Individual-based models for bacterial chemotaxis and variance reduced simulations

Mathias Rousset, Giovanni Samaey

INDIVIDUAL-BASED MODELS FOR BACTERIAL CHEMOTAXIS AND VARIANCE REDUCED SIMULATIONS *

MATHIAS ROUSSET † AND GIOVANNI SAMAEY ‡

Abstract. Velocity-jump models for the individual-based simulation of chemotaxis of bacteria with internal dynamics are discussed. We analyze a fine-scale process with internal dynamics, and a simpler process associated with a kinetic description. We first show that, in a diffusive asymptotics, both processes converge towards the same advection-diffusion process. Subsequently, we couple the jump times of both processes and analyze the variance of the difference between the two solutions. We show that this coupling yields an "asymptotic" variance reduction (control variate), in the sense that, in the diffusive asymptotics, the difference between the two processes has a variance which vanishes with the small parameter. Finally, this coupling is used to construct a "hybrid" scheme with reduced variance, by first computing a deterministic solution of the kinetic density description, and then simulating the coupled processes to evaluate the difference with the exact solution with internal dynamics.

*

†SIMPAF, INRIA Lille - Nord Europe, Lille, France, (mathias.rousset@inria.fr).

‡Department of Computer Science, K.U. Leuven, Celestijnenlaan 200A, 3001 Leuven, Belgium, (giovanni.samaey@cs.kuleuven.be).

1. Introduction

Generally, the motion of flagellated bacteria consists of a sequence of *run* phases, during which a bacterium moves in a straight line at constant speed. The bacterium changes direction in a *tumble* phase, which is much shorter than the run phase and acts as a reorientation. Hence, the motion of an individual bacterium is usually modeled as a velocity-jump process. To bias movement towards regions with high concentration of chemoattractant, the bacterium adjusts its turning rate to increase, resp. decrease, the chance of tumbling when moving in an unfavorable, resp. favorable, direction [2, 29]. Since many species are unable to sense chemoattractant gradients reliably due to their small size, this is often done via an intracellular mechanism that provides a type of memory [3]. This memory effect can be modeled using an ordinary differential equation (ODE), resulting in an individual-based model with internal state.

In this paper, we analyze individual-based simulations for chemotaxis of bacteria that incorporate such internal dynamics, using a slight generalization of a model that was first introduced by Erban and Othmer [11]. The goal of the paper is twofold. First, we analyze the model with internal dynamics, and a simpler, coarse process associated with a kinetic description. We show that, in a diffusive asymptotics, both converge towards the same advection-diffusion limit. We use a rigorous, probabilistic point of view in the asymptotic analysis, without moment closure assumptions. The proposed method also enables to deal with non-linear internal dynamics and non-linear jump rate mechanisms. Second, we use the simpler kinetic process as a control variate to significantly reduce the variance of the individual-based simulations. The variance reduction is achieved by coupling the velocity jump times of the individual-based simulation, and turns out to be "asymptotic" in the sense that the variance of the estimation vanishes in the diffusive asymptotics. The resulting variance reduced method can then be used to simulate the evolution of the bacterial density, by using a deterministic solver to compute the kinetic description, and using the coupled processes to compute the difference with the exact solution.

The idea of "asymptotic variance reduction" is a general and very recent idea in scientific computing that appears when using hybrid Monte-Carlo/PDE (partial differential equations) methods. Up to our knowledge, the only explicit attempt to develop "asymptotic variance reduction" in hybrid methods can be found in [6] and related papers, in the context of the Boltzmann equation. The most broad picture of the idea developed in the present paper is given by the following table:

Model	Numerics
Fine-scale Model–Particle description	Monte-Carlo method
Coarse Model–Particle description	Monte-Carlo method (coupled)
Coarse Model–Continuum description	Grid method

When both a continuum and a particle description are available for a coarse model that approximates a given fine-scale particle description, the idea is to construct a coupling between the coarse and fine-scale particle descriptions to obtain a variance reduced scheme for the fine-scale model. To this end, the fine-scale and coarse particle descriptions have to be simulated with similar schemes using the same random numbers. Evolution of the fine-scale model is then evaluated on the grid at hand by adding to the evolution of the coarse continuum description the difference between the two coupled particle descriptions.

In chemotaxis, the bacterial density is expected to satisfy an advection-diffusion equation on large space and time scales, in which a chemotactic sensitivity coeffi-

cient incorporates the effect of chemoattractant concentrations on the density fluxes. This assumption leads to the classical Keller–Segel equations (see [18], and [16, 17] for numerous historical references). Several works have considered the motion of a bacterium to be governed by a velocity-jump process corresponding to a kinetic description for the phase-space density [1, 23, 26]. These models can be shown to converge to a Keller–Segel equation in the appropriate drift-diffusion limit [5, 15, 24, e.g.]. Moreover, unlike the limiting Keller–Segel equation [4, 14, 25, e.g.], the kinetic description does not exhibit finite-time blow-up (which is believed to be unphysical) under certain biologically relevant assumptions on the turning kernel [5].

As in [11], we will supplement a velocity-jump process with a model for the evolution of the internal state of the bacteria as a function of the chemoattractant concentrations, and specify the turning rates as a function of these internal states. The internal variables then appear as extra dimensions in the kinetic equation. Using moment closure assumptions and an appropriate (diffusion) scaling, the Keller–Segel limit has been derived formally from the model with internal variables; the parameters of internal dynamics appearing in the expression for the chemotactic sensitivity [10, 11]. Due to the possibly high number of dimensions of the kinetic model with internal state, the evolution of the bacterial density away from the advection-diffusion limit can only be simulated using a stochastic particle method.

In this context, the "asymptotic variance reduction" method can be depicted using the following table:

Model	Numerics
Internal state–Individual description	Monte-Carlo method
Kinetic model–Individual description	Monte-Carlo method (coupled)
Kinetic model–Density description	Grid method

The bacterial density is then obtained by comparing the difference between the two coupled Monte-Carlo schemes, using a kernel density estimation [27, 28]. In this paper, we will assume that the computation of the density of bacteria under the kinetic description can be performed accurately with a grid method. Note that in dimension 3 or even 2, this can become cumbersome, and the grid computation should be carried out at the level of advection-diffusion description only, using a second level of coupling, between the velocity-jump process associated with the kinetic description, and the stochastic differential equation associated with the advection-diffusion limit. This is left for future work.

The paper is organized as follows. In Section 2, we discuss the different velocity-jump processes that will be considered. In Section 3, we show that, in the diffusive asymptotics and with an appropriate scaling (1) both the model with internal state and the kinetic model converge to an advection-diffusion equation; (2) the variance of the difference between the two coupled process vanishes with a given order. Section 4 discusses a discretization scheme for the velocity-jump processes that retains these continuum coupling properties. We illustrate the analysis in Section 5 and apply the method to a long time simulation in Section 6. Section 7 contains conclusions and a discussion of possibilities for future work.

2. Particle-based models for bacterial chemotaxis

2.1. The general model

In the models considered in the present paper, bacteria are sensitive to the concentration of m chemoattractants $(S_i(x))_{1\leq i\leq m}$, with $S_i(x)\geq 0$ for $x\in\mathbb{R}^d$. While we do not consider time dependence of $S(x)$ by production or consumption of chemoattractant by the bacteria, a generalization to this situation is straightforward. Bacteria

move with a constant speed v (run), and change direction at random instances in time (tumble), in an attempt to move towards regions with high chemoattractant concentrations. As in [10], we will describe this behavior by a velocity-jump process driven by some internal state $y \in \mathbb{Y} \subset \mathbb{R}^n$ of each individual bacterium, with $n \geq m$. The internal state models the memory of the bacterium and is subject to an evolution mechanism attracted by the chemoattractants concentrations:

$$S(x) := (S_1(x), \ldots, S_m(x), 0, \ldots, 0) \in \mathbb{R}^n,$$

where x is the present position of the bacterium and we have added some additional fixed (null) values to match the dimensions of S and y. The motion of each bacterium is then characterized by three physical scales:

- The typical length l_s of the chemoattractant concentration variations, and the associated time t_x, where l_s/t_x is the typical speed of a bacterium,
- The typical time t_y of the evolution of the internal state,
- The typical time t_λ between two changes of the bacterium velocity direction (tumbling).

Note that several typical time scales ($t_{1,y} \gg t_{2,y} \gg \ldots$) may in fact be required to describe the evolution of the internal state; $t_y := t_{1,y}$ is assumed to be the largest.

After scaling in dimensionless form using (l_s, t_λ) as a reference for length and time, the evolution of each bacterium individual position will be denoted:

$$t \mapsto X_t,$$

with normalized velocity:

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \frac{t_\lambda}{t_x} V_t, \qquad V_t \in \mathbb{V} = \mathbb{S}^{d-1},$$

with $\mathbb{S}^{d-1}$ the unit sphere in $\mathbb{R}^d$. The evolution of the internal state is denoted by $t \mapsto Y_t$. The internal state adapts to the local chemoattractant concentration through a known and numerically computable ordinary differential equation (ODE):

$$\frac{\mathrm{d}Y_t}{\mathrm{d}t} = \frac{t_\lambda}{t_y} F(Y_t, S(X_t)), \tag{2.1}$$

which is required to have a unique fixed point given by the chemoattractant concentrations with additional null values $S(x)$; more precisely, the equation $F(y, S(x)) = 0$ has a unique solution $y^* = S(x^*)$ for every fixed value $x^* \in \mathbb{R}^d$. In the above, the function F is normalized so that the largest time scale of the convergence of the ODE $\dfrac{\mathrm{d}Y_s}{\mathrm{d}s} = F(Y_t, S(x))$ to its stationary state is of order $\mathcal{O}(1)$ for any x. We also introduce the deviations from equilibrium $z = S(x) - y$. The evolution of these deviations is denoted as

$$t \mapsto Z_t = S(X_t) - Y_t,$$

and the associated evolution equation is given by

$$\frac{\mathrm{d}Z_t}{\mathrm{d}t} = \frac{t_\lambda}{t_x} \nabla S(X_t) V_t - \frac{t_\lambda}{t_y} F(Y_t, S(X_t)), \tag{2.2}$$

with, by convention, the gradient $\nabla S \in \mathbb{R}^{n\times d}$ for $S:\mathbb{R}^d \to \mathbb{R}^n$. The velocity of each bacterium is switched at random jump times $(T_n)_{n\geq 1}$ that are generated via a Poisson process with a time dependent rate given by $\lambda(Z_t)$, where $z \mapsto \lambda(z)$ is a smooth function satisfying

$$0 < \lambda_{\min} \leq \lambda(z) \leq \lambda_{\max}. \tag{2.3}$$

The new velocity $\mathcal{V}_n$ at time T_n is generated at random according to a centered probability distribution $\mathcal{M}(dv)$ with $\int v\mathcal{M}(dv) = 0$, typically

$$\mathcal{M}(dv) = \sigma_{\mathbb{S}^{d-1}}(dv),$$

where $\sigma_{\mathbb{S}^{d-1}}$ is the uniform distribution on the unit sphere.

The resulting fine-scale stochastic evolution of a bacterium is then described by a left continuous with right limits (lcrl) process

$$t \mapsto (X_t, V_t, Y_t),$$

which satisfies the following differential velocity-jump equation:

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = \dfrac{t_\lambda}{t_x} V_t \\ \dfrac{\mathrm{d}Y_t}{\mathrm{d}t} = \dfrac{t_\lambda}{t_y} F(Y_t, S(X_t)) \\ \int_{T_n}^{T_{n+1}} \lambda(Z_t)\mathrm{d}t = \theta_{n+1}, \qquad \text{with } Z_t := S(X_t) - Y_t \\ V_t = \mathcal{V}_n \quad \text{ for } t \in [T_n, T_{n+1}] \ , \end{cases} \tag{2.4}$$

with initial condition $X_0, V_0 \in \mathbb{R}^d$, $Y_0 \in \mathbb{R}^n$ and $T_0 = 0$. In (2.14), $(\theta_n)_{n\geq 1}$ denote i.i.d. random variables, with normalized exponential distribution, and $(V_n)_{n\geq 1}$ denote i.i.d. random variables with distribution $\mathcal{M}(dv)$.

REMARK 2.1. *In one spatial dimension, assuming that* $\mathcal{M}(dv) = \frac{1}{2}(\delta_{+1} + \delta_{-1})$, *the system* (2.14) *is equivalent to the following:*

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = \dfrac{t_\lambda}{t_x} V_t \\ \dfrac{\mathrm{d}Y_t}{\mathrm{d}t} = \dfrac{t_\lambda}{t_y} F(X_t, S(Y_t)) \\ \displaystyle\int_{T_n}^{T_{n+1}} 2\lambda(Z_t)\mathrm{d}t = \theta_{n+1}, \\ V_t = \mathcal{V}_{n+1} \quad \textit{ for } t \in [T_n, T_{n+1}] \textit{ with } \mathcal{V}_{n+1} = -\mathcal{V}_n, \end{cases} \tag{2.5}$$

in the sense that both processes have the same probability path distribution. This can be checked on the Markov infinitesimal generator of probability transitions, see (2.18) *and discussion below.*

2.2. Example

For concreteness, we provide a specific example, adapted from [10], which will also be used later on to illustrate our results numerically. We consider $m = 1$, i.e. there is only one chemoattractant $S_1(x)$. We first describe a cartoon dynamics exhibiting an excitation-adaptation behaviour. The internal state is two-dimensional, i.e. $n = 2$,

and $y=(y_1,y_2)$ satisfies the following ODE:

$$\begin{cases} \dfrac{\mathrm{d}y_1(t)}{\mathrm{d}t} = \dfrac{S_1(x)-y_1(t)}{t_a}, \\ \dfrac{\mathrm{d}y_2(t)}{\mathrm{d}t} = \dfrac{-S_1(x)+y_1(t)-y_2(t)}{t_e}, \end{cases} \tag{2.6}$$

in which t_a, resp. t_e, represent an adaptation, resp. excitation, time; and y_1, resp. y_2, represent the adapting, resp. exciting, variable. This model has a single fixed point $y^*=(y_1^*,y_2^*)=(S_1(x),0)=S(x)$, and the deviation variables are given by $z=(z_1,z_2)=(S_1(x)-y_1,-y_2)$. The variable y_1 therefore adapts "slowly" to the environment and memorizes it, while y_2 computes "faster" the lag $y_1-S_1(x)$, giving the response to changing environments. When the excitation time is much smaller than the adaptation time, $t_e \ll t_a$, internal dynamics reduce to a scalar equation

$$\frac{\mathrm{d}y_1(t)}{\mathrm{d}t} = \frac{S_1(x)-y_1}{t_a}, \tag{2.7}$$

while the exciting variable reduces to the difference $y_2=y_1-S_1(x)$ which is instantly learned by the bacteria.

