Individual-based models for bacterial chemotaxis in the diffusion asymptotics

Mathias Rousset, Giovanni Samaey

▶ To cite this version:

Individual-based models for bacterial chemotaxis in the diffusion aymptotics

Mathias Rousset[*] and Giovanni Samaey[†]

We discuss velocity-jump models for chemotaxis of bacteria with an internal state that allows the velocity jump rate to depend on the memory of the chemoattractant concentration along their path of motion. Using probabilistic techniques, we provide a pathwise result that shows that the considered process converges to an advection-diffusion process in the (long-time) diffusion limit. We also (re-)prove using the same approach that the same limiting equation arises for a related, simpler process with direct sensing of the chemoattractant gradient. Additionally, we propose a time discretization technique that retains these diffusion limits exactly, i.e., without error that depends on the time discretization. In the companion paper [21], these results are used to construct a coupling technique that allows numerical simulation of the process with internal state with asymptotic variance reduction, in the sense that the variance vanishes in the diffusion limit.

MSC: 35Q80, 92B05, 65C35.

Key words: bacterial chemotaxis, velocity-jump process, diffusion approximation.

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 takes much less time than the run phase and acts as a reorientation. To bias movement towards regions with high concentration of chemoattractant, the bacterium adjusts its turning rate to increase, resp. decrease, the probability of tumbling when moving in an unfavorable, resp. favorable, direction [2,23]. Since many species are unable to sense chemoattractant gradients reliably due to their small size, this adjustment is often done via an intracellular mechanism that allows the bacterium to retain information on the history of the chemoattractant concentrations along its path [4]. The resulting model can be formulated as a velocity-jump process, combined with an ordinary differential equation (ODE) that describes the evolution of an internal state that incorporates this memory effect [7,8]. This model will be called the "fine-scale" or "internal dynamics" model in this paper.

The probability density of the velocity-jump process evolves according to a kinetic equation, in which, besides position and velocity, the internal variables appear

[*]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).

as additional dimensions. A direct deterministic simulation of the density distribution of all the variables of the model is therefore prohibitively expensive. Hence, it is of interest to study the relation of this model with simplified, coarse-level descriptions of the bacteria dynamics.

In chemotaxis, a first coarse description is obtained by neglecting velocity and internal variables and considering the bacterial position density on large space and time scales. One then postulates an advection-diffusion equation for this bacterial position density, in which a chemotactic sensitivity coefficient incorporates the effect of chemoattractant concentrations gradients on the density fluxes. This assumption leads to the classical Keller–Segel equations (see [15], and [13,14] for numerous historical references). Several works have also considered the motion of a bacterium to be governed by a velocity-jump process corresponding to a kinetic description for the phase-space density [1,17,20]. In these models, the velocity jump rate of the bacteria depends on the local chemoattractant gradient. In the present paper, the model corresponding to such bacteria will be called a "coarse" or "direct gradient sensing" model. Unlike the limiting Keller–Segel equation, see e.g. [5,11,19], 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 [6]. These models also converge to a Keller–Segel equation in the appropriate drift-diffusion limit, see e.g. [6,12,18]. In [8], the kinetic equation associated to a model with internal state (similar to the model considered here) has also be shown to formally converge to a Keller–Segel equation using moment closure assumptions and an appropriate (diffusion) scaling; a similar result using an infinite moment expansion has been obtained in [26]. Then, the parameters of internal dynamics appear in the expression for the chemotactic sensitivity [7,8]. Existence results when the model is coupled to a mean-field production of chemo-attractants are also available [3], and analysis of the long-time behavior is carried out in [10]. Moreover, some recent studies have assessed the biological relevance of such a model through experimental validation of travelling pulses [22].

The present paper contributes to the analysis of velocity-jump processes for chemotaxis of bacteria with internal dynamics, and their numerical discretization, in the diffusive asymptotics. The main contributions are twofold :

- From an analysis point of view, we consider (fine-scale) velocity-jump processes with internal dynamics (2.5), with jump rates satisfying (2.4) and internal dynamics satisfying the two main assumptions (3.2) and (3.3) as well as other technical assumptions contained in Section 3.3. We rigorously prove, using probabilistic arguments, the *pathwise* convergence of the velocity-jump process described above to the stochastic differential equation (SDE) (4.1) in the long time limit (see Proposition 4.11). The technical steps of the proof are given in the beginning of Section 4. We also recall a convergence result for the (coarse) process with direct gradient sensing, see Proposition 4.3. For an appropriate choice of the model parameters, both

limits can be made identical.
- From a numerical point of view, we introduce a time-discretized model (see Section 5) for both the fine-scale and the coarse velocity-jump processes. When discretizing the fine-scale process with internal dynamics, particular attention is paid to the time discretization errors due to the approximate solution of the ODE describing the evolution of the internal state. As a result, the asymptotic analysis of the time-continuous case still holds in an exact fashion, ensuring that the advection-diffusion limits are preserved after time discretization.

Due to the possibly high number of dimensions of the kinetic model with internal state, the evolution of the bacterial density away from the diffusive limit may rather be simulated using a stochastic particle method. However, a direct stochastic particle-based simulation suffers from a large statistical variance, even in the diffusive asymptotic regime, raising the important issue of variance reduced simulation. In the companion paper [21], we will use the results of the analysis presented here to construct a numerical scheme that couples the fine-scale model for bacteria with internal dynamics and the simpler, coarse model for bacteria with direct gradient sensing to obtain a variance-reduced simulation of the fine-scale process. We show that the variance reduction is *asymptotic*, in the sense that the statistical variance scales as the error between the fine-scale and coarse description.

This paper is organized as follows. In Section 2, we discuss the velocity-jump processes that will be considered, introducing both the fine-scale process with internal state and a simpler, coarse process in which the internal state has been replaced by direct gradient sensing. We discuss the relevant asymptotic regimes in Section 3, and analyze the limits of the considered processes in these regimes in Section 4. Section 5 discusses a discretization scheme for the velocity-jump processes that retains the diffusion limit of the continuum processes. Section 6 contains conclusions and a more detailed discussion of the relation with the companion paper [21].

2. Particle-based models for bacterial chemotaxis {sec:particle}

2.1. *Model with internal state* {sec:model-internal}

We consider bacteria that are sensitive to the concentration of m chemoattractants $(\rho_i(x))_{i=1}^m$, with $\rho_i(x) \geq 0$ for $x \in \mathbb{R}^d$. While we do not consider time dependence of chemoattractants via production or consumption 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 [7], we 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. The internal state models the memory of the bacterium and is subject to an evolution mechanism attracted by a function $\psi : \mathbb{R}^m \to \mathbb{R}^n$ of the

chemoattractants concentrations,

$$S(x) := \psi(\rho_1(x), \ldots, \rho_m(x)) \in \mathbb{R}^n,$$

where x is the present position of the bacterium. A typical choice is $2m = n$, $\psi(\rho_1, \ldots, \rho_m) = (\rho_1, \ldots, \rho_m, 0, \ldots, 0)$, and the internal mechanism being such that the m first internal variables are memorizing the concentrations of chemoattractants, while the m last internal variables are computing the variations of concentrations along a bacterium trajectory (see Section 2.2). S and its two first derivatives are assumed to be smooth and bounded.

We use (l_s, t_λ) as a reference for length and time, where

- l_s is the typical length of the chemoattractant concentration variations (which we assume similar for all the different species of chemoattractants).
- t_λ is the typical time between two changes of the bacterium velocity direction (tumbling).

We also introduce the time t_x that indicates the typical time on which a bacterium observes changes in chemoattractant concentration; the typical speed of a bacterium is thus l_s/t_x. Anticipating the time-scale analysis in Section 3, we already introduce a small parameter

{eq:eps}
$$\epsilon := \frac{t_\lambda}{t_x} \ll 1, \tag{2.1}$$

indicating that we assume the typical time between two velocity changes to be much smaller than the typical time on which we observe macroscopic movement of the bacteria.

The model is now presented directly in dimensionless form. The evolution of each bacterium individual position is denoted as

$$t \mapsto X_t,$$

with velocity

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

where $\mathbb{S}^{d-1}$ represents 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 via an ordinary differential equation (ODE),

{eq:internal}
$$\frac{\mathrm{d}Y_t}{\mathrm{d}t} = F_\epsilon(Y_t, S(X_t)), \tag{2.2}$$

which is required to have a unique fixed point $y^* = S(x^*)$ for every fixed value $x^* \in \mathbb{R}^d$. We also introduce the deviations from equilibrium $z = S(x) - y$. The evolution of these deviations is denoted as

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

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

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

We will denote its expansion for small z as

$$\lambda(z) = \lambda_0 - b^T z + c_\lambda \mathcal{O}\left(|z|^k\right), \tag{2.4}$$ {eq:lin_rate}

with $k \geq 2$, $b \in \mathbb{R}^n$; the constant c_λ is used to keep track of the nonlinearity of λ in the analysis.

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:

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = \epsilon V_t \\[2mm] \dfrac{\mathrm{d}Y_t}{\mathrm{d}t} = F_\epsilon(Y_t, S(X_t)) \\[2mm] \displaystyle\int_{T_n}^{T_{n+1}} \lambda(Z_t)\mathrm{d}t = \theta_{n+1}, \qquad \text{with } Z_t := S(X_t) - Y_t \\[2mm] V_t = \mathcal{V}_n \quad \text{for } t \in [T_n, T_{n+1}] \ , \end{cases} \tag{2.5}$$ {eq:process_noscale}

with initial condition $X_0, V_0 \in \mathbb{R}^d$, $Y_0 \in \mathbb{R}^n$ and $T_0 = 0$. In (2.5), $(\theta_n)_{n\geq 1}$ denote i.i.d. random variables with normalized exponential distribution, and $(\mathcal{V}_n)_{n\geq 1}$ 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 (2.5) is equivalent to the following system:

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = \epsilon V_t \\[2mm] \dfrac{\mathrm{d}Y_t}{\mathrm{d}t} = F_\epsilon(X_t, S(Y_t)) \\[2mm] \displaystyle\int_{T_n}^{T_{n+1}} \lambda(Z_t)\mathrm{d}t = 2\theta_{n+1}, \\[2mm] V_t = \mathcal{V}_n \quad \text{for } t \in [T_n, T_{n+1}] \text{ with } \mathcal{V}_n = -\mathcal{V}_{n-1}, \end{cases} \tag{2.6}$$ {eq:process1d}

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 (3.5) and discussion below. (Note that, in this process, the velocity is always reversed. This is compensated by a factor 2 in the equation defining the jump times.)

2.2. *Example*

{ssec:example}

For concreteness, we provide a specific example, adapted from [7]. We consider $m = 1$, i.e. there is only one chemoattractant $\rho_1(x)$. 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:

{e:exc-ad}
$$\begin{cases} \dfrac{\mathrm{d}y_1(t)}{\mathrm{d}t} = \dfrac{\rho_1(x) - y_1(t)}{t_a}, \\ \dfrac{\mathrm{d}y_2(t)}{\mathrm{d}t} = \dfrac{-\rho_1(x) + y_1(t) - y_2(t)}{t_e}, \end{cases} \tag{2.7}$$

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 $y^* = (y_1^*, y_2^*) = (\rho_1(x), 0) =: S(x)$, and the deviation variables are given by $z = (z_1, z_2) = S(x) - y = (\rho_1(x) - y_1, -y_2)$. When the excitation time is much smaller than the adaptation time, $t_e \ll t_a$, the variable y_1 therefore adapts "slowly" to the environment and memorizes it, while y_2 computes "faster" the lag $y_1 - \rho_1(x)$, giving the response to changing environments. In the asymptotics $t_e/t_a \to 0$, the internal dynamics reduce to a scalar equation

{e:scalar-y}
$$\frac{\mathrm{d}y_1(t)}{\mathrm{d}t} = \frac{\rho_1(x) - y_1}{t_a}, \tag{2.8}$$

while the exciting variable reduces to the difference $y_2 = y_1 - \rho_1(x) = -z_1$ which is instantly learned by the bacteria. In this limit, we may eliminate the second internal state variable, and write $y \equiv y_1$ and $S(x) := \rho_1(x)$, and $z = S(x) - y$.

