal*, 28(5):817–843, 1979.

[44] J.-P. Fouque, J. Garnier, G. Papanicolaou, and K. Sø lna. *Wave Propagation and Time Reversal in Randomly Layered Media*, volume 56 of *Stochastic Modelling and Applied Probability*. Springer New York, New York, NY, 2007.

[45] J. P. Fouque, G. Papanicolaou, R. Sircar, and K. Sø lna. Singular Perturbations in Option Pricing. *SIAM Journal on Applied Mathematics*, 63(5):1648–1665, 2003.

[46] N. Fournier and S. Méléard. *A microscopic probabilistic description of a locally regulated population and macroscopic approximations*. The Annals of Applied Probability, 14(4) : 1880 - 1919, 2004.

[47] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*, volume 260 of *Grundlehren der mathematischen Wissenschaften*. Springer US, New York, NY, 1984.

[48] J.-F. Le Gall. *Mouvement brownien, martingales et calcul stochastique*. Vol 71, Mathématiques et Applications, Springer-Verlag, 2013.

[49] A. Genadot. Averaging for some simple constrained Markov processes. *Probability and Mathematical Statistics*, 39(1):139–158, 2019.

[50] D. Giorgi, S. Kaakai, and V. Lemaire. *IBMPopSim: Individual Based Model Population Simulation*, 2023. https://github.com/DaphneGiorgi/IBMPopSim, https://DaphneGiorgi.github.io/IBMPopSim/.

[51] A. Gonzalez Casanova and C. Smadi. On Λ-Fleming-Viot processes with general frequency-dependent selection. *Journal of Applied Probability*, 57(4):1162–1197, 2020.

[52] V. Grimm and S. F. Railsback. *Individual-based Modeling and Ecology:*. Princeton University Press, 2005.

[53] A. Gupta, J. A. J. Metz, and V. C. Tran. A New Proof for the Convergence of an Individual Based Model to the Trait Substitution Sequence. *Acta Applicandae Mathematicae*, 131(1):1–27, 2014.

[54] J. Hofbauer and K. Sigmund. *The theory of evolution and dynamical systems: mathematical aspects of selection*. Number 7 in London Mathematical Society student texts. Cambridge University Press, Cambridge [England] ; New York, 1988.

[55] J. Hofbauer and K. Sigmund. Adaptive dynamics and evolutionary stability. *Applied Mathematics Letters*, 3(4):75–79, 1990.

[56] E. Kisdi. Evolutionary Branching under Asymmetric Competition. *Journal of Theoretical Biology*, 197(2):149–162, 1999.

[57] T.G. Kurtz. Averaging for martingale problems and stochastic approximation. In I. Karatzas and D. Ocone, editors, *Applied Stochastic Analysis, Lecture Notes in Control and Information Sciences*, volume 177, pages 186–209. Springer-Verlag, Berlin/Heidelberg, 1992.

[58] H. J. Kushner. Jump-Diffusion Approximations for Ordinary Differential Equations with Wide-Band Random Right Hand Sides. *SIAM Journal on Control and Optimization*, 17(6):729–744, 1979.

[59] H. J. Kushner. A martingale method for the convergence of a sequence of processes to a jump-diffusion process. *Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 53(2):207–219, 1980.

[60] H. J. Kushner. *Approximation and weak convergence methods for random processes, with applications to stochastic systems theory.* Number 6 in The MIT Press series in signal processing, optimization, and control. MIT Press, Cambridge, Mass, 1984.

[61] R. Lande. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution*, 33(1):402–416, 1979.

[62] R. Law and U. Dieckmann. Moment approximations of individual-based models. pages 252–270. Cambridge University Press, Cambridge, 2002.

[63] H. Leman. Convergence of an infinite dimensional stochastic process to a spatially structured trait substitution sequence. *Stochastics and Partial Differential Equations: Analysis and Computations*, 4(4):791–826, 2016.

[64] A. Lorz, S. Mirrahimi, and B. Perthame. Dirac Mass Dynamics in Multidimensional Nonlocal Parabolic Equations. *Communications in Partial Differential Equations, Taylor & Francis*, 36(6):1071–1098, 2011.