For the turning rate $z\mapsto\lambda(z)$, we choose the following nonlinear strictly decreasing smooth function which depends on the scalar $\zeta=b^Tz$, for instance

$$\lambda(z)\equiv\lambda(\zeta)=2\lambda_0\left(\frac{1}{2}-\frac{1}{\pi}\arctan\left(\frac{\pi}{2\lambda_0}\zeta\right)\right). \tag{2.8}$$

For (2.6), we may choose

$$\begin{cases} z=(S_1(x),0)-y\in\mathbb{R}^2, \\ b=(0,\beta), \qquad \beta>0, \end{cases}$$

in which $\zeta=-\beta y_2>0$, resp. $\zeta=-\beta y_2<0$, give the information to the bacterium that it is moving in a favorable, resp. unfavorable, direction. A linear approximation of the turning rate is then given by

$$\lambda(z)\equiv\lambda_0-b^Tz. \tag{2.9}$$

When internal dynamics is governed by (2.7), the turning rate becomes a function of $\zeta=z_1=S(x)-y_1$.

2.3. Asymptotic regimes and kinetic formulation In the present work, the typical time t_x associated with chemoattractant variations as seen from the bacteria will be assumed to be large compared to the typical time of random change of velocity direction t_λ:

$$\epsilon:=\frac{t_\lambda}{t_x}\ll 1. \tag{2.10}$$

We may then assume that the ODE (2.2) driving the internal state is well approximated by a near equilibrium evolution equation in the sense that:

$$\bar{F}(y,S(x)):=\frac{t_\lambda}{t_y}F(Y_t,S(X_t))=-\tau^{-1}(y-S(x))+\mathcal{O}_\tau(|S(x)-y|^2), \tag{2.11}$$

where $\tau \in \mathbb{R}^{n\times n}$ is an invertible constant matrix with strictly positive real part of the spectrum.

REMARK 2.2. *Throughout the text, the Landau symbol $\mathcal{O}$ denotes a deterministic function that may depend on all the parameters of the model, but that is globally Lipshitz* uniformly *with respect to the variables/parameters of the model except λ_0, b, S and its derivatives. When F is non-linear, Landau symbols may also depend on τ, and we will use the notation $\mathcal{O}_\tau$. The same convention holds for constants denoted K or K_τ whose precise value may vary from line to line.* This yields the estimate:

$$\frac{\mathrm{d}Z_t}{\mathrm{d}t} = -\tau^{-1} Z_t + \epsilon \nabla S(X_t) V_t + \mathcal{O}_\tau(|Z_t|^2). \tag{2.12}$$

We will denote by $r_K(\tau^{-1}) > 0$ the optimal rate such that:

$$\forall \alpha > 0, \quad \left\| \mathrm{e}^{-\alpha \tau^{-1}} \right\| \leq K \mathrm{e}^{-\alpha r_K(\tau^{-1})}.$$

In what follows, we choose an arbitrary fixed constant K and denote simply $r \equiv r_K$ the optimal rate associated with K. The largest timescale of the internal state evolution can thus be defined by $t_y = \dfrac{t_\lambda}{r(\tau^{-1})}$. The turning rate is assumed to satisfy

$$\lambda(z) = \lambda_0 - b^T z + \mathcal{O}\left(|z|^2\right), \tag{2.13}$$

where $b \in \mathbb{R}^n$ is assumed to be independent of x. The dynamics of bacteria is then described by the following differential velocity-jump equation:

$$\begin{cases} \frac{\mathrm{d}X_t}{\mathrm{d}t} = \epsilon V_t \\ \frac{\mathrm{d}Y_t}{\mathrm{d}t} = \bar{F}(Y_t, S(X_t)) \\ \int_{T_n}^{T_{n+1}} \lambda(Z_t)\mathrm{d}t = \theta_{n+1}, \qquad \text{with } Z_t := S(X_t) - Y_t \\ V_t = \mathcal{V}_n \quad \text{ for } t \in [T_n, T_{n+1}] \ , \end{cases} \tag{2.14}$$

where $(\lambda, \bar{F})$ verify (2.13)-(2.11).

We then assume that the solution of the ODE driving the internal state in (2.14) satisfies the following long time behavior:

ASSUMPTION 2.3. *The solution of the equation* (2.14) *verify:*

$$\sup_{t\in[0,+\infty]} |Z_t| = \mathcal{O}_\tau\left(|Z_0| + \epsilon\right),$$

where the Landau symbol $\mathcal{O}$ is uniform in (x, y, v, ϵ). The Assumption 2.3 is motivated by the linear case, where such a behavior always holds. We will also consider the following linearity assumption:

ASSUMPTION 2.4. *The evolution equation of the internal state and the velocity jump rate are linear:*

$$\begin{cases} \frac{\mathrm{d}Z_t}{\mathrm{d}t} & = -\tau^{-1} Z_t + \epsilon \nabla S(X_t) V_t, \\ \lambda(z) & = \lambda_0 - b^T z. \end{cases}$$

Note that under Assumption 2.4, Assumption 2.3 is verified by applying Duhamel's integration formula. Indeed, the latter yields:

$$Z_t = \mathrm{e}^{-t\tau^{-1}} Z_0 + \epsilon \int_0^t \mathrm{e}^{-(t-t')\tau^{-1}} \nabla S(X_{t'}) V_{t'} \mathrm{d}t', \tag{2.15}$$

which gives:

$$\sup_{t\in[0,+\infty]}|Z_t| = \mathcal{O}\left(|Z_0| + \frac{\epsilon}{r(\tau^{-1})}\right), \tag{2.16}$$

where the Landau symbol $\mathcal{O}$ is indeed uniform in τ. This ensures that for small enough Z_0 and ϵ, the turning rate $\lambda(Z_t) \geq \lambda_{\min}$ for all $t \geq 0$.

In the non-linear case under Assumption 2.3, we will assume that τ is fixed. In the linear case given by Assumption 2.4, $\tau \equiv \tau_\epsilon$ may also depend on ϵ, and the following asymptotic regime will be considered

$$\epsilon \ll r(\tau_\epsilon^{-1}) \leq K, \tag{2.17}$$

which amounts to considering fast evolutions of the internal state compared to chemoattractant variations as seen from the bacteria $t_y \ll t_x$.

The probability distribution density of the process at time t with respect to the measure $\mathrm{d}x\,\mathcal{M}(\mathrm{d}v)\,\mathrm{d}y$ is denoted as $p(x,v,y,t)$, suppressing the dependence on ϵ for notational convenience, and evolves according to the Kolomogorov forward evolution equation (or master equation). In the present context, the latter is the following kinetic equation

$$\partial_t p + \epsilon v\cdot\nabla_x p + \mathrm{div}_y\left(\bar{F}(x,y)p\right) = \lambda(S(x)-y)\,(R(p)-p)\,, \tag{2.18}$$

where

$$R(p) := \int_{v\in\mathbb{S}^{d-1}} p(\cdot,v,\cdot)\,\mathcal{M}(dv)$$

is the operator integrating velocities wrt $\mathcal{M}$. The reader is referred to [12, Chapter 4] for the derivation of master equations associated to Markov jump processes.

2.4. Model with direct gradient sensing (control process)

The process with internal state (2.14) will be coupled[1] with a simplified velocity-jump process without internal state. The coupled process without internal state will be called the *control process* since it will be used in Section 3.3 to perform variance reduced simulations of (2.14).

The control process is a Markov process in position-velocity variables

$$t \mapsto (X_t^c, V_t^c),$$

which evolve according to the following differential velocity-jump equations:

$$\begin{cases} \dfrac{\mathrm{d}X_t^c}{\mathrm{d}t} = \epsilon V_t^c \\ \displaystyle\int_{T_n^c}^{T_{n+1}^c} \lambda^c(X_t^c, V_t^c)\mathrm{d}t = \theta_{n+1} \\ V_t^c = V_n \quad \text{for } t\in[T_n^c, T_{n+1}^c]\ , \end{cases} \tag{2.19}$$

with initial condition $X_0, V_0 \in \mathbb{R}^d$. In (2.19), $(\theta_n)_{n\geq 1}$ denote i.i.d. random variables, with normalized exponential distribution, and $(V_n)_{n\geq 1}$ denote i.i.d. random variables

[1] A coupling between two random processes, or more generally between two random variables, is a way of constructing them on the same probability space. Practically speaking, it amounts to choose a particular correlation between the two sets of random numbers generating them.

with distribution $\mathcal{M}(dv)$. The *coupling* between (2.19) and (2.14) precisely comes from the fact that the set of generating random variables $(\theta_n)_{n\geq 1}$ and $(V_n)_{n\geq 1}$ are the same. The turning rate of the control process is assumed to satisfy

$$\lambda^c(x,v) := \lambda_0 - \epsilon\, A^T(x)v + \mathcal{O}(\epsilon^2), \tag{2.20}$$

for some vector field $A(x)\in\mathbb{R}^d$.

The analysis in Section 3 will show that the appropriate choice for $A(x)$ is given by

$$A(x) = b^T \frac{\tau}{\lambda_0\tau + \mathrm{Id}} \nabla S(x), \tag{2.21}$$

where b, τ, and λ_0 were introduced in (2.11)-(2.13) as parameters of the process with internal state, and $\mathrm{Id}\in\mathbb{R}^{n\times n}$ is the identity matrix.

The model (2.20) may describe a large bacteria that are able to directly sense chemoattractant gradients. In the case $m=n=1$, the turning rate (2.20) is proportional to $\nabla S(x)v\in\mathbb{R}$, which can be interpreted as follows: the rate at which a bacterium will change its velocity direction depends on the alignment of the velocity with the gradient of the chemoattractant concentration $\nabla S(x)$, resulting in a transport towards areas with higher chemoattractant concentrations. The model can also be derived from the internal state model (2.14) in the special asymptotics where bacteria learn the value of the chemoattractants concentrations infinitely fast. We have the following lemma.

LEMMA 2.5. *Consider a sequence $(\tau_k,b_k)_{k\geq 0}$ such that $\|\tau_k\| \underset{k\to\infty}{\to} 0$ with $b_k^T\tau_k \to \bar{b}^T$ for some vector $\bar{b}\in\mathbb{R}^n$, and*

$$\sup_k \frac{\left\|\tau_k^{-1}\right\|}{r(\tau_k^{-1})} < +\infty. \tag{2.22}$$

Consider the process with internal state $t\mapsto(X_t^k,V_t^k,Y_t^k)$ solution of (2.14)*, and assume (2.11)-(2.13). The process converges almost surely, with respect to the Skorohod metric, towards the gradient sensing process solution of* (2.19) *with velocity-jump rate:*

$$\lambda_0 - \epsilon\bar{b}^T\nabla S(x)v.$$

A proof is given in Appendix A. It relies on the definition of the Skorohod metric and associated convergence and compacity criteria, see [12, 22]. In the present context of finite number of jumps, convergence with respect to the Skorohod topology is equivalent to convergence of the jump times and sizes, and uniform convergence away from jump times.

The distribution density of the control process at time t with respect to the measure $\mathrm{d}x\,\mathcal{M}(\mathrm{d}v)$ is denoted as $p^c(x,v,t)$, and evolves according to the kinetic/master equation:

$$\partial_t p^c + \epsilon v\cdot\nabla_x p^c = \left(R(\lambda^c p) - \lambda^c p\right), \tag{2.23}$$

where R is the operator integrating velocities wrt $\mathcal{M}$. Note that we have again suppressed the dependence of $p^c(x,v,t)$ on ϵ. (2.23) should be compared to (2.18).

2.5. Coupling, simulation and variance reduction Introducing an internal state mechanism increases the dimensionality of the system. It can then become cumbersome to simulate directly the kinetic equation (2.18) for the process with internal state (2.14) over diffusive time scales:

$$t = \bar{t}/\epsilon^2,$$

for $\bar{t} > 0$. Let us introduce the position bacterial density of the process with internal state (2.14)-(2.18)

$$n^\epsilon(x,\bar{t}) := n(x,\bar{t}) := \int_{\mathbb{V}} p(x,v,\bar{t}/\epsilon^2)\,\mathcal{M}(\mathrm{d}v), \tag{2.24}$$

and similarly the position bacterial density of the control process (2.19)-(2.23)

$$n^{c,\epsilon}(x,\bar{t}) := n^c(x,\bar{t}) := \int_{\mathbb{V}} p^c(x,v,\bar{t}/\epsilon^2)\,\mathcal{M}(\mathrm{d}v). \tag{2.25}$$

For given values of ϵ, we will therefore use an ensemble of N particles $\{X_t^i\}_{i=1}^N$, resp. $\{X_t^{i,c}\}_{i=1}^N$, that evolve according to (2.14), resp. (2.19). The density can then be computed via standard kernel density estimation [27, 28],

$$\hat{n}_N(x,\bar{t}) = \frac{1}{Nh}\sum_{i=1}^{N} K_h(x - X^i_{\bar{t}/\epsilon^2}), \tag{2.26}$$

and similarly for $\hat{n}_N^c$. Here K_h represents a suitable (e.g. Gaussian) kernel function; the bandwidth h can be chosen based on the data, for instance using the Silverman heuristic [28].