For the turning rate $z \mapsto \lambda(z)$, we choose a nonlinear strictly decreasing smooth function, depending on a scalar ζ that aims at measuring the difference between the chemoattractant concentration and the memory variable. When considering the two-dimensional internal dynamics, we choose $\zeta = z_2 = -y_2$, while, in the scalar case $\zeta = z$. As a consequence, $\zeta > 0$, resp. $\zeta < 0$, give the information to the bacterium that it is moving in a favorable, resp. unfavorable, direction. A specific choice of $\lambda(\zeta)$ is

{eq:rate}
$$\lambda(\zeta) = 2\lambda_0 \left(\frac{1}{2} - \frac{1}{\pi} \arctan\left(\frac{\pi}{2\lambda_0} \beta\zeta \right) \right), \qquad \beta > 0. \tag{2.9}$$

Note that this turning rate satisfies

$$\lambda(z) = \lambda_0 - \beta\zeta + \mathcal{O}\left(|\zeta|^3 \right).$$

{sec:modelgrad}

2.3. *Model with direct gradient sensing*

We now turn to a simplified model, in which the internal process (2.2), and the corresponding state variables, are eliminated. Instead, the turning rate depends directly on the chemoattractant gradient. The process with direct gradient sensing is a Markov process in position-velocity variables

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

which evolve according to the differential velocity-jump equations

$$\begin{cases} \dfrac{\mathrm{d}X_t^c}{\mathrm{d}t} = \epsilon V_t^c \\[2mm] \displaystyle\int_{T_n^c}^{T_{n+1}^c} \lambda_\epsilon^c(X_t^c, V_t^c)\mathrm{d}t = \theta_{n+1} \\[2mm] V_t^c = \mathcal{V}_n \quad \text{for } t \in [T_n^c, T_{n+1}^c]\,, \end{cases} \qquad (2.10) \quad \text{\{eq:cprocess\}}$$

with initial condition $X_0, \mathcal{V}_0 \in \mathbb{R}^d$. In (2.10), $(\theta_n)_{n \geq 1}$ denote i.i.d. random variables with normalized exponential distribution, and $(\mathcal{V}_n)_{n \geq 1}$ denote i.i.d. random variables with distribution $\mathcal{M}(dv)$. The turning rate of the process with direct gradient sensing is assumed to be of the form

$$\lambda_\epsilon^c(x,v) := \lambda_0 - \epsilon\, A_\epsilon^T(x)v + \mathcal{O}\left(\epsilon^2\right), \qquad (2.11) \quad \text{\{eq:control\_rate\}}$$

and to satisfy

$$0 < \lambda_{\min} \leq \lambda(x,v) \leq \lambda_{\max}. \qquad (2.12) \quad \text{\{eq:rateboundc\}}$$

The vector field $A_\epsilon(x) \in \mathbb{R}^d$ is usually a linear combination of the columns of $\nabla S(x) \in \mathbb{R}^{d \times n}$, and, as a consequence on the smoothness assumption on $S(x)$, satisfies that, for $\epsilon \to 0$, both A_ϵ and its derivatives approach a limiting vector field A_0 in uniform norm.

The model (2.11) may describe a large bacterium that is able to directly sense chemoattractant gradients. In the case $n = 1$, the turning rate (2.11) 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.

3. Asymptotic regimes and kinetic formulation

{sec:asympt}

3.1. *Notation*

Throughout the text, the Landau symbol $\mathcal{O}$ denotes a *deterministic* and globally Lipschitz function (with $\mathcal{O}(0) = 0$). Its precise value *may vary* from line to line, and may depend on the parameters of the model. However, its Lipschitz constant is uniform in ϵ. In the same way, we will denote generically by $C > 0$ a deterministic constant that may depend on all the parameters of the model except ϵ.

8   *Individual-based model of chemotaxis*

3.2. *Definition of time scales*

{sec:asympt_time}

Recall that we have introduced two time scales, characterized by (i) the typical time t_x on which a bacterium observes changes in chemoattractant concentration; and (ii) the typical time t_λ between two changes of the bacterium velocity direction (tumbling). We have assumed a large time-scale separation between these two time scales by introducing the small parameter $\epsilon = t_\lambda/t_x \ll 1$.

Let us now turn to the typical time t_y associated with the dynamics of the internal state. In principle, several typical time scales ($t_{1,y} \gg t_{2,y} \gg \ldots$) may be required to describe the evolution of the internal state; $t_y := t_{1,y}$ is assumed to be the largest one. The vector field driving the ODE (2.2) of the internal state is thus of order

$$F_\epsilon \sim \frac{t_\lambda}{t_y},$$

and we will assume that

{eq:F}
$$F_\epsilon \sim \frac{t_\lambda}{t_y} = \epsilon^{1-\delta}, \tag{3.1}$$

with $\delta > 1/k$, where k is defined by (2.4).

Consequently, we will consider three time scales :

- A fast time scale of order $\mathcal{O}(1)$ given by the rate of change of the velocity direction.
- A time scale of order $\mathcal{O}(1/\epsilon^{1-\delta})$ with $\delta > 1/k$ given by the internal state evolution; depending on δ, this time scale can be very fast, fast or intermediate.
- A slow time scale $\mathcal{O}(1/\epsilon)$ given by the evolution of the chemoattractant concentration as seen from the bacteria.

3.3. *Asymptotic assumptions*

{sec:ass_ass}

The loose assumption on the scale, equation (3.1) for the internal dynamics (2.2), needs to be made precise by the following set of assumptions. First, we assume that the ODE (2.2) driving the internal state is well approximated by a near-equilibrium evolution equation :

{a:1} Assumption 3.1. We have

{eq:tau_eps}
$$F_\epsilon(y,s) = -\tau_\epsilon^{-1}(y-s) + \epsilon^{1-\delta} c_F \mathcal{O}(|s-y|^2), \tag{3.2}$$

where $\tau_\epsilon \in \mathbb{R}^{n \times n}$ is an invertible constant matrix. The constant c_F is used to keep track of the dependence on the non-linearity of F_ϵ in the analysis.

We now formulate the necessary assumptions on the scale of the matrix τ_ϵ :

{a:2} Assumption 3.2. There is a constant $C > 0$ such that for any $t \geq 0$, one has

{eq:tau_scales}
$$\left\| e^{-t\tau_\epsilon^{-1}} \right\| \leq C e^{-t\epsilon^{1-\delta}/C}. \tag{3.3}$$

Assumption 3.2 is used to ensure exponential convergence of the linear ODE with time scale at least of order $\epsilon^{1-\delta}$. When τ_ϵ^{-1} is symmetric, Assumption 3.2 follows from the assumption that the eigenvalues of τ_ϵ^{-1} are strictly positive, with a lower bound of order $\mathcal{O}(\epsilon^{1-\delta})$.

Assumption 3.3. There is a constant $C > 0$ such that

$$\sup_{t \geq 0} \left\| t\tau_\epsilon^{-1} \mathrm{e}^{-t\tau_\epsilon^{-1}} \right\| \leq C\epsilon^{-1}.$$

{a:3}

Assumption 3.3 is necessary for technical reasons in Lemma 4.8. When τ_ϵ^{-1} is symmetric, then Assumption 3.3 follows from the assumption that the eigenvalues of τ_ϵ^{-1} are non-negative.

Finally, we assume that the solution of the ODE driving the internal state in (2.5) satisfies the following long time behavior :

Assumption 3.4. Consider a process $t \mapsto S_t$ such that $\sup_{t \in [0,+\infty]} |dS_t/dt| = \mathcal{O}(\epsilon)$, and assume $|S_0 - Y_0| = \mathcal{O}(\epsilon^\delta)$. Then the solution of

{ass:ODE}

$$\frac{dY_t}{dt} = F_\epsilon(Y_t, S_t),$$

satisfies

$$\sup_{t \in [0,+\infty]} |Y_t - S_t| = \mathcal{O}\left(\epsilon^\delta\right).$$

Assumption 3.4 is motivated by the case of linear ODEs; indeed, in that case, Assumption 3.4 is a consequence of Assumption 3.2, as is seen from the following Lemma :

Lemma 3.5. *Assume that*

$$F_\epsilon(y, s) = -\tau_\epsilon^{-1}(y - s),$$

then Assumption 3.2 implies Assumption 3.4.

Proof. Denoting $Z_t = S_t - Y_t$ in Assumption 3.4, Duhamel's integration formula yields

$$Z_t = \mathrm{e}^{-t\tau_\epsilon^{-1}} Z_0 + \int_0^t \mathrm{e}^{-(t-t')\tau_\epsilon^{-1}} \frac{dS_{t'}}{dt'} \mathrm{d}t',$$

so that by Assumption 3.2, we have

$$|Z_t| \leq |Z_0| + C\epsilon \int_0^t \mathrm{e}^{-(t-t')\epsilon^{1-\delta}/C} \mathrm{d}t'.$$
$$\leq |Z_0| + C\epsilon \times \epsilon^{\delta-1}\left(1 - \mathrm{e}^{-t\epsilon^{1-\delta}/C}\right),$$

and Assumption 3.4 follows. $\qquad\square$

In a similar fashion, Assumption 3.3 implies the following technical bound :

Lemma 3.6. *Consider a process $t \mapsto S_t$ such that $\sup_{t \in [0, +\infty]} |dS_t/dt| = \mathcal{O}(\epsilon)$. Assume that $|S_0 - Y_0| = \mathcal{O}(1)$, and that Assumption 3.1, Assumption 3.2, Assumption 3.3 and Assumption 3.4 hold. Then, we have*   {lem:techbound}

$$\sup_{t \in [0, +\infty]} \left| \tau_\epsilon^{-1} \mathrm{e}^{-\tau_\epsilon^{-1} t} (S_t - Y_t) \right| = \mathcal{O}(1). \tag{3.4}$$ {eq:techbound}

Proof. Let us denote $A_t := \tau_\epsilon^{-1} \mathrm{e}^{-\tau_\epsilon^{-1} t} (S_t - Y_t)$, which satisfies

$$\frac{dA_t}{dt} = -2\tau_\epsilon^{-1} A_t + \tau_\epsilon^{-1} \mathrm{e}^{-\tau_\epsilon^{-1} t} \left(\frac{dS_t}{dt} - F_\epsilon(Y_t, S_t) + \tau_\epsilon^{-1} (S_t - Y_t) \right).$$

Here again, Duhamel integration yields

$$A_t = \mathrm{e}^{-t 2\tau_\epsilon^{-1}} A_0 + \int_0^t \mathrm{e}^{-(t-t') 2\tau_\epsilon^{-1}} \tau_\epsilon^{-1} \mathrm{e}^{-\tau_\epsilon^{-1} t'} \left(\frac{dS_{t'}}{dt'} - F_\epsilon(Y_{t'}, S_{t'}) + \tau_\epsilon^{-1} (S_{t'} - Y_{t'}) \right) \mathrm{d}t'.$$

From Assumption 3.1 and Assumption 3.4, $\left| F_\epsilon(Y_{t'}, S_{t'}) - \tau_\epsilon^{-1}(S_{t'} - Y_{t'}) \right| = \mathcal{O}(\epsilon^{1+\delta}) \leq \mathcal{O}(\epsilon)$ so that

$$|A_t| \leq A_0 + \left\| \tau_\epsilon^{-1} \mathrm{e}^{-\tau_\epsilon^{-1} t} \right\| \int_0^t \left\| \mathrm{e}^{-(t-t') \tau_\epsilon^{-1}} \right\| \mathcal{O}(\epsilon) \mathrm{d}t'$$

$$\leq A_0 + \left\| t \tau_\epsilon^{-1} \mathrm{e}^{-\tau_\epsilon^{-1} t} \right\| \frac{1}{t \epsilon^{1-\delta}} (1 - \mathrm{e}^{-t \epsilon^{1-\delta}/C}) \mathcal{O}(\epsilon)$$

$$\leq \mathcal{O}(1) + \mathcal{O}(\epsilon^{-1}) \mathcal{O}(\epsilon) = \mathcal{O}(1),$$

where we have used in the above Assumption 3.2 and Assumption 3.3.   □

In the remainder of the paper, we will assume that the solution of (2.5) satisfies

$$|Z_0| = \mathcal{O}(\epsilon^\delta),$$

as well as:

$$\left| \tau_\epsilon^{-1} Z_0 \right| = \mathcal{O}(1),$$

which implies, according to Assumption 3.4 and Lemma 3.6, that

$$\sup_{t \in [0, +\infty]} (|Z_t|) = \mathcal{O}(\epsilon^\delta),$$

as well as

$$\sup_{t \in [0, +\infty]} \left(\left| \tau_\epsilon^{-1} \mathrm{e}^{-\tau_\epsilon^{-1} t} Z_t \right| \right) = \mathcal{O}(1).$$

Note that all the assumptions of the present section hold in the following case: (i) the ODE is linear (F_ϵ is linear), (ii) the associated matrix τ_ϵ is symmetric positive and satisfies

$$\tau_\epsilon^{-1} \geq C \epsilon^{1-\delta},$$

for some $\delta > 0$, (iii) the initial internal state is close to equilibrium in the sense that $\tau_\epsilon^{-1} Z_0 = \mathcal{O}(1)$. The assumptions of the present section can be seen as technical generalizations to non-linear ODEs with non-symmetric linear part.