[65] P. Marrow, R. Law, and C. Cannings. The coevolution of predator-prey interactions: ESSs and red queen dynamics. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 250(1328):133–141, 1992.

[66] S. Méléard and V. C. Tran. Trait Substitution Sequence process and Canonical Equation for age-structured populations. *Journal of Mathematical Biology*, 58(6):881–921, 2009.

[67] S. Méléard and V. C. Tran. Slow and fast scales for superprocess limits of age-structured populations. *Stochastic Processes and their Applications*, 122(1):250–276, 2012.

[68] J.A.J Metz, S.A.H. Geritz, G. Meszena, F.J.A Jacobs, and J.S. van Heerwaarden. Adaptive dynamics: A geometrical study of the consequences of nearly faithful reproduction. *Stochastic and spatial structures of dynamical systems, KNAW Verhandelingen, Afd. Natuurkunde, Eerste reeks*, 45:183–231, North–Holland, Amsterdam, 1996.

[69] J.A.J. Metz, R.M. Nisbet, and S.A.H. Geritz. How should we define fitness for general ecological scenarios ? *Trends in Ecology & Evolution*, 7(6):198–202, 1992.

[70] J. D. Murray. *Mathematical biology. 1: An introduction.* Number 17 in Interdisciplinary applied mathematics. Springer-Verlag GmbH, Berlin Heidelberg, softcover reprint of the hardcover 3rd edition 2002, corrected second printing edition, 2004.

[71] E.G. Nepomuceno, R.H.C. Takahashi, and L.A. Aguirre. Individual-based model (ibm): an alternative framework for epidemiological compartment models. *Brazilian Journal of Biometrics*, 34(1):133–162, 2016.

[72] M. Nowak and K. Sigmund. The evolution of stochastic strategies in the Prisoner's Dilemma. *Acta Applicandae Mathematicae*, 20(3):247–265, 1990.

[73] B. Oborny. Growth Rules in Clonal Plants and Environmental Predictability – A Simulation Study. *The Journal of Ecology*, 82(2):341–351, 1994.

[74] K. Oelschlager. Limit theorems for age-structured populations. *The Annals of Probability*, 18(1):290–318, 1990.

[75] G. C. Papanicolaou, D. Stroock, and S. R. S. Varadhan. Martingale approach to some limit theorems. In P. L. Chow, editor, *Duke turbulence conference*, number III in Duke University Mathematics Series, pages ii+120 pp. Duke University, Durham, NC, 1977.

[76] T. Paul. The canonical equation of adaptive dynamics in individual-based models with power law mutation rates. 2024. Publisher: arXiv Version Number: 2.

[77] G. A. Pavliotis and A. M. Stuart. *Multiscale methods: averaging and homogenization*. Number 53 in Texts in applied mathematics. Springer, New York, NY, 2008.

[78] B. Perthame and G. Barles. Dirac Concentrations in Lotka-Volterra Parabolic PDEs. *Indiana University Mathematics Journal*, 57(7):3275–3301, 2008.

[79] B. Perthame and M. Gauduchon. Survival thresholds and mortality rates in adaptive dynamics: conciliating deterministic and stochastic simulations. *Mathematical Medicine and Biology: A Journal of the IMA*, 27(3):195–210, 2009.

[80] P. Protter. *Stochastic Integration and Differential Equation*. Springer, 2004.

[81] J. Roughgarden. *The theory of coevolution*. Sinauer Associates, Sunderland Massachusetts, 1983.

[82] V. C. Tran. Large population limit and time behaviour of a stochastic particle model describing an age-structured population. *ESAIM: Probability and Statistics*, 12:345–386, 2008.

[83] D. Waxman and S. Gavrilets. 20 Questions on Adaptive Dynamics: Target review of Adaptive Dynamics. *Journal of Evolutionary Biology*, 18(5):1139–1154, 2005.

[84] D. Williams. *Probability with Martingales*. Cambridge University Press, 1 edition, 1991.

[85] S. Wright. Evolution in mendelian populations. *Genetics*, 16(2):97–159, 1931.