We will reduce the variance of this simulation by exploiting the coupling between the two ensembles. To this end, we will simulate (2.14) and (2.19) *using the same random numbers* $(\theta_n)_{n\geq 1}$ and $(V_n)_{n\geq 0}$, to obtain a strong correlation between X_t^i and $X_t^{i,c}$ for each realization. The important point is that a deterministic solution $n^c(x,t)$ for the control process may be obtained with deterministic grid-based methods, either by direct simulation of the kinetic equation or by obtaining a moment system that is accurate up to a desired order $\mathcal{O}(\epsilon^k)$. An improved estimate $\bar{n}_N(x,\bar{t})$ can then be obtained by solving the deterministic equation for $n^c(x,t)$ using a grid-based method, and to add the difference between the two coupled particle simulations,

$$\bar{n}_N(x,\bar{t}) = K_h * n^c(x,\bar{t}) + (\hat{n}_N(x,\bar{t}) - \hat{n}_N^c(x,\bar{t})), \tag{2.27}$$

where $*$ denotes the usual convolution operator. In (2.27), $\hat{n}_N(x,\bar{t})$ and $\hat{n}_N^c(x,\bar{t})$ are computed by simulating the two coupled processes, while $n^c(x,\bar{t})$ is computed by a seperate deterministic solver. By construction, the expectation $\mathbb{E}(\bar{n}_N(x,\bar{t})) = K_h * n(x,\bar{t}) + \mathcal{O}(\epsilon^k)$, while fluctuations are much reduced if $\hat{n}_N(x,\bar{t})$ and $\hat{n}_N^c(x,\bar{t})$ are strongly correlated. In Section 3.3, the variance of $\hat{n}_N(x,\bar{t}) - \hat{n}_N^c(x,\bar{t})$ will be estimated.

In Section 5, we will illustrate this result numerically. The simulations will reveal that the coupling degenerates rapidly as a function of diffusive time $\bar{t}$. We will therefore restore the coupling at fixed time instances $\bar{t}_n = n\,\delta\bar{t}_{pde}$, where $\delta\bar{t}_{pde}$ is the time step used to compute $n^c(x,\bar{t})$ with a deterministic solver. This is done as follows: Assume the simulation of the coupled processes have been performed on $[\bar{t}_{n-1},\bar{t}_n]$. At

$\bar{t}=\bar{t}_n$, we stop the simulation and compute the variance-reduced density estimator $\bar{n}_N(x,\bar{t}_n)$ using (2.27). We then reinitialize the control particles

$$X^{i,c}_{\bar{t}_n/\epsilon^2}=X^i_{\bar{t}_n/\epsilon^2}, \qquad V^{i,c}_{\bar{t}_n/\epsilon^2}=V^i_{\bar{t}_n/\epsilon^2}, \qquad i=1,\dots,N,$$

and reinitialize the current deterministic solution to be

$$n^c(x,\bar{t}_n):=\bar{n}_N(x,\bar{t}_n).$$

The simulation is then continued.

3. Advection-diffusion limits

3.1. Limit of the kinetic control process In this section, classical probabilistic arguments will be used to derive the pathwise advection-diffusion limit of the control process (2.19) when $\epsilon\to 0$. It should be noted that this diffusive limit as already been extensively studied in the context of bacterial chemotaxis, see e.g. [15] for justifications of Hilbert expansions at the PDE level, and [30] where the diffusive limit of some more general class of stochastic processes is suggested. However, we give the probabilistic arguments *in extenso* since they do not seem to appear as such in the literature, and moreover will be of interest for the analysis of both the model with internal state, and the coupling.

Diffusive times are denoted by

$$\bar{t}:=t\epsilon^2,$$

and the control process considered on diffusive time scales:

$$X^{c,\epsilon}_{\bar{t}}:=X^c_{\bar{t}/\epsilon^2}.$$

The latter converges for $\epsilon\to 0$ towards an advection-diffusion process, satisfying the stochastic differential equation (SDE):

$$\mathrm{d}X^{c,0}_{\bar{t}}=\left(\frac{DA(X^{c,0}_{\bar{t}})}{\lambda_0}\mathrm{d}\bar{t}+\left(\frac{2D}{\lambda_0}\right)^{1/2}\mathrm{d}W_{\bar{t}}\right), \tag{3.1}$$

where $\bar{t}\mapsto W_{\bar{t}}$ is a standard Brownian motion, and the diffusion matrix is given by the covariance of the Maxwellian distribution:

$$D=\int_{\mathbb{S}^{d-1}} v\otimes v\,\mathcal{M}(dv)\in\mathbb{R}^{d\times d}.$$

In particular, this result implies at the level of the Kolomogorov/master evolution equation, that the position bacterial density (2.25) satisfies the advection-diffusion equation

$$\partial_{\bar{t}}n^{c,0}=\frac{1}{\lambda_0}\mathrm{div}_x\left(D\nabla_x n^{c,0}-DA(x)n^{c,0}\right) \tag{3.2}$$

on diffusive time scales as $\epsilon\to 0$.

The proof relies on the perturbation analysis of the jump times difference in the following lemma.

LEMMA 3.1. *The difference between two jump times of the control process satisfies*

$$\Delta T^c_{n+1}:=T^c_{n+1}-T^c_n=\frac{\theta_{n+1}}{\lambda_0}+\epsilon\frac{\theta_{n+1}}{\lambda_0^2}A^T(X^c_{T^c_n})\mathcal{V}_n+(1+\theta_{n+1})\mathcal{O}(\epsilon^2), \tag{3.3}$$

where the Landau symbol $\mathcal{O}$ is uniform with respect to τ when A is given by (2.21).

Proof. Use (2.20) to get,

$$\begin{aligned}\lambda^c(X_t^c,V_t^c) &= \lambda_0 - \epsilon A^T(X_t^c)\mathcal{V}_n\,ds + \mathcal{O}(\epsilon^2)\\ &= \lambda_0 - \epsilon A^T(X_{T_n^c}^c)\mathcal{V}_n + (\|\nabla A\|_\infty + 1)\mathcal{O}(\epsilon^2).\end{aligned}$$

Integrating on time interval $[T_n^c, T_{n+1}^c]$ yields:

$$\begin{aligned}\Delta T_{n+1}^c &= \frac{\theta_{n+1}}{\lambda_0 - \epsilon A^T(X_{T_n^c}^c)\mathcal{V}_n} + \mathcal{O}(\epsilon^2)\\ &= \frac{\theta_{n+1}}{\lambda_0}\left(1 + \epsilon\frac{1}{\lambda_0}A^T(X_{T_n^c}^c)\mathcal{V}_n\right) + (1+\theta_{n+1})\mathcal{O}(\epsilon^2).\end{aligned}$$

Remark that when A is given by (2.21):

$$\sup_\tau\left(\|\nabla A\|_\infty + \|A\|_\infty\right) \le \frac{|b|}{\lambda_0}\left(\|\mathrm{Hess}(S)\|_\infty + \|\nabla S\|_\infty\right).$$

□ Then the following result holds:

PROPOSITION 3.2. *The process $\bar{t} \mapsto X_{\bar{t}}^{c,\epsilon} = X_{t/\epsilon^2}^c$ converges in distribution (with respect to the topology of uniform convergence) towards $\bar{t} \mapsto X_{\bar{t}}^{c,0}$, solution of the SDE (3.1). As a consequence, the density $n^{c,\epsilon}$ of the control process defined in* (2.25) *converges weakly towards $n^{c,0}$ (3.2).*

Proof. Let us introduce the process with random time change:

$$\tilde{X}_{\bar{t}}^{c,\epsilon} = X_0^c + \sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor - 1} \epsilon \Delta T_{n+1}^c \mathcal{V}_n. \tag{3.4}$$

We can compute the limit of the time changed process $\bar{t} \mapsto \tilde{X}_{\bar{t}}^{c,\epsilon}$. Applying standard results of weak convergence of Euler discretization of SDEs (see [21]) shows that $\bar{t} \to X_{\bar{t}}^{c,\epsilon}$ converges in distribution for the uniform convergence topology to the solution $\bar{t} \mapsto X_{\bar{t}}^{c,0}$ of an SDE. The diffusion coefficient of the latter SDE can be computed in (3.4) using the expansion (3.3) and the independence between the random variables $\mathcal{V}_n$, θ_{n+1}, and the past trajectory $(X_t^c)_{t \le T_n^c}$. This yields:

$$\epsilon^2 \mathbb{E}\left(\mathcal{V}_n \otimes \mathcal{V}_n \left(\Delta T_{n+1}^c\right)^2 | X_{T_n^c}^c\right) = \frac{\epsilon^2}{\lambda_0^2} D\,\mathbb{E}(\theta_{n+1}^2) + \mathcal{O}(\epsilon^3) = \frac{2\epsilon^2}{\lambda_0^2} D + \mathcal{O}(\epsilon^3);$$

as well as the drift coefficient:

$$\epsilon \mathbb{E}\left(\mathcal{V}_n \Delta T_{n+1}^c | X_{T_n^c}^c\right) = \epsilon^2 \frac{1}{\lambda_0^2} DA(X_{T_n^c}^c) + \mathcal{O}(\epsilon^3);$$

which yields in the end the SDE (3.1). To obtain (3.1), remark that the number of jumps is multiplied by λ_0 in the sum defining $\bar{t} \mapsto \tilde{X}_{\bar{t}}^{c,\epsilon}$, which amounts to multiply by λ_0 the drift and the diffusion coefficients.

It remains to show that $\bar{t} \to \tilde{X}_{\bar{t}}^{c,\epsilon}$ and $\bar{t} \to X_{\bar{t}}^{c,\epsilon}$ converge to the same limit. For this purpose, let us introduce the random time change

$$\alpha_{c,\epsilon}(\bar{t}) = \epsilon^2 T_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor}^c + \epsilon^2 \left(T_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor + 1}^c - T_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor}^c\right)\left(\lambda_0 \bar{t}/\epsilon^2 - \lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor\right) \tag{3.5}$$

$$= \sum_{n=1}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \epsilon^2 \Delta T_n^c + \epsilon^2 \left(\Delta T_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor + 1}^c\right)\left(\lambda_0 \bar{t}/\epsilon^2 - \lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor\right), \tag{3.6}$$

which is the linear interpolation of the jump times $(T_n^c)_{n\geq 0}$, and by construction is almost surely strictly increasing and continuous. Using (3.3), we get:

$$\alpha_{c,\epsilon}(\bar{t}) = \frac{\epsilon^2}{\lambda_0} \sum_{n=1}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \theta_n + \frac{\mathcal{O}(\epsilon^3)}{\lambda_0} \sum_{n=1}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} (\theta_n + 1),$$

so that $\bar{t} \mapsto \alpha_{c,\epsilon}(\bar{t})$ converges in distribution on any time interval $[0,T]$ for the uniform convergence topology towards the deterministic function $\bar{t} \mapsto \bar{t}$ (see e.g. [22]). The same hold true for the inverse time change $\bar{t} \mapsto \alpha_{c,\epsilon}^{-1}(\bar{t})$ by using for instance the inequality:

$$\begin{aligned} \sup_{\bar{t}\in[0,T]} \left|\alpha_{c,\epsilon}^{-1}(\bar{t}) - \bar{t}\right| \wedge 1 &= \sup_{\bar{t}\in[0,\alpha_{c,\epsilon}^{-1}(T)]} |\alpha_{c,\epsilon}(\bar{t}) - \bar{t}| \wedge 1 \\ &\leq \sup_{\bar{t}\in[0,2T]} |\alpha_{c,\epsilon}(\bar{t}) - \bar{t}| \wedge 1 + 1_{T\geq \alpha_{c,\epsilon}(2T)}. \end{aligned}$$

Applying the random time change to the process yields by construction:

$$\begin{aligned} X^{c,\epsilon}_{\alpha_{c,\epsilon}(\bar{t})} &= \tilde{X}^{c,\epsilon}_{\bar{t}} + \epsilon \mathcal{V}_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \left(\frac{\alpha_{c,\epsilon}(\bar{t})}{\epsilon^2} - T^c_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \right) \\ &= \tilde{X}^{c,\epsilon}_{\bar{t}} + \epsilon \mathcal{V}_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \left(\Delta T^c_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor + 1} \right) \left(\lambda_0 \bar{t}/\epsilon^2 - \lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor \right) \\ &= \tilde{X}^{c,\epsilon}_{\bar{t}} + \theta_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor + 1} \mathcal{O}(\epsilon). \end{aligned} \tag{3.7}$$

Since $\bar{t} \mapsto \alpha_{c,\epsilon}^{-1}(\bar{t})$ converges to a deterministic limit, the coupled process $\bar{t} \mapsto (X^{c,\epsilon}_{\alpha_{c,\epsilon}^{-1}(\bar{t})}, \alpha_{c,\epsilon}^{-1}(\bar{t}))$ also converges to $\bar{t} \mapsto (X^{c,0}_{\bar{t}}, \bar{t})$, and we can consider a Skorokhod embedding[2] associated with this convergence. Then we can write:

$$\begin{aligned} &\sup_{\bar{t}\in[0,T]} \left| X^{c,\epsilon}_{\bar{t}} - X^{c,0}_{\bar{t}} \right| \wedge 1 \\ &\quad \leq \sup_{\bar{t}\in[0,2T]} \left| X^{c,\epsilon}_{\alpha_{c,\epsilon}^{-1}(\bar{t})} - X^{c,0}_{\alpha_{c,\epsilon}^{-1}(\bar{t})} \right| \wedge 1 + \sup_{\bar{t}\in[0,2T]} \left| X^{c,0}_{\bar{t}} - X^{c,0}_{\alpha_{c,\epsilon}^{-1}(\bar{t})} \right| \wedge 1 + 1_{T\geq \alpha_{c,\epsilon}(2T)}, \end{aligned}$$

and the right hand side converges almost surely using the (uniform) continuity of $\bar{t} \mapsto X^{c,0}_{\bar{t}}$.