3.4. *Kinetic formulations*

The probability distribution density of the process with internal dynamics 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

$$\partial_t p + \epsilon v \cdot \nabla_x p + \mathrm{div}_y\left(F_\epsilon(y,S(x))p\right) = \lambda\left(S(x)-y\right)\left(R(p)-p\right), \qquad (3.5) \quad \texttt{\{e:kinetic\}}$$

where

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

is the operator integrating velocities with respect to $\mathcal{M}$. Similarly, the distribution density of the process with direct gradient sensing 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

$$\partial_t p^c + \epsilon v \cdot \nabla_x p^c = \left(R(\lambda_\epsilon^c(x,v)p) - \lambda_\epsilon^c(x,v)p\right). \qquad (3.6) \quad \texttt{\{e:kinetic\_control\}}$$

We refer to [9] for the derivation of master equations associated to Markov jump processes, and to [12,7,26,6,18] for an asymptotic analysis using moment closures.

4. Asymptotic analysis of individual-based models

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ `{sec:adv-diff}`

In this section, we investigate in more detail the asymptotic behavior of the process with internal dynamics and the simplified process with direct gradient sensing. We rigorously prove pathwise convergence to an advection-diffusion limit of both the process with direct gradient sensing (Section 4.1) and the process with internal dynamics (Section 4.2). Choosing the model parameters appropriately, both advection-diffusion limits can be made identical. The two proofs are based on an asymptotic expansion of the jump times (the time between two tumble phases), and a comparison with a simpler random walk. They can be split into 4 steps.

Step (i) Consider the evolution of the bacterium indexed by the number of tumble phases (jump times), as well as a proper Markovian random walk approximating the former. Show tightness of the coupled sequence (in fact their time continuous version obtained by linear interpolation).

Step (ii) Show, using a Gronwall argument, that the two processes have the same advection-diffusion limit. This step relies on an asympotic expansion of the jump times in ϵ.

Step (iii) Compute the advection-diffusion limit of the Markovian random walk.

Step (iv) Use a random time change to transfer the result to the bacteria position process itself.

4.1. *Diffusion limit of the process with direct gradient sensing*

{sec:limit-control}

We first derive a pathwise advection-diffusion limit of the process with direct gradient sensing (2.10) when $\epsilon \to 0$ using classical probabilistic arguments. It should be noted that this diffusive limit as already been extensively studied in the context of bacterial chemotaxis, see, e.g.,[12] for justifications of Hilbert expansions at the PDE level, and [24] where the diffusive limit of some more general class of stochastic processes is suggested. However, we give the probabilistic arguments *in extenso* since they will be of interest for the analysis of both the model with internal state, and of the coupling in the variance-reduced numerical scheme, discussed in the companion paper [21]. Moreover, they do not seem to appear as such in the literature.

Diffusive times are denoted by

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

and the process with direct gradient sensing considered on diffusive time scales by

$$X_{\bar{t}}^{c,\epsilon} := X_{\bar{t}/\epsilon^2}^c.$$

This section shows that this process converges for $\epsilon \to 0$ towards an advection-diffusion process, satisfying the stochastic differential equation (SDE)

{eq:cprocess_hydro}
$$\mathrm{d}X_{\bar{t}}^{c,0} = \left(\frac{DA_0(X_{\bar{t}}^{c,0})}{\lambda_0} \mathrm{d}\bar{t} + \left(\frac{2D}{\lambda_0} \right)^{1/2} \mathrm{d}W_{\bar{t}} \right), \tag{4.1}$$

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:

{eq:maxw}
$$D = \int_{\mathbb{S}^{d-1}} v \otimes v \, \mathcal{M}(dv) \in \mathbb{R}^{d \times d}. \tag{4.2}$$

In particular, this result implies at the level of the Kolomogorov/master evolution equation, that the position bacterial density

{eq:cdensity}
$$n^{c,\epsilon}(x,\bar{t}) := n^c(x,\bar{t}/\epsilon^2) := \int_{\mathbb{V}} p^c(x,v,\bar{t}/\epsilon^2) \, \mathcal{M}(dv) \tag{4.3}$$

satisfies the advection-diffusion equation

{eq:cdens_hydro}
$$\partial_{\bar{t}} n^{c,0} = \frac{1}{\lambda_0} \mathrm{div}_x \left(D\nabla_x n^{c,0} - DA_0(x) n^{c,0} \right) \tag{4.4}$$

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

The proof relies on a perturbation analysis of the jump time differences, which is given by the following lemma :

{lem:dt-control} **Lemma 4.1.** *The difference between two jump times of the process with direct gradient sensing satisfies*

{eq:deltaTc}
$$\Delta T_{n+1}^c := T_{n+1}^c - T_n^c = \frac{\theta_{n+1}}{\lambda_0} + \epsilon \frac{\theta_{n+1}}{\lambda_0^2} A_\epsilon^T (X_{T_n^c}^c) \mathcal{V}_n + \theta_{n+1} \mathcal{O}(\epsilon^2). \tag{4.5}$$

Proof. Using (2.11), we get

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

where we have used differentiability of $A_\epsilon(x)$, and the property $\|X_t^c - X_{T_n^c}^c\| \leq \epsilon(t - T_n^c)$. Integrating on the time interval $[T_n^c, T_{n+1}^c]$ yields

$$
\theta_{n+1} = \Delta T_{n+1}^c\left(\lambda_0 - \epsilon A_\epsilon^T(X_{T_n^c}^c)\mathcal{V}_n + \mathcal{O}(\epsilon^2)\right),
$$

so that

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

We also need the following technical lemma :

Lemma 4.2 (Random time change). *Let $(\bar{t} \mapsto \alpha_\epsilon(\bar{t}))_{\epsilon \geq 0}$ a sequence of continuous, strictly increasing random time changes in $\mathbb{R}^+$, which, for $\epsilon \to 0$, converges in distribution (on any time interval for the uniform convergence topology) towards the identity function $\bar{t} \mapsto \bar{t}$. Let $(\bar{t} \to X_\epsilon(\bar{t}))_{\epsilon \geq 0}$ a sequence of continuous random processes, such that* {l:techrandt}

$$
\bar{t} \to X_{\alpha_\epsilon(\bar{t})}^\epsilon
$$

converges in distribution when $\epsilon \to 0$ (on any time interval for the uniform convergence topology) towards a continuous process $\bar{t} \to X_{\bar{t}}^0$. Then the sequence $\left(\bar{t} \to X_{\bar{t}}^\epsilon\right)_{\epsilon \geq 0}$ also converges in distribution when $\epsilon \to 0$ (on any time interval for the uniform convergence topology) towards $\bar{t} \to X_{\bar{t}}^0$.

Proof. Since by assumption $(\bar{t} \mapsto \alpha^\epsilon(\bar{t}))_{\epsilon \geq 0}$ converges to a deterministic limit, the coupled process $\bar{t} \mapsto (X_{\alpha_\epsilon(\bar{t})}^\epsilon, \alpha_\epsilon(\bar{t}))$ also converges in distribution to $\bar{t} \mapsto (X_{\bar{t}}^0, \bar{t})$, and we can consider a Skorokhod embedding (a new probability space) associated with a ϵ-sequence where this convergence holds almost surely. Then we can write ($\wedge$ denotes the minimum, and $\mathbf{1}$ the identity function):

$$
\begin{aligned}
\sup_{\bar{t} \in [0,T]} &\left|X_{\bar{t}}^\epsilon - X_{\bar{t}}^0\right| \wedge 1 \\
&\leq \sup_{\bar{t} \in [0,T]}\left|X_{\bar{t}}^\epsilon - X_{\alpha_\epsilon^{-1}(\bar{t})}^0\right| \wedge 1 + \sup_{\bar{t} \in [0,T]}\left|X_{\bar{t}}^0 - X_{\alpha_\epsilon^{-1}(\bar{t})}^0\right| \wedge 1, \\
&\leq \sup_{\bar{s} \in [0,\alpha_\epsilon^{-1}(T)]}\left|X_{\alpha_\epsilon(\bar{s})}^\epsilon - X_{\bar{s}}^0\right| \wedge 1 + \sup_{\bar{s} \in [0,\alpha_\epsilon^{-1}(T)]}\left|X_{\alpha_\epsilon(\bar{s})}^0 - X_{\bar{s}}^0\right| \wedge 1 \\
&\leq \sup_{\bar{s} \in [0,2T]}\left|X_{\alpha^\epsilon(\bar{s})}^\epsilon - X_{\bar{s}}^0\right| \wedge 1 + \sup_{\bar{s} \in [0,2T]}\left|X_{\alpha_\epsilon(\bar{s})}^0 - X_{\bar{s}}^0\right| \wedge 1 + 2 \times \mathbf{1}_{\alpha^\epsilon(2T) \leq T}
\end{aligned}
$$

and the right hand side converges almost surely to 0 using the (uniform) continuity of $\bar{t} \mapsto X_{\bar{t}}^0$. □

Then the following result holds :

{pro:convcoupl}   **Proposition 4.3.** *Consider the gradient sensing process described in Section 2.3. 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 (4.1). As a consequence, the density $n^{c,\epsilon}$ of the process with direct gradient sensing defined in (4.3) converges weakly towards $n^{c,0}$, solution of (4.4).*

Proof. The proof requires the definition of some auxiliary processes. First, we perform a random time change on the process of positions of the bacterium, so that the jump times become deterministic, and will occur on regular time instances separated by ϵ^2/λ_0 on the diffusive time scale; the randomness is then transferred to the jump size. The auxiliary process is thus defined as

{eq:sum_X_c}
$$\tilde{X}_{\bar{t}}^{c,\epsilon} = X_0^c + \sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor - 1} \epsilon \Delta T_{n+1}^c \mathcal{V}_n + \underbrace{\left(\lambda_0 \bar{t}/\epsilon^2 - \lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor \right) \epsilon \Delta T_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor + 1}^c \mathcal{V}_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor}}_{r_1}.$$
(4.6)

The latter will be compared with a classical drifted random walk defined by the Markov chain

{eq:Markov}
$$\begin{cases} \Xi_0^c & = X_0^c \\ \Xi_{n+1}^c & = \Xi_n^c + \epsilon \dfrac{\theta_{n+1} \mathcal{V}_n}{\lambda_0} + \epsilon^2 \dfrac{\theta_{n+1} \mathcal{V}_n}{\lambda_0^2} A^T(\Xi_n^c) \mathcal{V}_n, \\ & = \Xi_n^c + \epsilon \left(\dfrac{\theta_{n+1}}{\lambda_0} \left(1 + \epsilon \dfrac{1}{\lambda_0} A^T(\Xi_n^c) \mathcal{V}_n \right) \right) \mathcal{V}_n, \end{cases}$$
(4.7)

to which we can associate a continuous process defined on a similar diffusive time scale by interpolation,

{eq:sum_Xi_c}
$$\tilde{\Xi}_{\bar{t}}^{c,\epsilon} = X_0^c + \left(\sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor - 1} \left(\Xi_{n+1}^c - \Xi_n^c \right) \right) + \underbrace{\left(\lambda_0 \bar{t}/\epsilon^2 - \lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor \right) \left(\Xi_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor + 1}^c - \Xi_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor}^c \right)}_{r_2}.$$
(4.8)

Now the computation of the diffusive limit of the *sequence* of processes $\left(\bar{t} \mapsto \tilde{X}_{\bar{t}}^{c,\epsilon} \right)_{\epsilon}$ for $\epsilon \to 0$ relies on four steps.

Step (i)   Show tightness of the coupled sequence $\left(\bar{t} \mapsto (\tilde{X}_{\bar{t}}^{c,\epsilon}, \tilde{\Xi}_{\bar{t}}^{\epsilon}) \right)_{\epsilon \geq 0}$. To this end, decompose the leading terms in these processes into a "martingale" term and a drift term.

Step (ii)   Show, using a Gronwall argument, that the process describing the bacterial position at jump times $\tilde{X}_{\bar{t}}^{c,\epsilon}$ and the classical Markovian random walk $\tilde{\Xi}_{\bar{t}}^{\epsilon}$ have the same advection-diffusion limit.

Step (iii) Compute the advection-diffusion limit of the Markovian random walk $\tilde{\Xi}_{\bar{t}}^{\epsilon}$ (and hence of the process at jump times $\tilde{X}_{\bar{t}}^{c,\epsilon}$).

Step (iv) Use a random time change to transfer the result on bacteria position process $X_{\bar{t}}^{c,\epsilon}$ itself.

Step (i). Tightness of the coupled sequence $\bar{t} \mapsto (\tilde{X}_{\bar{t}}^{c,\epsilon}, \tilde{\Xi}_{\bar{t}}^{\epsilon})$.