□

REMARK 3.3. *In one spatial dimension, one can describe the evolution of the distribution density $p^c(x,v,t)$ of the control process as follows [10]. We introduce the distributions of left-moving and right-moving particles $p^c_\pm(x,t) = p^c(x, v = \pm 1, t)$, and write, using* (2.18)*:*

$$\begin{cases} \partial_t p^c_+ + \epsilon \partial_x p^c_+ = -\dfrac{\lambda^c(x,+1)}{2} p^c_+ + \dfrac{\lambda^c(x,-1)}{2} p^c_- \\ \partial_t p^c_- - \epsilon \partial_x p^c_- = \dfrac{\lambda^c(x,+1)}{2} p^c_+ - \dfrac{\lambda^c(x,-1)}{2} p^c_- \end{cases}. \tag{3.8}$$

[2] A Skorokhod embedding associated with an extracted sequence $(\epsilon_n)_{n\geq 0}$ is a new probability space coupling the whole sequence of random processes; for each ϵ_n the distribution law of the process has to be respected on the one hand, and the sequence has to converge almost surely on the other hand (see [22, Chapter 3]).

If we assume $\lambda^c(x,v)$ to be a linear function,

$$\lambda^c(x,v) = \lambda_0 - \epsilon A(x)v, \tag{3.9}$$

is easy to see that equation (3.8) *is equivalent to*

$$\partial_t^2 n^c + \lambda_0 \partial_t n^c = \partial_x \left(\epsilon^2 \partial_x n^c - \epsilon^2 A(x) n^c \right), \tag{3.10}$$

which, on diffusive time-scales $\bar{t} = t\epsilon^2$, is equivalent to (3.2) *in the limit when $\epsilon \to 0$. Note, however, that the assumption* (3.9) *requires $|A(x)| \leq \lambda_0/\epsilon$ to retain positivity of the turning rate.*

3.2. Limit of the process with internal state

In the same way, a standard probabilistic diffusion approximation argument can be used to derive the pathwise diffusive limit of the process (2.14). We denote by

$$X_t^\epsilon = X_{t/\epsilon^2},$$

the process with internal state on diffusive time scales. When $A(x)$ is chosen according to (2.21), this process converges towards a solution of the same advection-diffusion SDE (3.1) as the control process when $\epsilon \to 0$. If we introduce the bacterial density as

$$n(x,t) = \int_{\mathbb{Y}} \int_{\mathbb{V}} p(x,v,y,t) \mathcal{M}(\mathrm{d}v) \mathrm{d}y, \tag{3.11}$$

this implies that n satisfies (3.2) on diffusive time scales as $\epsilon \to 0$. Moreover, in the linear Asumption 2.4, the result still holds for the relaxed regime:

$$\epsilon \ll r(\tau_\epsilon^{-1}) \leq K.$$

The proof relies on the following lemma:

LEMMA 3.4. *Suppose either Assumption 2.3 or 2.4 are satisfied, and let $|Z_0| = \mathcal{O}(\epsilon)$. Then the difference between two jump times*

$$\Delta T_{n+1} := T_{n+1} - T_n$$

satisfies:

$$\begin{aligned} \theta_{n+1} = {} & \lambda_0 \Delta T_{n+1} - b^T \left(Id - \mathrm{e}^{-\Delta T_{n+1} \tau^{-1}} \right) \tau Z_{T_n} \\ & - \epsilon b^T \left(\Delta T_{n+1} \tau - (\mathrm{Id} - \mathrm{e}^{-\Delta T_{n+1} \tau^{-1}}) \tau^2 \right) \nabla S(X_n) \mathcal{V}_n + \theta_{n+1}^2 \mathcal{O}_\tau \left(\epsilon^2 \right). \end{aligned} \tag{3.12}$$

Under Assumption 2.4, the Landau symbol $\mathcal{O}_\tau$ can be taken uniform in τ, $\mathcal{O}_\tau \equiv \mathcal{O}$.

Proof. Let us first consider the linear case (Assumption 2.4).

Duhamel's integration of (2.12) on $[T_n, T_{n+1}]$ with Assumption 2.4 yields

$$Z_t = \mathrm{e}^{-(t-T_n)\tau^{-1}} Z_{T_n} + \epsilon \int_{T_n}^t \mathrm{e}^{-(t-t')\tau^{-1}} \nabla S(X_{t'}) V_{t'} \mathrm{d}t' \tag{3.13}$$

Since S is smooth, $X_{t'}$ can be expanded near X_{T_n} up to an error of order $(t' - T_n) \times \epsilon$. It yields:

$$Z_t = \mathrm{e}^{-(t-T_n)\tau^{-1}} Z_{T_n} + \epsilon \left(Id - \mathrm{e}^{-(t-T_n)\tau^{-1}} \right) \tau \nabla S(X_{T_n}) V_{T_n} + \theta_{n+1} \mathcal{O}\left(\epsilon^2\right); \tag{3.14}$$

and integrating again, it gives:

$$\int_{T_n}^{T_{n+1}} Z_t dt = \left(Id - \mathrm{e}^{-(T_{n+1}-T_n)\tau^{-1}}\right)\tau\left(Z_{T_n} - \epsilon\tau\nabla S(X_{T_n})V_{T_n}\right) + \epsilon(T_{n+1}-T_n)\tau\nabla S(X_{T_n})V_{T_n} + \theta_{n+1}^2\mathcal{O}\left(\epsilon^2\right).$$

Finally, (2.13) gives

$$\theta_{n+1} = \lambda_0\Delta T_{n+1} + b^T\int_{T_n}^{T_{n+1}} Z_t dt + \theta_{n+1}^2\mathcal{O}\left(\epsilon^2\right),$$

and the result follows.

The proof is the same using Landau symbols $\mathcal{O}_\tau$ in the non-linear case (Assumption 2.3). □

The following estimate of jump times then follows:

LEMMA 3.5. *Suppose either Assumption 2.3 or 2.4 are satisfied, and assume moreover that* $\left\|\tau^{-1}\right\| \le K$ *is bounded and* $\dfrac{\epsilon}{r(\tau^{-1})}$ *is sufficiently small. The jump time variations can be written in the following form:*

$$\Delta T_{n+1} = \Delta T_{n+1}^0 + \epsilon\Delta T_{n+1}^1 + (\theta_{n+1}^3+1)\mathcal{O}_\tau\left(\epsilon^2\right), \tag{3.15}$$

where

$$\begin{aligned}\Delta T_{n+1}^0 &= \frac{\theta_{n+1}}{\lambda_0} + b^T\left(Id - \mathrm{e}^{-\Delta T_{n+1}^0\tau^{-1}}\right)\tau Z_{T_n}\\ &= \frac{\theta_{n+1}}{\lambda_0} + \theta_{n+1}\mathcal{O}_\tau(Z_n)\end{aligned} \tag{3.16}$$

and

$$\begin{aligned}\Delta T_{n+1}^1 &= \frac{1}{\lambda_0 - b^T\mathrm{e}^{-\Delta T_{n+1}^0\tau^{-1}}Z_{T_n}} b\left(\Delta T_{n+1}^0\tau - (\mathrm{Id} - \mathrm{e}^{-\Delta T_{n+1}^0\tau^{-1}})\tau^2\right)\nabla S(X_{T_n})V_{T_n}\\ &= \frac{1}{\lambda_0}b^T\left(\frac{\theta_{n+1}\tau}{\lambda_0} - \left(Id - \mathrm{e}^{-\frac{\theta_{n+1}}{\lambda_0}\tau^{-1}}\right)\tau^2\right)\nabla S(X_{T_n})V_{T_n} + (\theta_{n+1}^2+1)\mathcal{O}_\tau(Z_{T_n}).\end{aligned} \tag{3.17}$$

Under Assumption 2.4, the Landau symbol $\mathcal{O}_\tau$ *in the above is uniform in* τ.

Proof. Let us first write the proof in the linear case (Assumption 2.4). Let us re-write (3.12) with:

$$\begin{cases}\Psi_0(\Delta T_{n+1},\tau) = \lambda_0\Delta T_{n+1} + \left(Id - \mathrm{e}^{-\Delta T_{n+1}\tau^{-1}}\right)\tau Z_{T_n}\\ \Psi_1(\Delta T_{n+1},\tau) = b^T\left(\Delta T_{n+1}\tau - (\mathrm{Id} - \mathrm{e}^{-\Delta T_{n+1}\tau^{-1}})\tau^2\right)\nabla S(X_n)V_n.\end{cases}$$

The formulae (3.16) and (3.17) for ΔT_{n+1}^0 and ΔT_{n+1}^1 are solution to the Taylor expansion in ϵ of the relation (3.12), which amounts to solve $\Psi_0(\Delta T_{n+1}^0,\tau) = 0$, as well as $\partial_1\Psi_0(\Delta T_{n+1}^0,\tau)\Delta T_{n+1}^1 + \Psi_1(\Delta T_{n+1}^0,\tau) = 0$. To make the proof rigorous, we need to show that the latter equations have a unique solution, and to estimate the different terms involved in the computations. First of all, note that by construction $|\Delta T_{n+1}| \le K\theta_{n+1}$. Then the following key invertibility condition is satisfied:

$$|\partial_1\Psi_0(t,\tau)|^{-1} = \left|\frac{1}{\lambda_0 - b^T\mathrm{e}^{-t\tau^{-1}}Z_{T_n}}\right| \le K \tag{3.18}$$

using 2.16 in the linear case for $\frac{\epsilon}{r(\tau^{-1})}$ sufficiently small. Using the invertibility condition (3.18), the implicit function theorem ensures that $(\Delta T^0_{n+1}, \Delta T^1_{n+1})$ are well defined and satisfies $\left|\Delta T^0_{n+1}\right| \leq K\theta_{n+1}$.

Next some key estimates concern the function:

$$M(t) = t\tau - (\mathrm{Id} - \mathrm{e}^{-t\tau^{-1}})\tau^2,$$

which satisfies $M(0)=0$, $M'(0)=0$ and $\|M''(t)\| = \left\|\mathrm{e}^{-t\tau^{-1}}\right\| \leq K$ so that

$$\sup_{t\in[0,t_{\max}]} \|M(t)\| \leq K t_{\max}^2, \qquad \sup_{t\in[0,t_{\max}]} \|M'(t)\| \leq K t_{\max}. \tag{3.19}$$

This will yield estimates on ΔT^0_{n+1} and ΔT^1_{n+1}. First we get $\left|\Delta T^1_{n+1}\right| \leq K\theta^2_{n+1}$. As a consequence of (3.19)-(3.18), $\partial_1\Psi_0(t,\tau)^{-1}$ and $\Psi_1(t,\tau)$ remain bounded respectively by by K and $K\theta^2_{n+1}$ uniformly in τ and $t\in[0,K\theta_{n+1}]$. Inserting in the Taylor ϵ-expansion, these estimates show rigorously that $\Delta T_{n+1} = \Delta T^0_{n+1} + (1+\theta^2_{n+1})\mathcal{O}(\epsilon)$. Next, it is possible to check the same way that $\partial_1\Psi_0(t,\tau)$, $\partial_1^2\Psi_0(t,\tau)$, and $\partial_1\Psi_1(t,\tau)$ remain bounded respectively by K, $K\left\|\tau^{-1}\right\|$, and $K\theta_{n+1}$, uniformly in τ and $t\in[0,K\theta_{n+1}]$. Inserting in the expansion, these estimates show rigorously that $\Delta T_{n+1} = \Delta T^0_{n+1} + \epsilon\Delta T^1_{n+1} + (\theta^3_{n+1}+1)\mathcal{O}(\epsilon^2)$. This concludes the proof.

The proof is the same using Landau symbols $\mathcal{O}_\tau$ in the non-linear case (Assumption 2.3). □ Then the diffusive limit of the process with internal state can be computed as follows:

PROPOSITION 3.6. *Suppose either Assumption 2.3 is satisfied with a fixed τ, or Assumption 2.4 with $\epsilon \ll r(\tau_\epsilon^{-1}) \leq K$. Assume $|Z_0| = \mathcal{O}(\epsilon)$. Then the process $\bar{t} \mapsto X^\epsilon_{\bar{t}} = X_{t/\epsilon^2}$ converges in distribution (for uniform convergence topology) towards $\bar{t} \mapsto X^0_{\bar{t}}$ solution to the SDE (3.1). As a consequence, the density $n^\epsilon_{\bar{t}}$ of the process converges weakly towards $n^0_{\bar{t}} = n^{c,0}_{\bar{t}}$ solution of (3.2).*

Proof. Let us first write the proof in the linear case (Assumption 2.4). Let us consider again the process with random time change:

$$\tilde{X}^\epsilon_{\bar{t}} = X_0 + \sum_{n=0}^{[\lambda_0\bar{t}/\epsilon^2]} \epsilon V_n (T_{n+1} - T_n). \tag{3.20}$$

We use again standard results on diffusion approximation of Markov chains to compute the limit of the latter process ([21]). Since $\mathcal{V}_n$, θ_{n+1} and Z_n are independent, the bias of the term of the sum (3.20) is given using (3.17) by

$$\begin{aligned}\mathbb{E}((T_{n+1}-T_n)V_{T_n}|X_{T_n}) = b^T\mathbb{E}\left(\frac{\theta_{n+1}\tau}{\lambda_0} - \left(Id - \mathrm{e}^{-\frac{\theta_{n+1}}{\lambda_0}\tau^{-1}}\right)\tau^2\right)\mathbb{E}(\nabla S(X_{T_n})\cdot V_{T_n}V_{T_n}) \\ + \mathbb{E}(\mathcal{O}(Z_n)|X_{T_n});\end{aligned}$$

so that remarking that:

$$\mathbb{E}\left(\frac{\theta_{n+1}\tau}{\lambda_0} - \left(Id - \mathrm{e}^{-\frac{\theta_{n+1}}{\lambda_0\tau}}\right)\tau^2\right) = \frac{\tau^2}{\lambda_0\tau(\lambda_0\tau+\mathrm{Id})} = \frac{\tau}{\lambda_0(\lambda_0\tau+\mathrm{Id})}$$

we get the following formula:

$$\mathbb{E}((T_{n+1}-T_n)V_{T_n}|X_{T_n}) = \epsilon b^T \frac{\tau}{\lambda_0(\lambda_0\tau+\mathrm{Id})}\nabla S(X_{T_n})D + \epsilon\mathbb{E}(\mathcal{O}(Z_n)|X_{T_n}).$$

The variance term yields, using the estimates (3.19):

$$\mathbb{E}\left((T_{n+1}-T_n)^2 V_{T_n}\otimes V_{T_n}|X_{T_n}\right)=\frac{2}{\lambda_0}D+\mathbb{E}(\mathcal{O}(Z_n)|X_{T_n})+\mathcal{O}(\epsilon).$$

Using the fact that $\mathcal{O}(Z_n)=\mathcal{O}(\epsilon/r(\tau^{-1}))$ enables to conclude on the convergence of the process $\bar{t}\mapsto\tilde{X}^{\epsilon}_{\bar{t}}$. The fact that $\bar{t}\mapsto X^{\epsilon}_{\bar{t}}$ converges to the same limit is obtained by a random time change as done in the proof of Proposition 3.2.

The proof is the same using Landau symbols $\mathcal{O}_\tau$ in the non-linear case (Assumption 2.3). □

REMARK 3.7. *In one spatial dimension and with internal dynamics given by* (2.7)*, Erban and Othmer argued on formal arguments (a moment expansion) that the evolution of the density* $n(x,t)$ *on hyperbolic time-scales satisfies*

$$\partial_t^2 n+2\lambda_0\partial_t n=\partial_x\left(\epsilon^2\partial_x n-\frac{2b\epsilon^2\tau}{1+2\lambda_0\tau}S'(x)n\right),\tag{3.21}$$

in the limit when $\epsilon\to 0$ *[10]. It is easy to see that equation* (3.21) *is equivalent to* (3.1) *when choosing* $T(x)$ *in* (3.1) *according to* (2.21)*.*

3.3. Limit of the coupling. In this section, we show that the difference on diffusive time scales $\bar{t}=t/\epsilon^2$ between the two coupled processes in the scale separated regime (2.17) will remain of order $\mathcal{O}\left(\frac{\epsilon}{r(\tau_\epsilon^{-1})}\right)$ under Assumption 2.4 in the linear case, or $\mathcal{O}(\epsilon)_\tau$ under Assumption 2.3 in the non-linear case.

PROPOSITION 3.8. *Suppose either Assumption 2.3 or 2.4 are satisfied, and let* $|Z_0|=\mathcal{O}(\epsilon)$*. Consider the random time changes* $\bar{t}\mapsto\alpha_{c,\epsilon}(\bar{t})$ *defined in* (3.5)*, and* $\bar{t}\mapsto\alpha_\epsilon(\bar{t})$ *defined similarly for the process with internal state. The latter two processes converges in distribution under the uniform convergence topology towards the identity time function* $\bar{t}\to\bar{t}$*. Then the difference between the process with internal state and the coupling process satisfies:*

$$\mathbb{E}\left(\left(X_{\alpha_\epsilon(\bar{t})/\epsilon^2}-X^c_{\alpha_{c,\epsilon}(\bar{t})/\epsilon^2}\right)^2\right)\le K_{\tau,\bar{t}}\epsilon^2.\tag{3.22}$$

In particular, for any $\bar{t}\ge 0$*, it yields*

$$\lim_{\epsilon\to 0}\mathbb{E}\left(\left(X_{\bar{t}/\epsilon^2}-X^c_{\bar{t}/\epsilon^2}\right)^2\right)=0.\tag{3.23}$$

Under Assumption 2.4, the constant $K_{\tau,\bar{t}}$ *in* (3.22) *is of the form* $K_{\tau,\bar{t}}=\frac{K_{\bar{t}}}{r(\tau^{-1})^2}$*; and the convergence* (3.23) *occurs in the limit* $\epsilon/r(\tau_\epsilon^{-1})\to 0$*.*

Proof. Notations of the proof of Lemma 3.5 and Proposition 3.6 are used. We want to estimate precisely:

$$\begin{aligned}\mathbb{E}\left(\left|\tilde{X}^{c,\epsilon}_{\bar{t}}-\tilde{X}^{\epsilon}_{\bar{t}}\right|^2\right)&=\mathbb{E}\left(\left|\sum_{n=1}^{[\lambda_0\bar{t}/\epsilon^2]-1}\epsilon\mathcal{V}_n(\Delta T^c_{n+1}-\Delta T_{n+1})\right|^2\right)\\&\le 3\mathbb{E}\left|\sum_{n=1}^{[\lambda_0\bar{t}/\epsilon^2]-1}A_n\right|^2+3\mathbb{E}\left|\sum_{n=1}^{[\lambda_0\bar{t}/\epsilon^2]-1}B_n\right|^2+3\mathbb{E}\left|\sum_{n=1}^{[\lambda_0\bar{t}/\epsilon^2]-1}C_n\right|^2,\end{aligned}\tag{3.24}$$

$$\underbrace{\qquad\qquad}_{(i)}\qquad\underbrace{\qquad\qquad}_{(ii)}\qquad\underbrace{\qquad\qquad}_{(iii)}$$

where we have decomposed $\mathcal{V}_n(\Delta T^c_{n+1}-\Delta T_{n+1})$ using (3.16), (3.17), and (3.3) into:

- a leading term with null average :

$$A_n = \epsilon\theta_{n+1}\mathcal{O}(Z_{T_n})\mathcal{V}_n + \epsilon^2\frac{\mathcal{V}_n}{\lambda_0}\left(A^T(X_{T_n}) - b^T\left(\frac{\theta_{n+1}\tau}{\lambda_0} - \left(Id - \mathrm{e}^{-\frac{\theta_{n+1}}{\lambda_0}\tau^{-1}}\right)\tau^2\right)\nabla S(X_{T_n})\right)\mathcal{V}_n;$$

- a drift term:

$$B_n = \epsilon^2\frac{\mathcal{V}_n}{\lambda_0}\left(A^T(X_{T_n}) - A^T(X^c_{T_n})\right)\mathcal{V}_n;$$

- and a rest term:

$$C_n = \epsilon\theta^2_{n+1}\mathcal{O}(Z_{T_n}) + (1+\theta_{n+1})\mathcal{O}(\epsilon^2).$$

Using the independence of the random variables $(\mathcal{V}_n, \theta_n)_{n\geq 1}$, we can expand the square (i) and all non-diagonal elements have null average. Using the estimates (3.19), as well as (2.16) we get:

$$\mathbb{E}\left|\sum_{n=1}^{[\lambda_0\bar{t}/\epsilon^2]-1} A_n\right|^2 = \mathcal{O}\left(\frac{\epsilon^2}{r(\tau^{-1})^2}\right) + \mathcal{O}(\epsilon^2).$$

Next, we use Jensen inequality in (ii) to get:

$$\begin{aligned}
\mathbb{E}\left|\sum_{n=1}^{[\lambda_0\bar{t}/\epsilon^2]-1} B_n\right|^2 &\leq \lambda_0\bar{t}/\epsilon^2\mathbb{E}\left(\sum_{n=1}^{[\lambda_0\bar{t}/\epsilon^2]-1}|B_n|^2\right)\\
&\leq \frac{\bar{t}}{\lambda_0}\epsilon^2\sum_{n=1}^{[\lambda_0\bar{t}/\epsilon^2]-1}\mathbb{E}\left|A^T(X_{T_n}) - A^T(X^c_{T_n})\right|^2\\
&\leq 2\bar{t}\int_0^{\bar{t}}\mathbb{E}\left|A^T(\tilde{X}^\epsilon_{\bar{s}}) - A^T(\tilde{X}^{c,\epsilon}_{\bar{s}})\right|^2 d\bar{s} + \mathcal{O}(\epsilon^2)\\
&\leq 4\bar{t}\int_0^{\bar{t}}\mathbb{E}\left|\tilde{X}^\epsilon_{\bar{s}} - \tilde{X}^{c,\epsilon}_{\bar{s}}\right|^2 d\bar{s} + \mathcal{O}(\epsilon^2),
\end{aligned}$$

where in the last two lines we use the fact that $\|\nabla A\|_\infty$ is bounded uniformly in τ. Finally, (iii) is bounded using (2.16) by $\mathcal{O}\left(\frac{\epsilon^2}{r(\tau^{-1})^2}\right)$. Eventually, we get for $\bar{t}\leq\bar{T}$:

$$\mathbb{E}\left(\left|\tilde{X}^{c,\epsilon}_{\bar{t}} - \tilde{X}^\epsilon_{\bar{t}}\right|^2\right) \leq K_{\bar{T}}\int_0^{\bar{t}}\mathbb{E}\left|\tilde{X}^\epsilon_{\bar{s}} - \tilde{X}^{c,\epsilon}_{\bar{s}}\right|^2 d\bar{s} + \mathcal{O}(\frac{\epsilon^2}{r(\tau^{-1})^2}),$$

and a Gronwall argument yields the result.

Finally, we use the following difference:

$$X^{c,\epsilon}_{\alpha^c_\epsilon(\bar{t})} - X^\epsilon_{\alpha_\epsilon(\bar{t})} = \tilde{X}^{c,\epsilon}_{\bar{t}} - \tilde{X}^\epsilon_{\bar{t}} + \theta_{\lfloor\lambda_0\bar{t}/\epsilon^2\rfloor}\mathcal{O}(\epsilon)$$

to conclude. The convergence (3.23) follows from the convergence of the random time change and a Skorokhod embedding argument as already used in Propositions 3.2. □

This result shows that the variance of the coupling scales as the difference between the two models, hence the term "asymptotic" variance reduction. A consequence is

that the density of bacteria on the diffusive time scales satisfies under Assumption 2.4, for any continuous and bounded test function φ:

$$\left| \int_{\mathbb{R}^d} \varphi(x) \left(n(x,\bar{t}) - n^c(x,\bar{t}) \right) dx \right| \leq K_{\bar{t}} \frac{\epsilon}{r(\tau^{-1})},$$

and under Assumption 2.3:

$$\left| \int_{\mathbb{R}^d} \varphi(x) \left(n_{\bar{t}}(x,\bar{t}) - n^c(x,\bar{t}) \right) dx \right| \leq K_{\tau,\bar{t}} \epsilon.$$

4. Discretization of velocity-jump processes

To observe the derived coupling behaviour in simulations, the time discretization should retain these coupling properties. In particular, the jump times should be computed sufficiently accurately. In this section, we propose a time discretization of (2.14) and give a consistency result. The control process (2.19) can be discretized similarly. For ease of exposition, we consider the scalar equation (2.7) for the internal state; generalization to nonlinear systems of equations is suggested in Remark 4.1.

4.1. Linear turning rate.

First, assume a linearized jump rate

$$\lambda^l(z) = \lambda_0 - b^T z. \tag{4.1}$$

When the chemoattractant profile is linear, $S(x) = S + \nabla S x$, the ODE (2.7) can be explicitly solved for Z_t. For $t \in [T_n, T_{n+1})$, *i.e.* between two jumps, we then have

$$\begin{cases} X_t = X_{T_n} + \epsilon V_t \left(t - T_n \right) \\ Z_t = \exp(-(t-T_n)\tau^{-1}) Z_{T_n} + \epsilon\tau \left(\mathrm{Id} - \exp(-(t-T_n)\tau^{-1}) \right) \nabla S \mathcal{V}_n. \end{cases} \tag{4.2}$$

Similarly to the computation of Lemma 3.4, it yields using $\Delta T_{n+1} := T_{n+1} - T_n$,

$$\begin{aligned} \int_{T_n}^{T_{n+1}} \lambda^l(Z_t) \mathrm{d}t = \lambda_0 \left(\Delta T_{n+1} \right) - b^T \left(Id - \mathrm{e}^{-\Delta T_{n+1} \tau^{-1}} \right) \tau Z_{T_n} \\ - \epsilon b^T \left(\Delta T_{n+1} \tau - (\mathrm{Id} - \mathrm{e}^{-\Delta T_{n+1} \tau^{-1}}) \tau^2 \right) \nabla S \mathcal{V}_n. \end{aligned} \tag{4.3}$$

Using (4.3), the jump time T_{n+1} can be calculated exactly by solving

$$\int_{T_n}^{T_{n+1}} \lambda^l(Z_t) \mathrm{d}t = \theta_{n+1}, \tag{4.4}$$

using a Newton iteration. Hence, no time discretization error is made.