Let us recall that according to Kolmogorov's criterion, a sequence of continuous processes $\left(\bar{t} \mapsto \omega_{\bar{t}}^{\epsilon}\right)_{\epsilon \geq 0}$ is tight in the space $\mathcal{C}([0,\bar{T}], \mathbb{R}^d)$ of continuous processes endowed with uniform convergence if there exists a constant C independent of ϵ such that

$$\mathbb{E} |\omega_{\bar{t}}^{\epsilon} - \omega_{\bar{s}}^{\epsilon}|^4 \leq C (\bar{t} - \bar{s})^2, \qquad (4.9) \quad \text{\{eq:Kolmogorov\}}$$

for any $0 \leq \bar{s} \leq \bar{t} \leq \bar{T}$. In our case, checking (4.9) for the coupled process $\bar{t} \mapsto (\tilde{X}_{\bar{t}}^{c,\epsilon}, \tilde{\Xi}_{\bar{t}}^{c,\epsilon})$ is highly standard (see [25]). Remark that it is sufficient to check the criteria for each of the two separately. We recall the computation for the process $\bar{t} \mapsto \tilde{\Xi}_{\bar{t}}^{c,\epsilon}$, the case of $\bar{t} \mapsto \tilde{X}_{\bar{t}}^{c,\epsilon}$ being similar using (4.5). First, one can check (4.9) when $\bar{t} - \bar{s} \leq \lambda_0 \epsilon^2$. Indeed, in the latter case, only the rest r_2 (see equation (4.8)) is involved and

$$\mathbb{E} \left| \tilde{\Xi}_{\bar{t}}^{c,\epsilon} - \tilde{\Xi}_{\bar{s}}^{c,\epsilon} \right|^4 \leq C \frac{(\bar{t} - \bar{s})^4}{\epsilon^8} \epsilon^4 \leq C (\bar{t} - \bar{s})^2,$$

where we have used that

$$\left| \Xi_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor + 1}^c - \Xi_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor}^c \right| \leq C\epsilon. \qquad (4.10) \quad \text{\{eq:est-eps-here\}}$$

Assume now that $\bar{t} - \bar{s} > \lambda_0 \epsilon^2$. The terms coming from r_2 can be easily controlled, since, again using (4.10),

$$\mathbb{E} |r_2|^4 \leq C\epsilon^4 \leq C (\bar{t} - \bar{s})^2.$$

The terms of order ϵ in the definition (4.7) (defining a martingale) can be controlled by expanding the sum and eliminating the terms with null average. It yields:

$$\mathbb{E} \left| \sum_{n=\lfloor \lambda_0 \bar{s}/\epsilon^2 \rfloor}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor - 1} \epsilon \frac{\theta_{n+1} \mathcal{V}_n}{\lambda_0} \right|^4 \leq C\epsilon^4 \left(\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor - \lfloor \lambda_0 \bar{s}/\epsilon^2 \rfloor \right)^2$$

$$\leq C\epsilon^4 \left(\bar{t}/\epsilon^2 - \bar{s}/\epsilon^2 \right)^2 + C\epsilon^4$$

$$\leq C (\bar{t} - \bar{s})^4 \leq C (\bar{t} - \bar{s})^2.$$

In the last line, we use the fact that $(\bar{t} - \bar{s})^2 \leq 4\bar{T}^2 \leq C$. Finally the terms of order ϵ^2 in the definition (4.7) (defining a drift) can be controlled as

$$\mathbb{E} \left| \sum_{n=\lfloor \lambda_0 \bar{s}/\epsilon^2 \rfloor}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor - 1} \epsilon^2 \frac{\theta_{n+1} \mathcal{V}_n}{\lambda_0^2} A_\epsilon^T (\Xi_n^c) \mathcal{V}_n \right|^4 \leq C\epsilon^8 \left(\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor - \lfloor \lambda_0 \bar{s}/\epsilon^2 \rfloor \right)^4$$

$$\leq C\epsilon^8 \left(\bar{t}/\epsilon^2 - \bar{s}/\epsilon^2 \right)^4 + C\epsilon^8$$

$$\leq C (\bar{t} - \bar{s})^2,$$

where we again use the fact that $(\bar{t}-\bar{s})^2 \leq 4\bar{T}^2 \leq C$. We conclude that

$$\mathbb{E}\left|\tilde{\Xi}_{\bar{t}}^{c,\epsilon} - \tilde{\Xi}_{\bar{s}}^{c,\epsilon}\right|^4 \leq C(\bar{t}-\bar{s})^2,$$

and hence the sequence is tight.

   Step (ii). Comparison of the limits.

We can now use a Gronwall argument to show that $\bar{t} \mapsto \tilde{X}_{\bar{t}}^{c,\epsilon}$ and $\bar{t} \mapsto \tilde{\Xi}_{\bar{t}}^{\epsilon}$ have the same limit (in distribution) when $\epsilon \to 0$. First, since the coupled process is tight, we can use the Skorokhod embedding theorem, and consider a new probability space where an ϵ-sequence converge almost surely (in uniform norm) towards a limiting process $\bar{t} \mapsto (\tilde{X}_{\bar{t}}^{c,0}, \tilde{\Xi}_{\bar{t}}^{0})$. Then, we compare the two processes using (4.5) and (4.7). Incorporating the terms r_1 (see equation (4.6)) and r_2 in the sum, we get

$$\left|\tilde{X}_{\bar{t}}^{c,\epsilon} - \tilde{\Xi}_{\bar{t}}^{\epsilon}\right| \leq \sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \left(\epsilon^2 \frac{\theta_{n+1}}{\lambda_0^2}\left|\left(A_\epsilon(X_{T_n^c}^c) - A_\epsilon(\Xi_n^c)\right)\right| + \theta_{n+1}\mathcal{O}(\epsilon^3)\right),$$

so that taking expectation yields

$$\mathbb{E}\left|\tilde{X}_{\bar{t}}^{c,\epsilon} - \tilde{\Xi}_{\bar{t}}^{\epsilon}\right| \leq C \sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \epsilon^2 \mathbb{E}\left|\left(A_\epsilon(X_{T_n^c}^c) - A_\epsilon(\Xi_n^c)\right)\right| + \mathcal{O}(\epsilon).$$

Then, since the velocity of both processes is of $\mathcal{O}\epsilon$, and using the fact that A_ϵ is Lipschitz uniformly in ϵ, we get

$$\left|\left(A_\epsilon(X_{T_n^c}^c) - A_\epsilon(\Xi_n^c)\right)\right| \leq C\epsilon^{-2}\int_{n\epsilon^2/\lambda_0}^{(n+1)\epsilon^2/\lambda_0}\left|\tilde{X}_{\bar{s}}^{c,\epsilon} - \tilde{\Xi}_{\bar{s}}^{\epsilon}\right|d\bar{s} + \theta_{n+1}\mathcal{O}(\epsilon).$$

Finally we get

$$\mathbb{E}\left|\tilde{X}_{\bar{t}}^{c,\epsilon} - \tilde{\Xi}_{\bar{t}}^{\epsilon}\right| \leq C\int_0^{\bar{t}}\mathbb{E}\left|\tilde{X}_{\bar{s}}^{c,\epsilon} - \tilde{\Xi}_{\bar{s}}^{\epsilon}\right|d\bar{s} + \mathcal{O}(\epsilon),$$

and applying a Gronwall argument yields

$$\mathbb{E}\left|\tilde{X}_{\bar{t}}^{c,\epsilon} - \tilde{\Xi}_{\bar{t}}^{\epsilon}\right| \leq \mathcal{O}(\epsilon).$$

Taking the limit $\epsilon \to 0$ shows that $\tilde{X}_{\bar{t}}^{c,0} = \tilde{\Xi}_{\bar{t}}^{0}$ almost surely.

   Step (iii). Computation of the limit.

Standard results (see [25]) show that the sequence of processes $\bar{t} \mapsto \tilde{\Xi}_{\bar{t}}^{\epsilon}$ satisfying (4.7) converges in distribution for the uniform norm towards the process solution to the SDE with diffusion matrix

$$a := \lambda_0 \mathbb{E}\left(\frac{\theta_{n+1}^2 \mathcal{V}_n \otimes \mathcal{V}_n}{\lambda_0^2}\right) = \frac{2D}{\lambda_0},$$

and drift vector field

$$b(x) := \lambda_0 \mathbb{E}\left(\frac{\theta_{n+1}}{\lambda_0^2}\mathcal{V}_n A^T(x)\mathcal{V}_n\right) = \frac{DA(x)}{\lambda_0},$$

with D given by (4.2), which yields the SDE (4.1). By Step (ii), the same holds true for $\bar{t} \mapsto \tilde{X}_{\bar{t}}^{c,\epsilon}$.

Step (iv). Random time change.

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

$$\alpha_{c,\epsilon}(\bar{t}) = \epsilon^2 T_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor}^c + \epsilon^2 \left(T_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor+1}^c - T_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor}^c \right) \left(\lambda_0 \bar{t}/\epsilon^2 - \lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor \right) \qquad (4.11) \quad \texttt{\{eq:timechange\_c\}}$$

$$= \sum_{n=1}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \epsilon^2 \Delta T_n^c + \epsilon^2 \left(\Delta T_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor+1}^c \right) \left(\lambda_0 \bar{t}/\epsilon^2 - \lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor \right), \qquad (4.12)$$

which is exactly the linear interpolation of the jump times $(T_n^c)_{n \geq 0}$, and by construction is almost surely strictly increasing and continuous. Using (4.5), 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. [16]). By construction,

$$X_{\alpha_{c,\epsilon}(\bar{t})}^{c,\epsilon} = \tilde{X}_{\bar{t}}^{c,\epsilon}, \qquad (4.13) \quad \texttt{\{eq:alphatilde\_c\}}$$

so that we can apply the technical Lemma 4.2 to conclude. $\qquad \square$

Remark 4.4. In one spatial dimension, one can describe the evolution of the distribution density $p^c(x,v,t)$ of the process with direct gradient sensing as follows [7]. We introduce the distributions of left-moving and right-moving particles $p_{\pm}^c(x,t) = p^c(x,v=\pm 1,t)$, and write, using (3.5):

$$\begin{cases} \partial_t p_+^c + \epsilon \partial_x p_+^c = -\dfrac{\lambda_\epsilon^c(x,+1)}{2} p_+^c + \dfrac{\lambda_\epsilon^c(x,-1)}{2} p_-^c \\[2mm] \partial_t p_-^c - \epsilon \partial_x p_-^c = \dfrac{\lambda_\epsilon^c(x,+1)}{2} p_+^c - \dfrac{\lambda_\epsilon^c(x,-1)}{2} p_-^c \end{cases}. \qquad (4.14) \quad \texttt{\{eq:cprocess\_hydro1d}}$$

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

$$\lambda_\epsilon^c(x,v) = \lambda_0 - \epsilon A(x) v, \qquad (4.15) \quad \texttt{\{eq:lin-turn-1D\}}$$

is easy to see that equation (4.14) is equivalent to

$$\partial_t^2 n^c + \lambda_0 \partial_t n^c = \partial_x \left(\epsilon^2 \partial_x n^c - \epsilon^2 A(x) n^c \right), \qquad (4.16)$$

which, on diffusive time-scales $\bar{t} = t\epsilon^2$, is equivalent to (4.4) in the limit when $\epsilon \to 0$. Note, however, that the assumption (4.15) requires $|A(x)| \leq \lambda_0/\epsilon$ to retain positivity of the turning rate. For ϵ small enough compared to $A(x)$ (containing $\nabla S(x)$) this requirement is always fulfilled.

4.2. *Diffusion limit of the process with internal state*

{sec:limit-internal}

In the same way, a standard probabilistic diffusion approximation argument can be used to derive the pathwise diffusive limit of the process with internal state (2.5). We denote by

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

the process with internal state on diffusive time scales. We will show that this process converges towards a solution of the same advection-diffusion SDE (4.1) as the process with direct gradient sensing (2.10) for $\epsilon \to 0$, provided $A(x)$ is chosen according to

{eq:control_field}
$$A_0(x) = b^T \lim_{\epsilon \to 0} \frac{\tau_\epsilon}{\lambda_0 \tau_\epsilon + \mathrm{Id}} \nabla S(x), \tag{4.17}$$

assuming that the limit exists. The parameters b, τ_ϵ, and λ_0 in equation (4.17) were introduced in (3.2)-(2.4) as parameters of the process with internal state, and $\mathrm{Id} \in \mathbb{R}^{n \times n}$ is the identity matrix. If we introduce the bacterial density as

$$n(x,t) = \int_{\mathbb{Y}} \int_{\mathbb{V}} p(x,v,y,t) \mathcal{M}(\mathrm{d}v) \mathrm{d}y, \tag{4.18}$$

this implies that n satisfies (4.4) on diffusive time scales as $\epsilon \to 0$.