When $S(x)$ is a nonlinear function of x, exact integration of Y_t is no longer possible. We therefore define a numerical solution $(X_t^{\delta t}, V_t^{\delta t}, Y_t^{\delta t})$ as follows. Between jumps, we discretize the simulation in steps of size δt and denote by $(X_{n,k}^{\delta t}, Z_{n,k}^{\delta t})$ the solution at $t_{n,k} = T_n^{\delta t} + k\delta t$. The numerical solution for $t \in [t_{n,k}, t_{n,k+1}]$ is given by

$$\begin{cases} X_t^{\delta t} = X_{n,k}^{\delta t} + \epsilon \mathcal{V}_n \left(t - t_{n,k} \right) \\ Z_t^{\delta t} = \exp(-(t-t_{n,k})\tau^{-1}) Z_{n,k}^{\delta t} + \epsilon\tau \left(\mathrm{Id} - \exp(-(t-t_{n,k})\tau^{-1}) \right) \nabla S(X_{n,k}^{\delta t}) \mathcal{V}_n. \end{cases} \tag{4.5}$$

We denote by $K \geq 0$ the integer such that the simulated jump time $T_{n+1}^{\delta t} \in [t_{n,K}, t_{n,K+1}]$. To find $T_{n+1}^{\delta t}$, we first approximate the integral $\int_{T_n^{\delta t}}^{T_{n+1}^{\delta t}} \lambda^l(Z_t^{\delta t})\mathrm{d}t$ using

$$\int_{T_n^{\delta t}}^{T_{n+1}^{\delta t}} \lambda^l(Z_t^{\delta t})\mathrm{d}t = \sum_{k=0}^{K-1} \int_{t_{n,k}}^{t_{n,k+1}} \lambda^l(Z_t^{\delta t})\mathrm{d}t + \int_{t_{n,K}}^{T_{n+1}^{\delta t}} \lambda^l(Z_t^{\delta t})\mathrm{d}t, \tag{4.6}$$

and then compute:

$$\int_{t_{n,k}}^{t_{n,k+1}} \lambda^l(Z_t^{\delta t})\mathrm{d}t = \lambda_0 \delta t - b^T \left(Id - \mathrm{e}^{-\delta t \tau^{-1}} \right) \tau Z_{T_n}^{\delta t} - \epsilon b^T \left(\delta t \tau - (\mathrm{Id} - \mathrm{e}^{-\delta t \tau^{-1}})\tau^2 \right) \nabla S(X_{n,k}^{\delta t}) \mathcal{V}_n. \tag{4.7}$$

The jump time $T_{n+1}^{\delta t}$ can then be computed as the solution of

$$\int_{t_{n,K}}^{T_{n+1}^{\delta t}} \lambda^l(Z_t^{\delta t})\mathrm{d}t = \theta_{n+1} - \sum_{k=0}^{K-1} \int_{t_{n,k}}^{t_{n,k+1}} \lambda^l(Z_t^{\delta t})\mathrm{d}t, \tag{4.8}$$

again using a Newton procedure. The use of (4.5) as a discretization scheme has a crucial consequence: the analysis performed in Section 3 for the continuous time case, still holds *in an exact fashion* in the time discretized case. This is mainly due to the fact that the key estimate (3.14) still holds. It is left to the reader to check that all the proofs performed in Section 3 are not affected by the discretization.

REMARK 4.1. *The discretization scheme* (4.5) *can be generalized to a non-linear evolution equation for the internal state. It is however necessary to check that a time discrete version of Assumption 2.3 still holds in this case, and then that* (3.14) *remains true.*

We also give the following consistency property, which is of independent interest:

LEMMA 4.2. *Assume* $T_n^{\delta t} = T_n$, $X_{T_n} = X_{T_n}^{\delta t}$, $\mathcal{V}_n = V_{T_n}^{\delta t}$ *and* $Y_{T_n}^{\delta t} = Y_{T_n}$, *then*

$$\left| T_{n+1} - T_{n+1}^{\delta t} \right| \leq K \frac{\epsilon^2}{r(\tau^{-1})} \delta t, \tag{4.9}$$

and consequently,

$$\left| X_{T_{n+1}} - X_{T_{n+1}^{\delta t}}^{\delta t} \right| \leq K \frac{\epsilon^3}{r(\tau^{-1})} \delta t. \tag{4.10}$$

$$\tag{4.11}$$

The dependence of the constant K is precised in Remark 2.2. The proof is given in Appendix B. Note that the dependance in $\frac{\epsilon^3}{r(\tau^{-1})}\delta t$ shows, at least at the consistency level, that the error due to discretization remains smaller than the difference between the coupling process and the process with internal state.

4.2. Nonlinear turning rate. We now consider the nonlinear turning rate as defined in (2.8). We again discretize in time to obtain the time-discrete solution (4.5). The jump time $T_{n+1}^{\delta t}$ is now computed by linearizing (2.8) in each time step,

$$\lambda^{\delta t}\left(Z_t^{\delta t}, Z_{n,k}^{\delta t}\right) = \lambda(Z_{n,k}^{\delta t}) + \frac{\mathrm{d}\lambda(Z_{n,k}^{\delta t})}{\mathrm{d}z}\left(Z_t^{\delta t} - Z_{n,k}^{\delta t}\right) \tag{4.12}$$

and approximating the integral

$$\int_{T_n^{\delta t}}^{T_{n+1}^{\delta t}} \lambda(Z_t^{\delta t}, Z^{\delta t}_{n,\lfloor (t-T_n^{\delta t})/\delta t \rfloor})\mathrm{d}t = \sum_{k=0}^{K-1} \int_{t_{n,k}}^{t_{n,k+1}} \lambda(Z_t^{\delta t}, Z^{\delta t}_{n,k})\mathrm{d}t + \int_{t_{n,K}}^{T_{n+1}^{\delta t}} \lambda(Z_t^{\delta t}, Z^{\delta t}_{n,K})\mathrm{d}t \tag{4.13}$$

$$= \theta_{n+1}, \tag{4.14}$$

The jump time $T_{n+1}^{\delta t}$ can then again be computed using a Newton procedure as in (4.8). Here again, the analysis performed in Section 3 for the continuous time case, still holds *in an exact fashion* in the time discretized case. The consistency property similar to Lemma 4.2 still holds in the case of non-linear turning rate.

LEMMA 4.3. *Assume* $T_n^{\delta t} = T_n$, $X_{T_n} = X_{T_n}^{\delta t}$, $\mathcal{V}_n = V_{T_n}^{\delta t}$ *and* $Y_{T_n}^{\delta t} = Y_{T_n}$, *then*

$$\left| T_{n+1} - T_{n+1}^{\delta t} \right| \leq K \frac{\epsilon^2}{r(\tau^{-1})} \delta t, \tag{4.15}$$

and consequently,

$$\left| X_{T_{n+1}} - X_{T_{n+1}^{\delta t}}^{\delta t} \right| \leq K \frac{\epsilon^3}{r(\tau^{-1})} \delta t + K \frac{\epsilon^3 \left\| \tau^{-1} \right\|}{r(\tau^{-1})^2} \delta t^3. \tag{4.16}$$

$$\tag{4.17}$$

The dependence of the constant K is precised in Remark 2.2.

The proof is given in appendix C.

5. Numerical illustration

In this section, we demonstrate the validity of the analysis above. For the numerical experiments, we restrict ourselves to one space dimension. In this case, the kinetic equation corresponding to the control process reduces to the system (3.8) of two PDEs, which is straightforward to simulate using finite differences. We fix a scalar bimodal chemoattractant concentration field

$$S(x) = \alpha \left(\exp\left(-\delta (x-\xi)^2 \right) + \exp\left(-\delta (x-\eta)^2 \right) \right), \tag{5.1}$$

in which $\alpha = \delta = 1$, $\xi = 7.5$ and $\eta = 12.5$.

For the process with internal state, we use internal dynamics that are given by the scalar cartoon model (2.7). The corresponding turning rate is given by (2.8) with $\beta = 1$. For the control process, we choose the linear turning rate (2.20), with parameters (2.21) such that the two processes have the same hydrodynamic limit. The physical domain is $x \in [0,20]$, and we use reflecting boundary conditions, *i.e.* the bacterial velocity is reversed when $x = 0$ or $x = 20$.

In a first experiment, we simulate a single bacterium evolving according to the fine-scale model (2.14), in which we set the parameters to $\epsilon = 5 \cdot 10^{-2}$, $t_a = 10$, $\lambda_0 = 1$ and $b = 1$. We compare this evolution to that of a bacterium that satisfies the corresponding control process (2.19). Both simulations are performed using the same random numbers starting from the initial position $X_0 = X_0^c = 7.5$ and $V_0 = V_0^c = +1$. The time step $\delta t = 1$. The results are shown in figure 5.1. We see a very good coupling initially, which degrades over time. Note the time shift in the short time picture, which is the numerical illustration of the difference of time change

$$\alpha_\epsilon(\bar{t}) \neq \alpha_{c,\epsilon}(\bar{t});$$

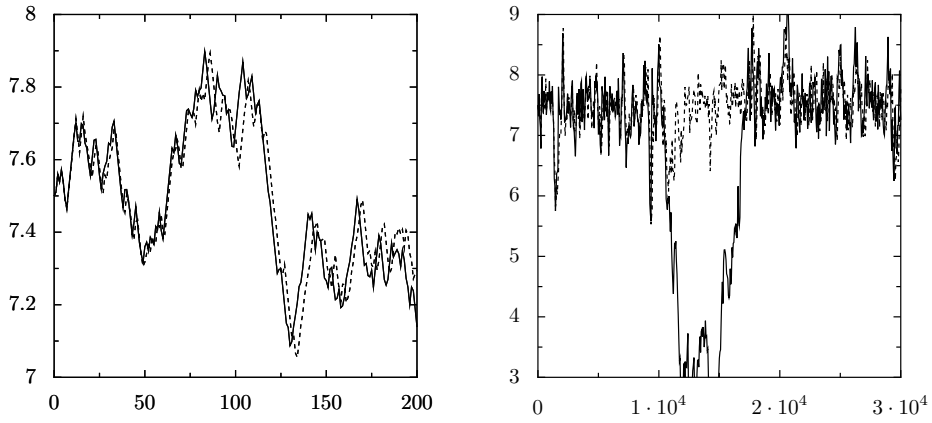

FIG. 5.1. *Evolution of a single bacterium evolving according to the fine-scale process (2.14) (solid line) and the corresponding control process (2.19) (dashed) on short (left) and long (right) time-scales.*

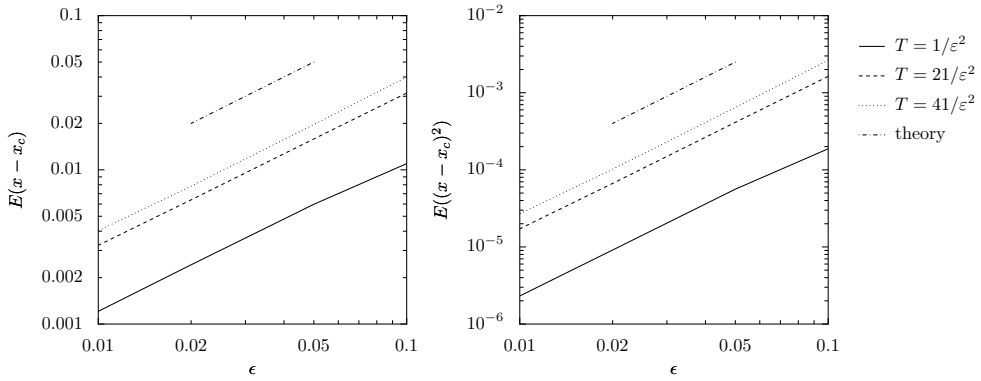

FIG. 5.2. *Empirical mean (left) and variance (right) of the difference between the fine-scale process (2.14) and the corresponding control process (2.19) as a function of ϵ for different values of the reporting time T. The theoretical slope is indicated with a dashdotted line. The sample size is $N = 10000$.*

as defined in Section 3. In the long time picture, coupling is completely lost at some point, when the bacterium with internal state "escapes" the chemoattactant peak at $x = 7.5$.

Next, we repeat the experiment using $N = 10000$ particles and compute the empirical mean and variance of the coupling, i.e.

$$E(X_{\bar{t}} - X^c_{\bar{t}}) = \frac{1}{N}\sum_i^N \left(X^i_{\bar{t}} - X^{i,c}_{\bar{t}}\right), \text{ resp., } \quad E\left((X_{\bar{t}} - X^c_{\bar{t}})^2\right) = \frac{1}{N}\sum_i^N \left(X^i_{\bar{t}} - X^{i,c}_{\bar{t}}\right)^2.$$

As fine-scale parameters, we choose $t_a = 1$, $\lambda_0 = b = 1$ and several values of ϵ.

Figure 5.2 shows the dependence in ϵ of the coupling, by plotting the empirical mean and variance defined above as a function of ϵ for different values of the reporting time T. The results shown in figure 5.2 are in clear accordance with the theoretical slope predicted by the asymptotic analysis. Figure 5.3 shows that the mean and

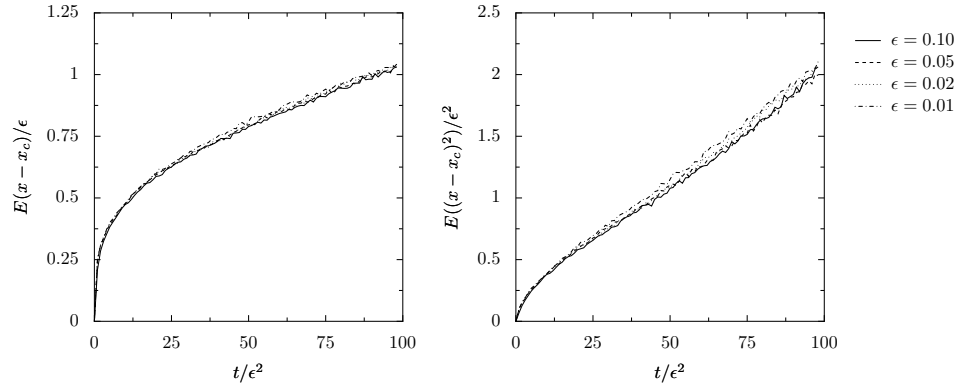

FIG. 5.3. *Evolution of the empirical mean (left) and variance (right) of the difference between the fine-scale process (2.14) and the corresponding control process (2.19) as a function of* $\bar{t} = t/\epsilon^2$ *for different values of* ϵ. *The sample size is* $N = 10000$.

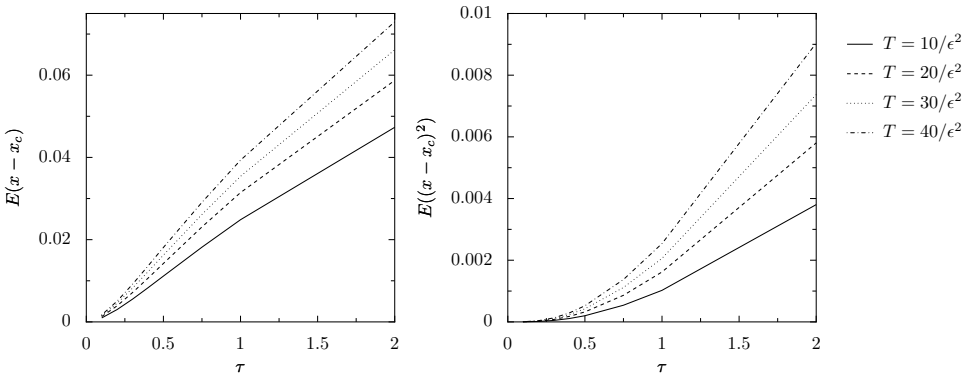

FIG. 5.4. *Empirical mean (left) and variance (right) of the difference between the fine-scale process (2.14) and the corresponding control process (2.19) as a function of* t_a *for different values of the reporting time* T. *The sample size is* $N = 10000$.

variance of the coupling difference as a function of time. Note the seemingly linear dependance of the mean and variance with respect to time on diffusive time scales. This behavior has not been analyzed mathematically in Section 3. It is probably due to: (i) sufficiently short diffusive times; (ii) the specific double-well form of the chemoattractant potential.

Finally, we illustrate the dependence on t_a. To this end, we again simulate $N = 10000$ particles choosing $\lambda_0 = 1$, $\epsilon = 0.1$ and $X_0 = 7.5$, for different values of t_a. The results in figure 5.4 show that the variance quickly increases with $t_a (> 1)$.

6. Simulation on diffusive time scales

In this section, we consider a simulation of the density of an ensemble of particles, with and without variance reduction and/or reinitialization, as described in Section 2.5. We again restrict ourselves to one space dimension and fix the chemoattractant concentration field as (5.1), with the same parameters as in the previous section ($\alpha = \delta = 1$, $\xi = 7.5$ and $\eta = 12.5$). The same model is used ((2.7)-(2.8), (2.20)-(2.21)),

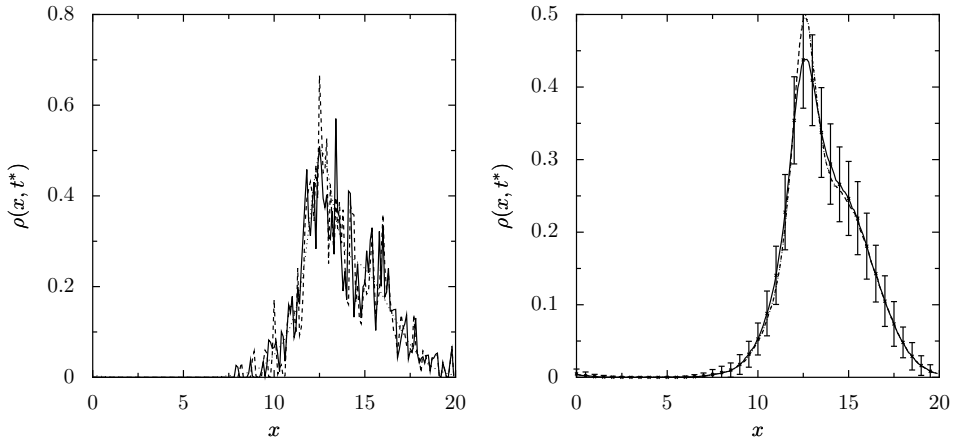

FIG. 6.1. *Bacterial density as a function of space at $t=100/\epsilon^2$ without variance reduction. Left: one realization. Right: mean over 200 realizations and 95% confidence interval. The solid line is the estimated density from a particle simulation using the process with internal state; the dashed line is estimated from a particle simulation using the control process. Both used $N=3750$ particles. The dotted line is the solution of the deterministic PDE (2.25).*

but with domain $x\in[0,20]$ with periodic boundary conditions. The parameters are $\epsilon=0.2$, $\lambda_0=1$, $t_a=5$, $\beta=1$, $\delta t=1$.

All simulations are performed with $N=3750$ particles. The initial positions are uniformly distributed in the interval $x\in[13,15]$; the initial velocities are chosen uniformly, i.e. each particle has an equal probability of having an initial velocity of $\pm\epsilon$. The initial condition for the internal variable is chosen to be in local equilibrium, i.e. $Y_0^i=S(X_0^i)$. The initial positions and velocities of the control particles are chosen to be identical.

We discretize the continuum description (3.8) on a mesh with $\Delta x=0.1$ using a third-order upwind-biased scheme, and perform time integration using the standard fourth order Runge–Kutta method with time step $\delta t_{pde}=10^{-1}$. The initial condition is given as

$$p^+(x,0)=p^-(x,0)=\begin{cases}1, & x\in[13,15],\\ 0, & \text{otherwise.}\end{cases} \tag{6.1}$$

First, we simulate both stochastic processes up to time $\bar{t}=100$ ($t=100/\epsilon^2$) and estimate the density of each of these processes $\hat{n}_N(x,\bar{t})$, resp. $\hat{n}_N^c(x,\bar{t})$, without variance reduction (see Section 2.5). The density is obtained via kernel density estimation using a Gaussian kernel function with bandwidth $h=5\cdot10^{-2}$. Figure 6.1 (left) shows the results for a single realization. We see that, given the fluctuations on the obtained density, it is impossible to conclude on differences between the two models. This observation is confirmed by computing the average density of both processes over 200 realizations. The mean densities are shown in figure 6.1 (right), which also reveals that the mean density of the control process is within the 95% confidence interval of the process with internal state. Both figures also show the density that is computed using the continuum description, which coincides with the mean of the density of the control particles.

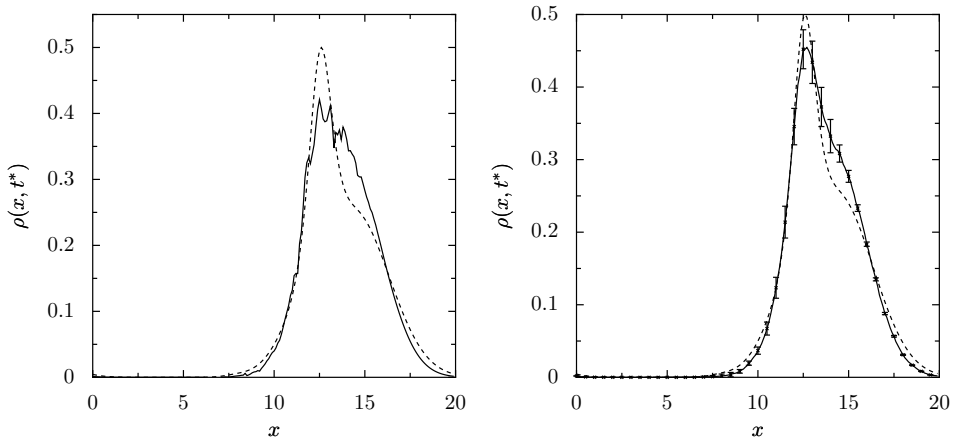

FIG. 6.2. *Bacterial density as a function of space at* $t = 100/\epsilon^2$ *with variance reduction and reinitialization. Left: variance reduced density estimation of one realization with* $N = 3750$ *particles (solid) and deterministic solution for the control process (2.25) (dashed). Right: mean over* 200 *realization and* 95% *confidence interval (solid) and the deterministic solution for the control process (2.25) (dashed).*

Next, we compare the variance reduced estimation (2.27) with the density of the control PDE. We reinitialize the control particles after each coarse-scale step, i.e. each k steps of the particle scheme, where $k\delta t = \delta t_{pde}$, (here $k = 2$). The results are shown in figure 6.2. We see that, using this reinitialization, the difference between the behaviour of the two processes is visually clear from one realization (left figure). Also, the resulting variance is such that the density of the control PDE is no longer within the 95% confidence interval of the variance reduced density estimation (right figure). We see that there is a significant difference between both models: the density corresponding to the control process is more peaked, indicating that bacteria that follow the control process are more sensitive to sudden changes in chemoattractant gradient. This difference can be interpreted from the fact that the bacteria with internal state do not adjust themselves instantaneously to their environment, but instead with a time constant t_a.

Finally, we compare the variance reduced estimation (2.27) with the density of the control PDE without performing any reinitialization of the control process to restore the coupling. Figure 6.3 (right) shows the mean estimation of the density over 200 realizations, as well as the density of the control process. As evidenced by the error bars, simulating a single realization of the process, even with variance reduction, is not able to reliably reveal this difference. This phenomenon is due to the degeneracy of the coupling on long diffusive times. This is also illustrated in figure 6.3 (left), which compares the variance reduced density estimation of a single realization with the density of the control PDE. Note that, as predicted by the analysis, the variance is much larger in regions where $\nabla S(x)$ is large.

7. Conclusions and discussion

In future work, we shall study the development of truly "asymptotic preserving" schemes in the diffusion asymptotics; in the sense that the computational cost of the simulation of processes is independent of the small parameter ϵ. This will require to deal with two kinds of difficulties: (i) We will need to use an asymptotic preserving

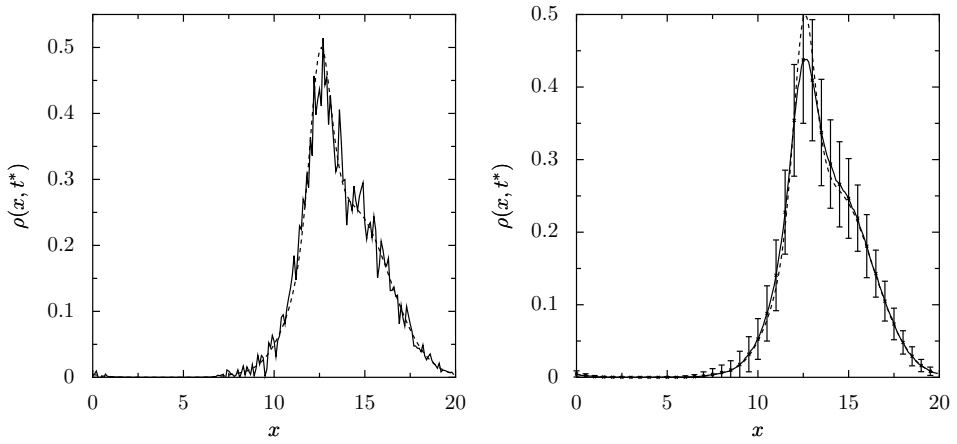

FIG. 6.3. *Bacterial density as a function of space at* $t = 100/\epsilon^2$ *with variance reduction. Left: variance reduced density estimation of one realization with* $N = 3750$ *particles (solid) and deterministic solution for the control process (2.25) (dashed). Right: mean over* 200 *realization and* 95% *confidence interval (solid) and the deterministic solution for the control process (2.25) (dashed).*

method to solve on the grid the density evolution of the control model; and (ii) We will need to extrapolate forward in time the state of the fine-scale simulation.

The first difficulty implies that, instead of solving the full kinetic equation associated with the control process, we only solve its hydrodynamic limit. This may imply the use of a second level of variance reduction, coupling the control velocity jump process and its limiting drift-diffusion process.

The second difficulty, extrapolation in time, is related to the equation-free [19, 20] and HMM [7, 8] types of methodologies; ideas of this type can be traced back to Erhenfest [9]. One approach is to use a "coarse projective integration" method [13, 19]. One then extrapolates the bacterial density over a projective time step on diffusive time scales, after which the projected density needs to be "lifted" to an ensemble of individual bacteria. For such methods, besides the variance of the obtained results, also the effects of reconstructing the velocities and internal variables have to be systematically studied.

Acknowledgements. The authors thank Radek Erban, Thierry Goudon, Yannis Kevrekidis and Tony Lelièvre for interesting discussions that eventually led to this work. This work was performed during a research stay of GS at SIMPAF (INRIA - Lille). GS would like to thank the whole SIMPAF team for the hospitality and the Research Foundation – Flanders (FWO – Vlaanderen) for funding through a travel grant. GS is a Postdoctoral Fellow of the Research Foundation – Flanders. This work was partially supported by the Research Foundation – Flanders through Research Project G.0130.03 and by the Interuniversity Attraction Poles Programme of the Belgian Science Policy Office through grant IUAP/V/22 (GS). The scientific responsibility rests with its authors.

REFERENCES

[1] W Alt. Biased random-walk models for chemotaxis and related diffusion approximations. *J Math Biol*, 9(2):147–177, 1980.