The proof proceeds along the same lines as the proof of Proposition 4.3. However, the derivation of an asymptotic expansion of the jump time differences becomes more involved. We need the following auxiliary definitions and lemmas :

{def:m} **Definition 4.5.** We denote by $m : \mathbb{R} \to \mathbb{R}^{n \times n}$ the function

{eq:m(t)}
$$m(t) := t\tau_\epsilon - \left(\mathrm{Id} - \mathrm{e}^{-t\tau_\epsilon^{-1}} \right) \tau_\epsilon^2, \tag{4.19}$$

whose derivative is given by

$$m'(t) = \tau_\epsilon \left(\mathrm{Id} - \mathrm{e}^{-t\tau_\epsilon^{-1}} \right).$$

Lemma 4.6. *For all $t \in \mathbb{R}$, we have*

$$\|m'(t)\| \le t,$$

as well as

$$\|m(t)\| \le \frac{t^2}{2}.$$

Proof. We have $m(0) = m'(0) = 0$ and $m''(t) = \mathrm{e}^{-t\tau_\epsilon^{-1}}$. By Assumption 3.1 on τ_ϵ, we have

$$\|m''(t)\| \le 1,$$

so that

$$\|m'(t)\| = \left\| \int_0^t m''(s)\,ds \right\| \le \int_0^t ds = t,$$

as well as

$$\|m(t)\| = \left\| \int_0^t m'(s)\,ds \right\| \leq \int_0^t s\,ds = \frac{t^2}{2}. \qquad \square$$

Lemma 4.7. *Let us denote (∇^2 denotes the Hessian) as*   {lem:DTasymp}

$$c_S = \left\| \nabla^2 S \right\|_\infty, \tag{4.20} \quad \text{\{eq:nonlin\_signal\}}$$

as well as the error term due to non-linearities of the problem as (recall that by assumption $\delta < 1/k$)

$$\mathrm{r}(\epsilon) = c_\lambda \epsilon^{k\delta} + c_F \epsilon^{1+\delta} + c_S \epsilon^2 < \mathcal{O}(\epsilon), \tag{4.21} \quad \text{\{eq:nl\_err\}}$$

where the constants are defined in resp. equations (2.4), (3.2), and (4.20). The difference between two jump times,

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

satisfies the equation

$$\theta_{n+1} = \lambda_0 \Delta T_{n+1} - b^T m'(\Delta T_{n+1}) Z_{T_n} \\ - \epsilon b^T m(\Delta T_{n+1}) \nabla S(X_{T_n}) \mathcal{V}_n + \left(\theta_{n+1}^3 + \theta_{n+1} \right) \mathcal{O}(\mathrm{r}(\epsilon)). \tag{4.22} \quad \text{\{eq:DT1\}}$$

Proof. Throughout the proof, we will use the estimate $|\Delta T_{n+1}| \leq C\theta_{n+1}$, without always mentioning explicitly. Duhamel's integration of the ODE in (2.5) on $[T_n, T_{n+1}]$ yields

$$Z_t = e^{-(t-T_n)\tau_\epsilon^{-1}} Z_{T_n} + \epsilon \int_{T_n}^t e^{-(t-t')\tau_\epsilon^{-1}} \nabla S(X_{t'}) V_{t'} dt' \\ + \int_{T_n}^t e^{-(t-t')\tau_\epsilon^{-1}} \left(\tau_\epsilon^{-1} Z_{t'} - F_\epsilon(Y_{t'}, S(X_{t'})) \right) dt'. \tag{4.23} \quad \text{\{eq:Duhamel\_lin\_2\}}$$

Since $X_{t'} = X_{T_n} + \epsilon(t' - T_n)\mathcal{V}_n$, with $t' \leq T_{n+1}$, we get,

$$|\nabla S(X_{t'}) - \nabla S(X_{T_n})| = \theta_{n+1} c_S \mathcal{O}(\epsilon),$$

using c_S as defined in (4.20) and the estimate $t' - T_n \leq \Delta T_{n+1} \leq C\theta_{n+1}$. Moreover, Assumption 3.1, combined with Assumption 3.4, ensures that

$$\left| \tau_\epsilon^{-1} Z_{t'} - F_\epsilon(Y_{t'}, S(X_{t'})) \right| = \epsilon^{1-\delta} c_F \mathcal{O}(|Z_{t'}|^{2\delta}) = c_F \mathcal{O}(\epsilon^{1+\delta}).$$

Inserting these two estimates in (4.23), yields

$$Z_t = e^{-(t-T_n)\tau_\epsilon^{-1}} Z_{T_n} + \epsilon \int_{T_n}^t e^{-(t-t')\tau_\epsilon^{-1}} dt' \nabla S(X_{T_n}) \mathcal{V}_{T_n} + (\theta_{n+1}^2 + \theta_{n+1}) \mathcal{O}\left(c_F \epsilon^{1+\delta} + c_S \epsilon^2 \right)$$

$$= e^{-(t-T_n)\tau_\epsilon^{-1}} Z_{T_n} + \epsilon \left(Id - e^{-(t-T_n)\tau_\epsilon^{-1}} \right) \tau_\epsilon \nabla S(X_{T_n}) \mathcal{V}_{T_n} + (\theta_{n+1}^2 + \theta_{n+1}) \mathcal{O}\left(c_F \epsilon^{1+\delta} + c_S \epsilon^2 \right).$$

$$\tag{4.24} \quad \text{\{eq:key\_Duhamel\}}$$

Integrating again, we get

$$\int_{T_n}^{T_{n+1}} Z_t \mathrm{d}t = \left(Id - \mathrm{e}^{-(T_{n+1}-T_n)\tau_\epsilon^{-1}}\right)\tau_\epsilon\left(Z_{T_n} - \epsilon\tau_\epsilon\nabla S(X_{T_n})\mathcal{V}_{T_n}\right)$$
$$+ \epsilon(T_{n+1}-T_n)\tau_\epsilon\nabla S(X_{T_n})\mathcal{V}_{T_n} + (\theta_{n+1}^3 + \theta_{n+1}^2)\mathcal{O}\left(c_F\epsilon^{1+\delta} + c_S\epsilon^2\right)$$
$$= m'(\Delta T_{n+1})Z_{T_n} + \epsilon m(\Delta T_{n+1})\nabla S(X_{T_n})\mathcal{V}_{T_n} + (\theta_{n+1}^3 + \theta_{n+1}^2)\mathcal{O}\left(c_F\epsilon^{1+\delta} + c_S\epsilon^2\right),$$

with $m(t)$ and $m'(t)$ defined as in Definition 4.5. Finally, (2.4) gives

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

and the result (4.22) follows.   □

The following estimate of jump times then follows :

{lem:DTestim}   **Lemma 4.8.** *The jump time variations can be written in the following form :*

{eq:DTestim}
$$\Delta T_{n+1} = \Delta T_{n+1}^0 + \epsilon\Delta T_{n+1}^1 + (\theta_{n+1}^6 + \theta_{n+1})\mathcal{O}\left(\mathrm{r}(\epsilon) + \epsilon^2\right), \tag{4.25}$$

where

{eq:estimDT0}
$$\Delta T_{n+1}^0 = \frac{\theta_{n+1}}{\lambda_0} + \frac{b^T}{\lambda_0}m'(\Delta T_{n+1}^0)Z_{T_n}, \tag{4.26}$$

{eq:estimDT1}
$$\Delta T_{n+1}^1 = \frac{1}{\lambda_0 - b^T\mathrm{e}^{-\Delta T_{n+1}^0\tau_\epsilon^{-1}}Z_{T_n}}b^T m(\Delta T_{n+1}^0)\nabla S(X_{T_n})\mathcal{V}_n, \tag{4.27}$$

or, for short,

{eq:estimDT0_bis}
$$\Delta T_{n+1}^0 = \frac{\theta_{n+1}}{\lambda_0} + \theta_{n+1}\mathcal{O}(\epsilon^\delta) \tag{4.28}$$

{eq:estimDT1_bis}
$$\Delta T_{n+1}^1 = \frac{1}{\lambda_0}b^T m\left(\frac{\theta_{n+1}}{\lambda_0}\right)\nabla S(X_{T_n})\mathcal{V}_n + \theta_{n+1}^2\mathcal{O}(\epsilon^\delta), \tag{4.29}$$

where the term of order ϵ^δ in (4.28) is independent of $\mathcal{V}_n$.

Proof. Let us first introduce the functions

$$\begin{cases} \Psi_0(t) = \lambda_0 t - b^T m'(t)Z_{T_n} \\ \Psi_1(t) = -b^T m(t)\nabla S(X_{T_n})\mathcal{V}_n, \end{cases}$$

Using these functions, we can rewrite the result (4.22) of Lemma 4.7, as

{eq:DT1bis}
$$\theta_{n+1} = \Psi_0(\Delta T_{n+1}) + \epsilon\Psi_1(\Delta T_{n+1}) + \left(\theta_{n+1}^3 + \theta_{n+1}\right)\mathcal{O}(\mathrm{r}(\epsilon)). \tag{4.30}$$

Also, the definitions (4.26)-(4.27) of $(\Delta T_{n+1}^0, \Delta T_{n+1}^1)$ can be written as

{eq:psi_DT}
$$\begin{cases} \Psi_0(\Delta T_{n+1}^0) = \theta_{n+1} \\ \Delta T_{n+1}^1 = -\dfrac{\Psi_1(\Delta T_{n+1}^0)}{\Psi_0'(\Delta T_{n+1}^0)}. \end{cases} \tag{4.31}$$

Taking the first line, add ϵ times the second line, and rearrange, one has

{eq:why_this}
$$\theta_{n+1} = \Psi_0(\Delta T_{n+1}^0) + \epsilon\Psi_0'(\Delta T_{n+1}^0)\Delta T_{n+1}^1 + \epsilon\Psi_1(\Delta T_{n+1}^0). \qquad (4.32)$$

We then consider the expansion

$$\Delta T_{n+1} = \Delta T_{n+1}^0 + \epsilon\Delta T_{n+1}^1 + R_{n+1}.$$

The proof of the Lemma then reduces to showing that R_{n+1} is small. Anticipating the result, the idea is to Taylor expand $\Psi_0(\Delta T_{n+1})$ and $\Psi_1(\Delta T_{n+1})$, use these Taylor expansions in (4.31), and compare to (4.32). We start by using a double Taylor expansion at order 1 for $\Psi_0(\Delta T_{n+1})$:

$$\begin{aligned}
\Psi_0(\Delta T_{n+1}) &= \Psi_0(\Delta T_{n+1}^0 + R_{n+1}) + \epsilon\Delta T_{n+1}^1\Psi_0'(\Delta T_{n+1}^0 + R_{n+1}) \\
&\quad + \frac{(\epsilon\Delta T_{n+1}^1)^2}{2}\Psi_0''(\Delta T_{n+1}^0 + R_{n+1} + \gamma_3\epsilon\Delta T_{n+1}^1), \\
&= \Psi_0(\Delta T_{n+1}^0) + R_{n+1}\Psi_0'(\Delta T_{n+1}^0 + \gamma_1 R_{n+1}) + \epsilon\Delta T_{n+1}^1\Psi_0'(\Delta T_{n+1}^0) \\
&\quad + \epsilon\Delta T_{n+1}^1 R_{n+1}\Psi_0''(\Delta T_{n+1}^0 + \gamma_2 R_{n+1}) + \frac{(\epsilon\Delta T_{n+1}^1)^2}{2}\Psi_0''(\Delta T_{n+1}^0 + R_{n+1} + \gamma_3\epsilon\Delta T_{n+1}^1),
\end{aligned}$$

and a simple Taylor expansion at order 0 for $\Psi_1(\Delta T_{n+1})$:

$$\Psi_1(\Delta T_{n+1}) = \Psi_1(\Delta T_{n+1}^0) + (\epsilon\Delta T_{n+1}^1 + R_{n+1})\Psi_1'(\Delta T_{n+1}^0 + \gamma_4(\epsilon\Delta T_{n+1}^1 + R_{n+1})),$$

where $\gamma_i \in [0,1], i = 1\ldots4$ are used to apply the intermediate value theorem to the rest of the Taylor expansions. Inserting the Taylor expansions in the identity (4.30), and comparing to (4.32) simplifies to

$$\begin{aligned}
0 = {}& \Psi_0'(\Delta T_{n+1}^0 + \gamma_1 R_{n+1})R_{n+1} \\
&+ \epsilon\Delta T_{n+1}^1 R_{n+1}\Psi_0''(\Delta T_{n+1}^0 + \gamma_2 R_{n+1}) + \frac{(\epsilon\Delta T_{n+1}^1)^2}{2}\Psi_0''(\Delta T_{n+1}^0 + R_{n+1} + \gamma_3\epsilon\Delta T_{n+1}^1) \\
&+ (\epsilon^2\Delta T_{n+1}^1 + \epsilon R_{n+1})\Psi_1'(\Delta T_{n+1}^0 + \gamma_4(\epsilon\Delta T_{n+1}^1 + R_{n+1})) + (\theta_{n+1}^3 + \theta_{n+1})\mathcal{O}(\mathrm{r}(\epsilon)),
\end{aligned}$$
$$(4.33) \quad \{\text{eq:DT1ter}\}$$