[2] W Alt. Orientation of cells migrating in a chemotactic gradient. *Adv Appl Probab*, 12(3):566–566, 1980.
[3] A Bren and M Eisenbach. How signals are heard during bacterial chemotaxis: Protein-protein interactions in sensory signal propagation. *Journal of bacteriology*, 182(24):6865–6873, 2000.
[4] MP Brenner, P Constantin, LP Kadanoff, A Schenkel, and SC Venkataramani. Diffusion, attraction and collapse. *Nonlinearity*, 12(4):1071–1098, 1999.
[5] FACC Chalub, PA Markowich, B Perthame, and Christian Schmeiser. Kinetic models for chemotaxis and their drift-diffusion limits. *Monatsh Math*, 142(1-2):123–141, 2004.
[6] G Dimarco and L Pareschi. Hybrid multiscale methods II. kinetic equations. *SIAM Multiscale Model. Simul.*, 6(4):1169–1197, 2008.
[7] W E and B Engquist. The heterogeneous multi-scale methods. *Comm. Math. Sci.*, 1(1):87–132, 2003.
[8] W E, B Engquist, X Li, W Ren, and E Vanden-Eijnden. Heterogeneous multiscale methods: A review. *Commun Comput Phys*, 2(3):367–450, 2007.
[9] P. Ehrenfest and T. Ehrenfest. The conceptual foundations of the statistical approach in mechanics, cornell university press (1959). *Enc. Math. Wiss (English trans. by MJ Moravcsik*, 4:32, 1911.
[10] R Erban and HG Othmer. From individual to collective behavior in bacterial chemotaxis. *SIAM Journal on Applied Mathematics*, 65(2):361–391, 2004.
[11] R Erban and HG Othmer. From signal transduction to spatial pattern formation in e-coli: A paradigm for multiscale modeling in biology. *Multiscale Model Sim*, 3(2):362–394, 2005.
[12] SN Ethier and TG Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics, 1986.
[13] CW Gear, IG Kevrekidis, and C Theodoropoulos. "Coarse" integration/bifurcation analysis via microscopic simulators: micro-Galerkin methods. *Computers and Chemical Engineering*, 26:941–963, 2002.
[14] MA Herrero, E Medina, and JJL Velazquez. Self-similar blow-up for a reaction-diffusion system. *J Comput Appl Math*, 97(1-2):99–119, 1998.
[15] T Hillen and HG Othmer. The diffusion limit of transport equations derived from velocity-jump processes. *SIAM Journal on Applied Mathematics*, 61(3):751–775, 2000.
[16] Dirk Horstmann. From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. I. *Jahresber. Deutsch. Math.-Verein.*, 105(3):103–165, 2003.
[17] Dirk Horstmann. From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. II. *Jahresber. Deutsch. Math.-Verein.*, 106(2):51–69, 2004.
[18] E Keller and L Segel. Initiation of slime mold aggregation viewed as an instability. *Journal on Theoretical Biology*, 26:399–415, 1970.
[19] IG Kevrekidis, C. W Gear, JM Hyman, PG Kevrekidis, O Runborg, and C Theodoropoulos. Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level tasks. *Comm. Math. Sciences*, 1(4):715–762, 2003.
[20] IG Kevrekidis and G Samaey. Equation-free multiscale computation: Algorithms and applications. *Annu Rev Phys Chem*, 60:321–344, 2009.
[21] P Kloeden and E Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1992.
[22] HJS Kushner. *Approximation and weak convergence methods for random processes, with applications to stochastic systems theory*. MIT Press, Cambridge, MA, 1984.
[23] HG Othmer, SR Dunbar, and W Alt. Models of dispersal in biological-systems. *J Math Biol*, 26(3):263–298, 1988.
[24] HG Othmer and T Hillen. The diffusion limit of transport equations ii: Chemotaxis equations. *SIAM Journal on Applied Mathematics*, 62(4):1222–1250, 2002.
[25] HG Othmer and A Stevens. Aggregation, blowup, and collapse: The ABC's of taxis in reinforced random walks. *SIAM Journal on Applied Mathematics*, 57(4):1044–1081, 1997.
[26] C Patlak. Random walk with persistence and external bias. *B Math Biol*, 15:311–338, 1953.
[27] DW Scott. *Multivariate density estimation: theory, practice, and visualization*. Wiley, 1992.
[28] BW Silverman. *Density estimation for statistics and data analysis*, volume 26 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, 1986.
[29] AM Stock. A nonlinear stimulus-response relation in bacterial chemotaxis. *P Natl Acad Sci Usa*, 96(20):10945–10947, 1999.
[30] DW Stroock. Some stochastic processes which arise from a model of the motion of a bacterium. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 28:303–315, 1973/74.

Appendix A. Limit of the control process with infinitely fast learning.

LEMMA A.1. *Consider a sequence* $(\tau_k, b_k)_{k\geq 0}$ *with* $\|\tau_k\| \underset{k\to\infty}{\to} 0$, *and* $b_k^T \tau_k \to \bar{b}^T$ *for some vector* $\bar{b} \in \mathbb{R}^n$, *and*

$$\sup_k \frac{\left\|\tau_k^{-1}\right\|}{r(\tau_k^{-1})} < +\infty. \tag{A.1}$$

Consider the process with internal state $t \mapsto (X_t^k, V_t^k, Y_t^k)$ *solution of* (2.14), *and assume (2.11)-(2.13). The process converges almost surely, with respect to the Skorohod metric, towards the gradient sensing process solution of* (2.19) *with velocity-jump rate:*

$$\lambda(x,v) = \lambda_0 - \epsilon\,\lambda = \lambda_0 - \epsilon \bar{b}^T \nabla S(x) v.$$

Proof. Consider a given realisation of the exponentially distributed random numbers $(\theta_n)_{n\geq 1}$, which satisfy $\theta_n > 0$ for all $n \geq 0$. First, since the the rate $\lambda(z)$ is bounded from below by $\lambda_{\min}$, the number of jumps is bounded uniformly in k. Moreover, the size of velocity jumps $\left|V_t^k - V_{t^-}^k\right|$ is bounded by 2, and the norm of the time derivative

$$\frac{d}{dt}(X_t^k, V_t^k) = (\epsilon V_t^k, 0)$$

is bounded by ϵ. Thus, classical criteria of relative compactness in the Skorohod space apply, and a sub-sequence also denoted $(\tau_k)_{k\geq 0} 0$ can be extracted such that the associated path $t \mapsto (X_t^k, V_t^k)$ converges in the Skorohod sense towards the process $t \mapsto (X_t^\infty, V_t^\infty)$.

Now applying Duhamel's formula in (2.11) yields:

$$Z_t^k = \mathrm{e}^{-\tau_k^{-1} t} Z_0 + \int_0^t \mathrm{e}^{-\tau_k^{-1}(t-s)} \left(\epsilon \nabla S(X_s^k) V_s^k + \mathcal{O}\left(\left|Z_s^k\right|^2\right)\right) ds; \tag{A.2}$$

note that $\sup_t |Z_t|$ is bounded uniformly in τ_k on any time interval $[0,T]$. (A.2) yields:

$$\left|\tau_k^{-1}\left(Z_t^k - \mathrm{e}^{-\tau_k^{-1} t} Z_0\right)\right| \leq K \int_0^t \left\|\tau_k^{-1}\right\| \mathrm{e}^{-r(\tau_k^{-1})(t-s)} \left(1 + \mathcal{O}\left(\left|Z_s^k\right|^2\right)\right) ds,$$

so that for any $t > 0$, $\lim_{k\to+\infty} Z_t^k = 0$. Now, using again (A.1), this gives :

$$\lim_{k\to+\infty} \int_0^t \tau_k^{-1} \mathrm{e}^{-\tau_k^{-1}(t-s)} \epsilon \nabla S(X_s^k) V_s^k\, ds = \epsilon \nabla S(X_t^\infty) V_t^\infty,$$

for any time t different from the jump times of the limit process $t \mapsto (X_t^\infty, V_t^\infty)$. As a consequence, using dominated convergence in (A.2):

$$\tau_k^{-1} Z_t^k \underset{k\to+\infty}{\to} \epsilon \nabla S(X_t^\infty) V_t^\infty,$$

and the turning rate of the process in (2.13) satisfies:

$$\lambda(Z_t^k) \underset{k\to+\infty}{\to} \lambda_0 - \epsilon \bar{b}^T \nabla S(X_t^\infty) V_t^\infty.$$

and the limit process verify (2.19) by applying the dominated convergence theorem to the integrals $\int_{T_n^k}^{T_{n+1}^k} \lambda(Z_t^k)\mathrm{d}t$ where $(T_n^k)_{n\geq 0}$ are the jump times of the process indexed by k. □

Appendix B. Proof of Lemma 4.2. Let us make the proof first in the case of a linear turning rate (4.1). Using the assumption $T_n = T_n^{\delta t}$ and $X_{T_n} = X_{T_n}^{\delta t}$, we get straightforwardly:

$$X_t = X_t^{\delta t}, \quad \forall t \in [T_n, T_{n+1} \wedge T_{n+1}^{\delta t}].$$

Consider Duhamel's formula 3.13, and write similarly for the time discretized case for $t \in [T_n, T_{n+1}^{\delta t}]$:

$$Z_t^{\delta t} = \mathrm{e}^{-(t-T_n)\tau^{-1}} Z_{T_n}^{\delta t} + \epsilon \int_{T_n}^{t} \mathrm{e}^{-(t-t')\tau^{-1}} \nabla S(X^{\delta t}_{n,\lfloor \frac{t_{n,0}-t'}{\delta t} \rfloor}) \mathcal{V}_n \mathrm{d}t'. \tag{B.1}$$

The difference between the two Duhamel's integrals then yields for $t \in [T_n, T_{n+1} \wedge T_{n+1}^{\delta t}]$, using the assumption $Z_{T_n}^{\delta t} = Z_{T_n}$:

$$\left| Z_t - Z_t^{\delta t} \right| \leq \frac{\epsilon^2}{r(\tau^{-1})} \delta t \left\| \mathrm{Hess} S \right\|_\infty .$$

We now proceed to find a bound on $\left| T_{n+1} - T_{n+1}^{\delta t} \right|$. We start by writing

$$\int_{T_n}^{T_{n+1}} \lambda^l(Z_t)\mathrm{d}t = \int_{T_n^{\delta t}}^{T_{n+1}^{\delta t}} \lambda^l(Z_t^{\delta t})\mathrm{d}t = -\log U_n;$$

and using the assumption $T_n = T_n^{\delta t}$, we obtain then:

$$\begin{aligned} -\int_{T_n}^{T_{n+1} \wedge T_{n+1}^{\delta t}} b^T \left(Z_t - Z_t^{\delta t} \right) \mathrm{d}t &= \int_{T_{n+1} \wedge T_{n+1}^{\delta t}}^{T_{n+1}^{\delta t}} \lambda^l(Z_t^{\delta t})\mathrm{d}t - \int_{T_{n+1} \wedge T_{n+1}^{\delta t}}^{T_{n+1}} \lambda^l(Z_t)\mathrm{d}t \\ &= \left(T_{n+1}^{\delta t} - T_{n+1} \wedge T_{n+1}^{\delta t} \right) \lambda^l(Z_{t_1^*}^{\delta t}) + \left(T_{n+1} - T_{n+1} \wedge T_{n+1}^{\delta t} \right) \lambda^l(Z_{t_2^*}^{\delta t}). \end{aligned}$$

for some values of $t_1^* \in [T_{n+1} \wedge T_{n+1}^{\delta t}, T_{n+1}^{\delta t}]$ and $t_2^* \in [T_{n+1} \wedge T_{n+1}^{\delta t}, T_n^{\delta t}]$. Since $\lambda_{\min} \leq \lambda^l(z) \leq \lambda_{\max}$, we have

$$\begin{aligned} \left| T_{n+1}^{\delta t} - T_{n+1} \right| &\leq K \int_{T_n}^{T_{n+1} \wedge T_{n+1}^{\delta t}} \left| Z_t - Z_t^{\delta t} \right| \mathrm{d}t \\ &\leq K \theta_{n+1} \frac{\epsilon^2}{r(\tau^{-1})} \delta t \left\| \mathrm{Hess} S \right\|_\infty . \end{aligned}$$

Finally, note that

$$\left| X_t - X_t^{\delta t} \right| \leq \epsilon (t - T_n)$$

for $t \geq T_n$ to get the estimate on $\left| X_{T_{n+1}} - X^{\delta t}_{T_{n+1}^{\delta t}} \right|$. This concludes the proof for the linear turning rate case (4.1).

Appendix C. Proof of Lemma 4.3. We now turn to the case of a non-linear turning rate (2.8) discretized by (4.12). Similarly to the linear case we can show that:

$$\left| T_{n+1} - T_{n+1}^{\delta t} \right| \leq K \int_{T_n}^{T_{n+1} \wedge T_{n+1}^{\delta t}} \left| \lambda(Z_t) - \lambda^{\delta t}(Z_t^{\delta t}, Z^{\delta t}_{n,\lfloor \frac{t_{n,0}-t}{\delta t} \rfloor}) \right| \mathrm{d}t.$$

We then consider the following bound for $t \in [t_{n,k}, t_{n,k+1}]$

$$\left|\lambda(Z_t^{\delta t}) - \lambda^{\delta t}(Z_t^{\delta t}, Z_{n,k}^{\delta t})\right| \leq \frac{\|\mathrm{Hess}\,\lambda\|_\infty}{2} \left(Z_t^{\delta t} - Z_{n,k}^{\delta t}\right)^2.$$

Then remarking

$$Z_t^{\delta t} = \mathrm{e}^{-(t-t_{n,k})\tau^{-1}} Z_{n,k}^{\delta t} + \epsilon\tau \left(\mathrm{Id} - \mathrm{e}^{\left(-(t-t_{n,k})\tau^{-1}\right)}\right) \nabla S(X_{n,k}^{\delta t}),$$

so that

$$Z_t^{\delta t} - Z_{n,k}^{\delta t} = -\left(\mathrm{Id} - \exp\left(-\frac{t-t_{n,k}}{\tau}\right)\right) \left(Z_{n,k}^{\delta t} - \tau\epsilon\nabla S(X_{n,k}^{\delta t})\right),$$

from which we obtain

$$\begin{aligned}\int_{t_{n,k}}^{t_{n,k+1}} \left(Z_t^{\delta t} - Z_{n,k}^{\delta t}\right)^2 \mathrm{d}t &\leq K\delta t^3 \left(\left\|\tau^{-1}\right\| Z_{n,k}^{\delta t} + \epsilon\right)^2 \\ &\leq K\delta t^3 \frac{\epsilon^2 \left\|\tau^{-1}\right\|^2}{r(\tau^{-1})^2}.\end{aligned}$$

In the last line of the above, we have used as in the non discretized case the estimate

$$\left|Z_t^{\delta t}\right| \leq K\left(\left|Z_{T_n}^{\delta t}\right| + \frac{\epsilon}{r(\tau^{-1})}\right) \leq K\frac{\epsilon}{r(\tau^{-1})}$$

which follows from Duhamel's formula (B.1) and the assumption $\left|Z_{T_n}^{\delta t}\right| = \mathcal{O}(\frac{eps}{r(\tau^{-1})})$.

On the other hand, using Duhamel's formula as in the proof of Lemma 4.2 above yields

$$\left|\lambda(Z_t) - \lambda(Z_t^{\delta t})\right| \leq K\left|Z_t - Z_t^{\delta t}\right| \leq K\delta t\frac{\epsilon^2}{r(\tau^{-1})},$$

Finally we get:

$$\left|T_{n+1} - T_{n+1}^{\delta t}\right| \leq K\theta_{n+1}\delta t\frac{\epsilon^2}{r(\tau^{-1})} + K\theta_{n+1}\delta t^3\frac{\epsilon^2\left\|\tau^{-1}\right\|^2}{r(\tau^{-1})^2}.$$

This concludes the proof.