We now need some estimates on the terms involved in (4.33) to control R_{n+1}. First of all, we need some crude bounds on jump times. Since rates are bounded above and below by $\lambda_{\min}$ and $\lambda_{\max}$, we have

$$|\Delta T_{n+1}| \le C\theta_{n+1}.$$

Then, $\Psi_0'(t) = \lambda_0 - \lambda_0 b^T e^{-t\tau_\epsilon^{-1}} Z_{T_n}$, so that, using Assumption 3.4, there is a constant C such that

$$\frac{1}{2}\lambda_0 \le \sup_{t \in [0,+\infty[} \Psi_0'(t) \le \frac{3}{2}\lambda_0,$$

for $\epsilon \le 1/C$. As a consequence, ΔT_{n+1}^0 is well-defined by the implicit definition (4.26) and one has obviously

$$\left|\Delta T_{n+1}^0\right| \le C\theta_{n+1}.$$

In the same way

$$\left|\Delta T_{n+1}^1\right| \le Cm(\Delta T_{n+1}^0) \le C\theta_{n+1}^2,$$

and thus by definition:

$$|R_{n+1}| \le C(\theta_{n+1} + \epsilon\theta_{n+1}^2).$$

Then we need some bounds on the derivatives of Ψ_0 and Ψ_1. Using Lemma 3.4, we get

$$|\Psi_0''(t)| = \lambda_0 \left| b^T \tau_\epsilon^{-1} \mathrm{e}^{-t\tau_\epsilon^{-1}} Z_{T_n} \right| \le C.$$

Eventually, for any $t \ge 0$,

$$|\Psi_1'(t)| \le Cm'(t) \le Ct^2.$$

Inserting all the previous estimates in (4.33) to control R_{n+1}, we get

$$|R_{n+1}| \le C\epsilon(\theta_{n+1}^4 + \theta_{n+1}^2) + \left(\theta_{n+1}^3 + \theta_{n+1}\right)\mathcal{O}(\mathrm{r}(\epsilon)).$$

Inserting a second time in (4.33) to control R_{n+1}, it yields:

$$|R_{n+1}| \le C\epsilon^2(\theta_{n+1}^6 + \theta_{n+1}^2) + \left(\theta_{n+1}^5 + \theta_{n+1}\right)\mathcal{O}(\mathrm{r}(\epsilon)),$$

and the result follows. $\qquad\square$

Remark 4.9. It is crucial to note that the term of order ϵ^δ in (4.26)-(4.28) is independent of $\mathcal{V}_n$. This will ensure that the latter contributes to martingale terms in sums of the form $\sum_n \Delta T_n \mathcal{V}_n$.

Finally, the drift of the limiting diffusion (see Proposition below) will be computed using the following lemma :

{lem:av_m} **Lemma 4.10.** *Let θ be an exponential random variable of mean 1. Then:*

{eq:av_m}
$$\mathbb{E}\left(m\left(\frac{\theta}{\lambda_0}\right)\right) = \frac{1}{\lambda_0}\frac{\tau_\epsilon}{\mathrm{Id} + \lambda_0\tau_\epsilon}. \tag{4.34}$$

Proof. It is a straightforward computation:

$$
\begin{aligned}
\mathbb{E}\left(m\left(\frac{\theta}{\lambda_0}\right)\right) &= \int_0^{+\infty} m(t/\lambda)e^{-t}dt \\
&= \frac{\tau_\epsilon}{\lambda_0}\int_0^{+\infty} te^{-t}\,dt - \tau_\epsilon^2\int_0^{+\infty} \mathrm{e}^{-t}\,dt + \tau_\epsilon^2\int_0^{+\infty} \mathrm{e}^{-(\mathrm{Id}+\frac{\tau_\epsilon^{-1}}{\lambda_0})t}\,dt \\
&= \frac{\tau_\epsilon}{\lambda_0} - \tau_\epsilon^2 + \tau_\epsilon^2\left(\mathrm{Id} + \tau_\epsilon^{-1}/\lambda_0\right)^{-1} \\
&= \frac{\tau_\epsilon}{\mathrm{Id} + \lambda_0\tau_\epsilon}\left(\frac{1}{\lambda_0}\left(\mathrm{Id} + \lambda_0\tau_\epsilon\right)\left(\mathrm{Id} - \lambda_0\tau_\epsilon\right) + \lambda_0\tau_\epsilon^2\right) \\
&= \frac{1}{\lambda_0}\frac{\tau_\epsilon}{\mathrm{Id} + \lambda_0\tau_\epsilon}.
\end{aligned}
$$

□

Using these lemmas, we can prove the following result on the diffusive limit of the process with internal state :

{pro:convproc}   **Proposition 4.11.** *Consider the process with internal state define in Section 2.1, and the assumptions of Section 3.3. Assume $\lim_{\epsilon \to 0} \dfrac{\tau_\epsilon}{\lambda_0 \tau_\epsilon + \mathrm{Id}}$ exists, and that $\delta > 1/k$. Then, the process $\bar{t} \mapsto X_{\bar{t}}^\epsilon = X_{t/\epsilon^2}$ converges in distribution (for uniform convergence topology) towards $\bar{t} \mapsto X_{\bar{t}}^0$ solution to the SDE (4.1) with drift coefficient (4.17). As a consequence, the density $n_{\bar{t}}^\epsilon$ of the process converges weakly towards $n_{\bar{t}}^0 = n_{\bar{t}}^{c,0}$ solution of (4.4) with drift coefficient (4.17).*

Proof. Since the proof is very similar to the proof of Proposition 4.3, we will skip similar arguments and refer to the latter whenever possible. As in the proof of Proposition 4.3, we perform a random time change on the process of positions of the bacterium, so that a jump (tumble phase) become deterministic, and will occur on regular small time interval of length ϵ^2/λ_0. The latter is thus defined as:

$$\tilde{X}_{\bar{t}}^\epsilon = X_0 + \sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor - 1} \epsilon \Delta T_{n+1} \mathcal{V}_n + \underbrace{\left(\lambda_0 \bar{t}/\epsilon^2 - \lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor \right) \epsilon \Delta T_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor + 1} \mathcal{V}_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor}}_{r_1}.$$

(4.35)   {eq:sum_X}

To compute the diffusion limit $\tilde{X}_{\bar{t}}^\epsilon$, we will compare it to a classical drifted random walk defined by the Markov chain

$$\begin{cases} \Xi_0 = X_0 \\ \Xi_{n+1} = \Xi_n + \epsilon \dfrac{\theta_{n+1} \mathcal{V}_n}{\lambda_0} + \epsilon^2 \dfrac{\mathcal{V}_n}{\lambda_0} b^T m \left(\dfrac{\theta_{n+1}}{\lambda_0} \right) \nabla S(\Xi_n) \mathcal{V}_n. \end{cases}$$

(4.36)   {eq:Markov_int}

To the latter, we can associate a continuous process defined on diffusive time scales by interpolation

$$\tilde{\Xi}_{\bar{t}}^\epsilon = X_0 + \left(\sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor - 1} \Xi_{n+1} - \Xi_n \right) + \underbrace{\left(\lambda_0 \bar{t}/\epsilon^2 - \lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor \right) \left(\Xi_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor + 1} - \Xi_{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \right)}_{r_2}.$$

(4.37)   {eq:sum_Xi}

Now the computation of the diffusive limit of the sequence of processes $\bar{t} \mapsto \tilde{X}_{\bar{t}}^\epsilon$ relies again on four steps.

Step (i). Tightness of the coupled sequence $\bar{t} \mapsto (\tilde{X}_{\bar{t}}^\epsilon, \tilde{\Xi}_{\bar{t}}^\epsilon)$.
The proof of tightness is strictly similar to the proof in Proposition 4.3, so we leave it to the reader.

Step (ii). Comparison of the limits.
We can now use a Gronwall argument to show that $\bar{t} \mapsto \tilde{X}_{\bar{t}}^\epsilon$ and $\bar{t} \mapsto \tilde{\Xi}_{\bar{t}}^\epsilon$ have the same limit (in distribution) when $\epsilon \to 0$. First, since the coupled process is tight, we can consider a new probability space where a ϵ-sequence converges almost surely

(in uniform norm) towards the limiting process $\bar{t} \mapsto (\tilde{X}^0_{\bar{t}}, \tilde{\Xi}^0_{\bar{t}})$. Then, we compare the two processes using (4.25)-(4.35) and (4.36)-(4.37). For this purpose, we incorporate the rest terms r_1 and r_2 in the sum in (4.35) and (4.37), and we get the estimate

$$
\left| \tilde{X}^\epsilon_{\bar{t}} - \tilde{\Xi}^\epsilon_{\bar{t}} \right| \leq \left| \sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \epsilon \mathcal{V}_n \left(\Delta T^0_{n+1} - \frac{\theta_{n+1}}{\lambda_0} \right) \right|
$$
$$
+ \sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \epsilon^2 \mathcal{V}_n \left(\Delta T^1_{n+1} - \frac{b^T}{\lambda_0} m(\frac{\theta_{n+1}}{\lambda_0}) \nabla S(\Xi_n) \mathcal{V}_n \right) + \sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} (\theta^6_{n+1} + \theta_{n+1}) \mathcal{O}(\epsilon^3 + \epsilon \mathrm{r}(\epsilon)).
$$

Denoting

$$
M_n = \mathcal{V}_n b^T m'(\Delta T^0_{n+1}) Z_{T_n},
$$

and using the estimates in Lemma 4.8, we get

$$
\left| \tilde{X}^\epsilon_{\bar{t}} - \tilde{\Xi}^\epsilon_{\bar{t}} \right| \leq \left| \sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \epsilon M_n \right|
$$
$$
+ \sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \epsilon^2 \frac{\theta_{n+1}}{\lambda_0} \left| b^T m(\frac{\theta_{n+1}}{\lambda_0}) \right| \left| \nabla S(X_{T^c_n}) - \nabla S(\Xi_n) \right| + \sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} (\theta^6_{n+1} + \theta_{n+1}) \mathcal{O}(\epsilon^3 + \epsilon^{2+\delta} + \epsilon \mathrm{r}(\epsilon)).
$$

Taking expectation and using Jensen's inequality, as well as the fact that S is Lipschitz, then gives

$$
\mathbb{E} \left| \tilde{X}^\epsilon_{\bar{t}} - \tilde{\Xi}^\epsilon_{\bar{t}} \right| \leq \left(\mathbb{E} \left| \sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \epsilon M_n \right|^2 \right)^{1/2}
$$
$$
+ C \epsilon^2 \sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \mathbb{E} \left| X_{T^c_n} - \Xi_n \right| + \mathcal{O}(\epsilon + \epsilon^\delta + \epsilon^{-1} \mathrm{r}(\epsilon)).
$$

Using $|M_n| \leq C \theta_{n+1} \epsilon^\delta$ (which follows from Assumption 3.4) and the independence of the random variables $(\mathcal{V}_n)_{n \geq 0}$, we get

$$
\mathbb{E} \left| \sum_{n=0}^{\lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor} \epsilon M_n \right|^2 \leq C \lfloor \lambda_0 \bar{t}/\epsilon^2 \rfloor \times \epsilon^2 \times \epsilon^{2\delta} = \mathcal{O}(\epsilon^{2\delta}).
$$

Controlling the derivative of both processes on the time interval $[n\epsilon^2/\lambda_0, (n+1)\epsilon^2/\lambda_0]$ yields

$$
\left| \tilde{X}_{T^c_n} - \Xi_n \right| \leq \lambda_0 \epsilon^{-2} \int_{n\epsilon^2/\lambda_0}^{(n+1)\epsilon^2/\lambda_0} \left| \tilde{X}^\epsilon_{\bar{s}} - \tilde{\Xi}^\epsilon_{\bar{s}} \right| d\bar{s} + \theta_{n+1} \mathcal{O}(\epsilon).
$$

In the end, we get

$$
\mathbb{E} \left| \tilde{X}^\epsilon_{\bar{t}} - \tilde{\Xi}^\epsilon_{\bar{t}} \right| \leq C \int_0^{\bar{t}} \mathbb{E} \left| \tilde{X}^\epsilon_{\bar{s}} - \tilde{\Xi}^\epsilon_{\bar{s}} \right| d\bar{s} + \mathcal{O}(\epsilon + \epsilon^\delta + \epsilon^{-1} \mathrm{r}(\epsilon))
$$

and applying a Gronwall argument yields (expliciting r(ϵ)):

{eq:final_bound}
$$\mathbb{E}\left|\tilde{X}_{\bar{t}}^{\epsilon} - \tilde{\Xi}_{\bar{t}}^{\epsilon}\right| = \mathcal{O}((1+c_S)\epsilon + (1+c_F)\epsilon^{\delta} + c_{\lambda}\epsilon^{k\delta-1}). \tag{4.38}$$

Taking the limit $\epsilon \to 0$ shows that $\tilde{X}_{\bar{t}}^{0} = \tilde{\Xi}_{\bar{t}}^{0}$ almost surely.

Step (iii). Computation of the limit.

Standard results (see [25]) show that the sequence of processes $\bar{t} \mapsto \tilde{\Xi}_{\bar{t}}^{\epsilon}$ satisfying (4.36) converges in distribution for the uniform norm towards the process that is solution of the SDE with diffusion matrix :

$$a = \lambda_0 \mathbb{E}\left(\frac{\theta_{n+1}^2 \mathcal{V}_n \otimes \mathcal{V}_n}{\lambda_0^2}\right) = 2D/\lambda_0,$$

and drift vector field, using (4.34), see Lemma 4.10 :

$$b(x) = \lim_{\epsilon \to 0} \lambda_0 \mathbb{E}\left(\frac{\mathcal{V}_n}{\lambda_0} b^T m(\frac{\theta_{n+1}}{\lambda_0})\nabla S(x)\mathcal{V}_n\right)$$
$$= \lim_{\epsilon \to 0} D\left(b^T \frac{\tau_{\epsilon}}{\mathrm{Id}+\lambda_0\tau_{\epsilon}}\nabla S(x)\right)/\lambda_0,$$
$$\xrightarrow{\epsilon \to 0} DA(x)/\lambda_0,$$

with $A(x)$ defined in equation 4.17, yielding the SDE (4.1). By Step (ii), the same holds true for $\bar{t} \mapsto \tilde{X}_{\bar{t}}^{\epsilon}$.

Step (iv). Random time change.

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. We introduce the random time change

$$\alpha_{\epsilon}(\bar{t}) = \epsilon^2 T_{\lfloor \lambda_0\bar{t}/\epsilon^2 \rfloor} + \epsilon^2 \left(T_{\lfloor \lambda_0\bar{t}/\epsilon^2 \rfloor+1} - T_{\lfloor \lambda_0\bar{t}/\epsilon^2 \rfloor}\right)\left(\lambda_0\bar{t}/\epsilon^2 - \lfloor \lambda_0\bar{t}/\epsilon^2 \rfloor\right) \tag{4.39}$$

{eq:timechange}

$$= \sum_{n=1}^{\lfloor \lambda_0\bar{t}/\epsilon^2 \rfloor} \epsilon^2 \Delta T_n + \epsilon^2 \left(\Delta T_{\lfloor \lambda_0\bar{t}/\epsilon^2 \rfloor+1}\right)\left(\lambda_0\bar{t}/\epsilon^2 - \lfloor \lambda_0\bar{t}/\epsilon^2 \rfloor\right), \tag{4.40}$$

which is the linear interpolation of the jump times $(T_n)_{n\geq 0}$, and by construction is almost surely strictly increasing and continuous. Lemma 4.8 implies that $\Delta T_n = \frac{\theta_{n+1}}{\lambda_0} + (\theta_{n+1}^6 + \theta_{n+1})\mathcal{O}\left(\epsilon + \epsilon^{1+\delta} + \mathrm{r}(\epsilon)\right)$. Therefore, we get

$$\alpha_{\epsilon}(\bar{t}) = \frac{\epsilon^2}{\lambda_0}\sum_{n=1}^{\lfloor \lambda_0\bar{t}/\epsilon^2 \rfloor}\theta_n + \frac{\mathcal{O}(\epsilon^{2+\delta} + \epsilon^3 + \epsilon^2\mathrm{r}(\epsilon))}{\lambda_0}\sum_{n=1}^{\lfloor \lambda_0\bar{t}/\epsilon^2 \rfloor}(\theta_n^6 + \theta_{n+1}),$$

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., [16]). By construction,

$$X_{\alpha_{\epsilon}(\bar{t})}^{\epsilon} = \tilde{X}_{\bar{t}}^{\epsilon}, \tag{4.41}$$

{eq:alphatilde}

so that we can apply the technical Lemma 4.2 to conclude. $\qquad\square$

Remark 4.12. In one spatial dimension and with internal dynamics given by (2.8), Erban and Othmer argued on formal arguments (a moment expansion) that the evolution of the density $n(x,t)$ on hyperbolic time-scales satisfies

$$\partial_t^2 n + \lambda_0 \partial_t n = \partial_x \left(\epsilon^2 \partial_x n - \frac{b \epsilon^2 \tau_\epsilon}{1 + \lambda_0 \tau_\epsilon} S'(x) n \right), \qquad (4.42) \quad \{\texttt{eq:process\_hydro}\}$$

in the limit when $\epsilon \to 0$ [7]. It is easy to see that equation (4.42) is equivalent to (4.1) in this limit when choosing $A(x)$ in (4.1) according to (4.17).

Remark 4.13. Another relation between the process with internal dynamics and the process with direct gradient sensing is obtained when, for fixed ϵ, the bacteria learn the value of the chemoattractants concentrations infinitely fast. To investigate this limit, we may construct a sequence $(\tau_\epsilon{}^k)_{k=1}^\infty$, such that $\lim_{k \to \infty} \|\tau_\epsilon{}^k\|^{-1} = \infty$, and show that, for $k \to \infty$, the process with internal state converges to a process with direct gradient sensing, with jump rate given by $\lambda_\epsilon^c(x,v) = \lambda_\epsilon(\nabla S(x)v)$, i.e. the turning rate for the process with internal dynamics, evaluated for the deviation $z = \nabla S(x)v$.

{sec:discr}   ## 5. Discretization of velocity-jump processes

To observe the asymptotic behaviour in numerical simulations, the time discretization should approach the same diffusion limit as the time-continuous problem. This requires, in particular, that the discretized jump times are computed sufficiently accurately. In this section, we propose a time discretization of (2.5) and give a consistency result. The process with direct gradient sensing (2.10) can be discretized similarly. For ease of exposition, we consider the scalar equation (2.8) for the internal state; generalization to nonlinear systems of equations is suggested in Remark 5.1.

5.1. *Linear turning rate*

First, assume a linearized jump rate

{eq:jump-lin}
$$\lambda(z) = \lambda_0 - b^T z. \qquad (5.1)$$

When the chemoattractant profile is linear, $S(x) = S_0 + (\nabla S_0) \cdot x$, the ODE (2.8) can be explicitly solved for Z_t. For $t \in [T_n, T_{n+1})$, *i.e.* between two jumps, we then have

{e:discr-lin}
$$\begin{cases} X_t = X_{T_n} + \epsilon V_t (t - T_n) \\ Z_t = e^{-(t - T_n) \tau_\epsilon{}^{-1}} Z_{T_n} + \epsilon \tau_\epsilon \left(\mathrm{Id} - e^{-(t - T_n) \tau_\epsilon{}^{-1}} \right) (\nabla S_0) \mathcal{V}_n. \end{cases} \qquad (5.2)$$

Similarly to the computation in the proof of Lemma 4.7, we get, using $\Delta T_{n+1} := T_{n+1} - T_n$,

$$\int_{T_n}^{T_{n+1}} \lambda(Z_t) \mathrm{d}t = \lambda_0 (\Delta T_{n+1}) - b^T \left(Id - e^{-\Delta T_{n+1} \tau_\epsilon{}^{-1}} \right) \tau_\epsilon Z_{T_n}$$

{e:jump-lin}
$$- \epsilon b^T \left(\Delta T_{n+1} \tau_\epsilon - (\mathrm{Id} - e^{-\Delta T_{n+1} \tau_\epsilon{}^{-1}}) \tau_\epsilon{}^2 \right) (\nabla S_0) \mathcal{V}_n. \quad (5.3)$$

The jump time T_{n+1} can be calculated exactly from (5.3) by solving

$$\int_{T_n}^{T_{n+1}} \lambda(Z_t)\mathrm{d}t = \theta_{n+1}, \tag{5.4}$$

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

$$\begin{cases} X_t^{\delta t} = X_{n,k}^{\delta t} + \epsilon \mathcal{V}_n (t - t_{n,k}) \\ Z_t^{\delta t} = \exp(-(t - t_{n,k})\tau_\epsilon^{-1}) Z_{n,k}^{\delta t} + \epsilon \tau_\epsilon \left(\mathrm{Id} - \exp(-(t - t_{n,k})\tau_\epsilon^{-1}) \right) \nabla S(X_{n,k}^{\delta t}) \mathcal{V}_n. \end{cases} \tag{5.5}$$
{e:discr-nonlin}

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 $\int_{T_n^{\delta t}}^{T_{n+1}^{\delta t}} \lambda(Z_t^{\delta t})\mathrm{d}t$ using

$$\int_{T_n^{\delta t}}^{T_{n+1}^{\delta t}} \lambda(Z_t^{\delta t})\mathrm{d}t = \sum_{k=0}^{K-1} \int_{t_{n,k}}^{t_{n,k+1}} \lambda(Z_t^{\delta t})\mathrm{d}t + \int_{t_{n,K}}^{T_{n+1}^{\delta t}} \lambda(Z_t^{\delta t})\mathrm{d}t, \tag{5.6}$$

and then compute:

$$\int_{t_{n,k}}^{t_{n,k+1}} \lambda(Z_t^{\delta t})\mathrm{d}t = \lambda_0 \delta t - b^T \left(Id - \mathrm{e}^{-\delta t \tau_\epsilon^{-1}} \right) \tau_\epsilon Z_{T_n}^{\delta t}$$
$$- \epsilon b^T \left(\delta t \tau_\epsilon - (\mathrm{Id} - \mathrm{e}^{-\delta t \tau_\epsilon^{-1}})\tau_\epsilon^{2} \right) \nabla S(X_{n,k}^{\delta t})\mathcal{V}_n. \tag{5.7}$$

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

$$\int_{t_{n,K}}^{T_{n+1}^{\delta t}} \lambda(Z_t^{\delta t})\mathrm{d}t = \theta_{n+1} - \sum_{k=0}^{K-1} \int_{t_{n,k}}^{t_{n,k+1}} \lambda(Z_t^{\delta t})\mathrm{d}t, \tag{5.8}$$
{e:lin-Newton}

again using a Newton procedure. The use of (5.5) as a discretization scheme has a crucial consequence: the analysis performed in Section 4 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 (4.24) still holds. It is left to the reader to check that all the proofs performed in Section 4 are not affected by this discretization.

Remark 5.1. The discretization scheme (5.5) 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 3.4 still holds in this case, and also that (4.24) remains true.
{rem:nonlin_num}

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

Lemma 5.2. *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, using*   {lem:jump-lin}
the time discretization (5.5), *we have*

$$\left| T_{n+1} - T_{n+1}^{\delta t} \right| \leq C\epsilon^2 \delta t \theta_{n+1}^2, \tag{5.9}$$

and consequently,

$$\left| X_{T_{n+1}} - X_{T_{n+1}^{\delta t}}^{\delta t} \right| \leq C\epsilon^3 \delta t \theta_{n+1}^2. \tag{5.10}$$

$$\tag{5.11}$$

Proof. Using the assumption $T_n = T_n^{\delta t}$ and $X_{T_n} = X_{T_n}^{\delta t}$, we get straightforwardly

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

so that

$$\left| \nabla S(X_t) - \nabla S\left(X_{n, \lfloor (t_{n,0} - t')/\delta t \rfloor}^{\delta t} \right) \right| \leq C\delta t \epsilon.$$

Consider Duhamel's formula (4.23), and write a similar expression for the time discretized case for $t \in [T_n, T_{n+1}^{\delta t}]$:

{eq:Duhamel_num} $$Z_t^{\delta t} = e^{-(t-T_n)\tau_\epsilon^{-1}} Z_{T_n}^{\delta t} + \epsilon \int_{T_n}^t e^{-(t-t')\tau_\epsilon^{-1}} \nabla S\left(X_{n, \lfloor (t_{n,0} - t')/\delta t \rfloor}^{\delta t} \right) \mathcal{V}_n dt'. \tag{5.12}$$

Using the assumption $Z_{T_n}^{\delta t} = Z_{T_n}$, the difference between the two Duhamel's integrals then yields

{eq:dZ} $$\left| Z_t - Z_t^{\delta t} \right| \leq C\epsilon^2 \delta t \theta_{n+1}, \tag{5.13}$$

for $t \in [T_n, T_{n+1} \wedge T_{n+1}^{\delta t}]$.

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

$$\int_{T_n}^{T_{n+1}} \lambda(Z_t) dt = \int_{T_n^{\delta t}}^{T_{n+1}^{\delta t}} \lambda(Z_t^{\delta t}) dt = \theta_{n+1};$$

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

$$-\int_{T_n}^{T_{n+1} \wedge T_{n+1}^{\delta t}} b^T \left(Z_t - Z_t^{\delta t} \right) dt = \int_{T_{n+1} \wedge T_{n+1}^{\delta t}}^{T_{n+1}^{\delta t}} \lambda(Z_t^{\delta t}) dt - \int_{T_{n+1} \wedge T_{n+1}^{\delta t}}^{T_{n+1}} \lambda(Z_t) dt$$

$$= \left(T_{n+1}^{\delta t} - T_{n+1} \wedge T_{n+1}^{\delta t} \right) \lambda(Z_{t_1^*}^{\delta t}) + \left(T_{n+1} - T_{n+1} \wedge T_{n+1}^{\delta t} \right) \lambda(Z_{t_2^*}^{\delta t}).$$

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 $\lambda_{\min} \leq \lambda(z) \leq \lambda_{\max}$, we get

$$\left| T_{n+1}^{\delta t} - T_{n+1} \right| \leq C \int_{T_n}^{T_{n+1} \wedge T_{n+1}^{\delta t}} \left| Z_t - Z_t^{\delta t} \right| dt$$

$$\leq C\theta_{n+1}^2 \epsilon^2 \delta t.$$

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 $\left| X_{T_{n+1}} - X^{\delta t}_{T^{\delta t}_{n+1}} \right|$. This concludes the proof. $\qquad\square$

Note that the dependence in $\epsilon^3 \delta t$ shows, at least at the consistency level, that the error due to discretization remains of a higher order in ϵ than the difference between the process with internal state and its advection-diffusion limit.

5.2. *Nonlinear turning rate*

We now consider the nonlinear turning rate as defined in (2.9). We again discretize in time to obtain the time-discrete solution (5.5). The jump time $T^{\delta t}_{n+1}$ is now computed by linearizing (2.9) in each time step,

$$\lambda^{\delta t}\left(Z^{\delta t}_t, Z^{\delta t}_{n,k}\right) = \lambda(Z^{\delta t}_{n,k}) + \frac{\mathrm{d}\lambda(Z^{\delta t}_{n,k})}{\mathrm{d}z}\left(Z^{\delta t}_t - Z^{\delta t}_{n,k}\right) \qquad (5.14) \quad \texttt{\{eq:nonlinrate\_disc\}}$$

and approximating the integral

$$\int_{T^{\delta t}_n}^{T^{\delta t}_{n+1}} \lambda^{\delta t}(Z^{\delta t}_t, Z^{\delta t}_{n,\lfloor (t-T^{\delta t}_n)/\delta t\rfloor})\mathrm{d}t = \sum_{k=0}^{K-1} \int_{t_{n,k}}^{t_{n,k+1}} \lambda^{\delta t}(Z^{\delta t}_t, Z^{\delta t}_{n,k})\mathrm{d}t + \int_{t_{n,K}}^{T^{\delta t}_{n+1}} \lambda^{\delta t}(Z^{\delta t}_t, Z^{\delta t}_{n,K})\mathrm{d}t$$

$$= \theta_{n+1}, \qquad (5.15)$$

The jump time $T^{\delta t}_{n+1}$ can then again be computed using a Newton procedure as in (5.8). Here again, the analysis performed in Section 4 for the continuous time case, still holds *in an exact fashion* in the time discretized case. Also, a consistency property similar to Lemma 5.2 holds in the case of non-linear turning rate.

Lemma 5.3. *Consider a non-linear turning rate (2.9) discretized by (5.14) and* $\quad$ `{lem:jump-nonlin}`
(5.5). Assume $T^{\delta t}_n = T_n$, $X_{T_n} = X^{\delta t}_{T_n}$, $\mathcal{V}_n = V^{\delta t}_{T_n}$ *and* $Y^{\delta t}_{T_n} = Y_{T_n}$, *then*

$$\left| T_{n+1} - T^{\delta t}_{n+1} \right| \leq C\epsilon^2(\delta t\theta^2_{n+1} + \delta t^2\theta_{n+1}), \qquad (5.16)$$

and consequently,

$$\left| X_{T_{n+1}} - X^{\delta t}_{T^{\delta t}_{n+1}} \right| \leq C\epsilon^3(\delta t\theta^2_{n+1} + \delta t^2\theta_{n+1}). \qquad (5.17)$$

Proof. Similarly to the linear case we can show that

$$\left| T_{n+1} - T^{\delta t}_{n+1} \right| \leq C \int_{T_n}^{T_{n+1}\wedge T^{\delta t}_{n+1}} \left| \lambda(Z_t) - \lambda^{\delta t}\left(Z^{\delta t}_t, Z^{\delta t}_{n,\lfloor \frac{t_{n,0}-t}{\delta t}\rfloor}\right) \right|\mathrm{d}t. \qquad (5.18) \quad \texttt{\{eq:diff1\}}$$

We then estimate the righthand side of (5.18). This is done in two steps :

- Step (i) Estimate difference between turning rate and locally linearized turning rate for the discretized deviations $\left| \lambda(Z^{\delta t}_t) - \lambda^{\delta t}(Z^{\delta t}_t, Z^{\delta t}_{n,k}) \right|$.
- Step (ii) Estimate difference between turning rates for the discretized and time-continuous deviations $\left| \lambda(Z_t) - \lambda(Z^{\delta t}_t) \right|$.

Step (i): $\left|\lambda(Z_t^{\delta t}) - \lambda^{\delta t}(Z_t^{\delta t}, Z_{n,k}^{\delta t})\right|$.

A Taylor expansion for $t \in [t_{n,k}, t_{n,k+1}]$ yields

$$\left|\lambda(Z_t^{\delta t}) - \lambda^{\delta t}(Z_t^{\delta t}, Z_{n,k}^{\delta t})\right| \le C\left(Z_t^{\delta t} - Z_{n,k}^{\delta t}\right)^2.$$

Then, remarking that, for $t \in [t_{n,k}, t_{n,k+1}]$, we have

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

we get

$$\left|Z_t^{\delta t} - Z_{n,k}^{\delta t}\right| = \left|\left(\mathrm{Id} - \exp\left(-\frac{t-t_{n,k}}{\tau_\epsilon}\right)\right)\left(Z_{n,k}^{\delta t} - \tau_\epsilon\epsilon\nabla S(X_{n,k}^{\delta t})\right)\right|,$$
$$= \left|m'(t-t_{n,k})(eps\nabla S(X_{n,k}^{\delta t}) - \tau_\epsilon^{-1} Z_{n,k}^{\delta t})\right|$$
$$\le C\delta t\epsilon.$$

In the last line of the above, we have used the estimate

$$\left|Z_t^{\delta t}\right| = \mathcal{O}(\epsilon^\delta)$$

as in the non discretized case, which follows from Duhamel's formula (5.12) and the assumption $\left|Z_{T_n}^{\delta t}\right| = \mathcal{O}(\epsilon^\delta)$. Finally, we obtain

$$\int_{t_{n,k}}^{t_{n,k+1}} \left|\lambda(Z_t^{\delta t}) - \lambda^{\delta t}(Z_t^{\delta t}, Z_{n,k}^{\delta t})\right| \mathrm{d}t \le C\delta t^3\epsilon^2,$$

and thus:

$$\int_{T_n}^{T_{n+1}\wedge T_{n+1}^{\delta t}} \left|\lambda(Z_t) - \lambda^{\delta t}\left(Z_t^{\delta t}, Z_{n,\lfloor\frac{t_{n,0}-t}{\delta t}\rfloor}^{\delta t}\right)\right| \mathrm{d}t \le C\delta t^2\epsilon^2\theta_{n+1}$$

Step (ii): $\left|\lambda(Z_t) - \lambda(Z_t^{\delta t})\right|$.

Using Duhamel's formula as in the proof of Lemma 5.2, and using (5.13) above yields

$$\left|\lambda(Z_t) - \lambda(Z_t^{\delta t})\right| \le C\left|Z_t - Z_t^{\delta t}\right| \le C\delta t\epsilon^2\theta_{n+1}.$$

Finally we get, using (5.18),

$$\left|T_{n+1} - T_{n+1}^{\delta t}\right| \le C\epsilon^2(\delta t\theta_{n+1}^2 + \delta t^2\theta_{n+1}).$$

This concludes the proof. $\square$

6. Conclusions and discussion

{sec:concl}

This paper proved rigorously, using probabilistic arguments, the convergence of a velocity-jump process for chemotaxis of bacteria with internal state to an advection-diffusion limit, and showed that the same advection-diffusion limit can be obtained as the diffusion limit of a simpler, coarse process with direct gradient sensing. From a numerical point of view, we introduced a time-discretized model that retains these diffusion limits.

A direct simulation using stochastic particles presents a large statistical variance, even in the diffusive asymptotic regime, where the limiting behavior of the bacterial density is known explicitly. Consequently, it is difficult to assess accurately how the solutions of the fine-scale model differ from their advection-diffusion limit in intermediate regimes. Therefore, numerical results are deferred to the companion paper [21], where we will use the results of this paper to construct a hybrid scheme for the simulation of the fine-scale process with "asymptotic" variance reduction. The hybrid scheme simulates both the process with internal state and the simpler process with direct gradient sensing using the same random numbers, and exploits a pathwise coupling between the two trajectories. The variance reduction that is "asymptotic" in the sense that the variance vanishes in the diffusion limit.

Acknowledgements

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. Nikolaos Bournaveas and Vincent Calvez. Global existence for the kinetic chemotaxis model without pointwise memory effects, and including internal variables. *Kinet. Relat. Models*, 1(1):29–48, 2008.
4. 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.
5. MP Brenner, P Constantin, LP Kadanoff, A Schenkel, and SC Venkataramani. Diffusion, attraction and collapse. *Nonlinearity*, 12(4):1071–1098, 1999.
6. 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.
7. R Erban and HG Othmer. From individual to collective behavior in bacterial chemotaxis. *SIAM Journal on Applied Mathematics*, 65(2):361–391, 2004.
8. 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.
9. SN Ethier and TG Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics, 1986.

10. J. Fontbona, H. Guérin, and F. Malrieu. Quantitative estimates for the long time behavior of a PDMP describing the movement of bacteria. *ArXiv e-prints*, June 2010.
11. 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.
12. 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.
13. D Horstman. From 1970 until present: the Keller–Segel model in chemotaxis and its consequences I. *Jahresber. Deutsh. Math.-Verein*, 105(3):103–165, 2003.
14. D Horstman. From 1970 until present: the Keller–Segel model in chemotaxis and its consequences II. *Jahresber. Deutsh. Math.-Verein*, 106(2):51–69, 2004.
15. E Keller and L Segel. Initiation of slime mold aggregation viewed as an instability. *Journal on Theoretical Biology*, 26:399–415, 1970.
16. HJS Kushner. *Approximation and weak convergence methods for random processes, with application to stochastic systems theory*. MIT Press, Cambridge, MA, 1984.
17. HG Othmer, SR Dunbar, and W Alt. Models of dispersal in biological-systems. *J Math Biol*, 26(3):263–298, 1988.
18. HG Othmer and T Hillen. The diffusion limit of transport equations ii: Chemotaxis equations. *SIAM Journal on Applied Mathematics*, 62(4):1222–1250, 2002.
19. 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.
20. C Patlak. Random walk with persistence and external bias. *B Math Biol*, 15:311–338, 1953.
21. M Rousset and G Samaey. Simulating individual-based models of bacterial chemotaxis with asymptotic variance reduction. *Mathematical Models and Methods in Applied Sciences*, 2010 (submitted).
22. Jonathan Saragosti, Vincent Calvez, Nikolaos Bournaveas, Axel Buguin, Pascal Silberzan, and Benoit Perthame. Mathematical description of bacterial traveling pulses. *PLoS Comput Biol*, 6(8):e1000890, 08 2010.
23. AM Stock. A nonlinear stimulus-response relation in bacterial chemotaxis. *P Natl Acad Sci Usa*, 96(20):10945–10947, 1999.
24. Daniel W. Stroock. Some stochastic processes which arise from a model of the motion of a bacterium. *Probability Theory and Related Fields*, 28:305–315, 1974. 10.1007/BF00532948.
25. Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.
26. Chuan Xue and Hans G. Othmer. Multiscale models of taxis-driven patterning in bacterial populations. *SIAM J. Appl. Math.*, 70(1):133–167, 2009.