DERIVATION OF THE BACTERIAL RUN-AND-TUMBLE KINETIC MODEL : QUANTITATIVE AND STRONG CONVERGENCE RESULTS

ALAIN BLAUSTEIN[1]

ABSTRACT. During the past century, biologists and mathematicians investigated two mechanisms underlying bacteria motion: the run phase during which bacteria move in straight lines and the tumble phase in which they change their orientation. When surrounded by a chemical attractant, experiments show that bacteria increase their run time as moving up concentration gradients, leading to a biased random walk towards favorable regions. This observation raises the following question, which has drawn intense interest from both biological and mathematical communities:

what cellular mechanisms enable bacteria to feel concentration gradients ?

In this article, we investigate an asymptotic regime that was proposed to explain this ability thanks to internal mechanisms. More precisely, we derive the run-and-tumble kinetic equation with concentration's gradient dependent tumbling rate from a more comprehensive model, which incorporates internal cellular mechanisms. Our result improves on previous investigations, as we obtain strong convergence towards the gradient dependent kinetic model with quantitative and formally optimal convergence rates. The main ingredient consists in identifying a set of coordinates for the internal cellular dynamics in which concentration gradients arise explicitly. Then, we use relative entropy methods in order to capture quantitative measurement of the distance between the model incorporating cellular mechanisms and the one with concentration gradient dependent tumbling rate.

CONTENTS

1. INTRODUCTION

Quantitative studies of chemotaxis date back to the 19th century with the pioneering works of Engelmann and Pfeffer, followed by numerous other scientists [8, 40, 1]. At that time, very little was known about the underlying mechanisms driving the movement of organisms toward and away from a chemical. In 1953, C. Patlak [36] established the link between macroscopic motion of sensitive organisms and biased velocity jumps at the microscopic level. He proposed a quite abstract framework in which he derived the Patlak-Keller-Segel model using a kinetic description of organisms. This approach was then extended in the early 80's by W. Alt [3] who specifically focused on chemo-sensitive cells and simplified Patlak's arguments. He proposed a slightly more detailed equation than the kinetic run-and-tumble model as we know today and formally derived Patlak-Keller-Segel type model in the appropriate scaling. H. Othmer *et al.* [35] simplified further the analysis in the case where cells are not affected by the presence of chemicals. They proposed the kinetic run-and-tumble model that is used today. This model prescribes the dynamics of the kinetic distribution $\bar{p}(t, \boldsymbol{x}, \boldsymbol{v})$ of bacteria for all time

$t \geq 0$, over the phase space $(\boldsymbol{x}, \boldsymbol{v}) \in \mathbb{R}^d \times V$, where the spatial variable $\boldsymbol{x}$ lies in $\mathbb{R}^d$ whereas velocities $\boldsymbol{v}$ lie in a smooth open subset V of $\mathbb{R}^d$. The distribution function $\bar{p}(t, \boldsymbol{x}, \boldsymbol{v})$ satisfies a Boltzmann type equation which reads

$$(1.1) \qquad \frac{\partial \bar{p}}{\partial t} + \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} \bar{p} = Q[\bar{p}].$$

On the left-hand side of (1.1), the transport operator $\boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} \bar{p}$ describes the "running" phase during which bacteria move according to the straight lines generated by their velocities $\boldsymbol{v}$. On the right-hand side of (1.1), the operator $Q[\bar{p}]$ describes the "tumbling" phase within which bacteria modify their velocity. Numerous experiments [34, 6, 47, 5, 7, 29, 12] enabled to measure the parameters involved in the tumbling process of various types of cells, including tumbling rate and post-tumbling velocity distribution. It appears that bacteria such as *E. Coli* decrease their tumbling rate proportionally to the spatio-temporal variations of the logarithmic chemo-attractant concentration [34, 6, 12, 33, 32], leading to a biased random walk towards more favorable regions. Thus, the effective external signal $M(t, \boldsymbol{x})$ may be interpreted as the logarithm $\ln(S(t, \boldsymbol{x}))$ of the concentration of chemical $S(t, \boldsymbol{x})$ at time $t \geq 0$ and position $\boldsymbol{x} \in \mathbb{R}^d$. It influences the tumbling process through its spatio-temporal variations, or path-wise gradient, defined as

$$(1.2) \qquad D_t M(t, \boldsymbol{x}, \boldsymbol{v}) = \partial_t M(t, \boldsymbol{x}) + \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} M(t, \boldsymbol{x}), \quad \forall (t, \boldsymbol{x}, \boldsymbol{v}) \in \mathbb{R}^+ \times \mathbb{R}^d \times V.$$

The tumbling operator $Q[\bar{p}]$ then takes the following general form [3, 35, 20, 30, 14, 16]:

$$(1.3) \qquad Q[\bar{p}] = \int_{V \times \mathbb{R}} \bar{\Lambda}\left(D_t M(t, \boldsymbol{x}, \boldsymbol{v}'), \boldsymbol{v}, \boldsymbol{v}'\right) \bar{p}\left(t, \boldsymbol{x}, \boldsymbol{v}'\right) - \bar{\Lambda}\left(D_t M(t, \boldsymbol{x}, \boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v}\right) \bar{p}\left(t, \boldsymbol{x}, \boldsymbol{v}\right) \mathrm{d} \boldsymbol{v}',$$

where the tumbling kernel $\bar{\Lambda}(D_t M, \boldsymbol{v}, \boldsymbol{v}')$ describes the probability for a bacteria with velocity $\boldsymbol{v}'$ to transfer to velocity $\boldsymbol{v}$ for a given effective signal variation $D_t M$.

This phenomenological model raises the natural question of understanding the internal mechanisms which enable cells to feel and respond to external spatio-temporal variations, despite the gap that separates their body size and the characteristic scale at which these variations occur.

Attempts to answer such a question date back to the 50's for the reaction to light stimulations of Phycomyces [22]. In the case of chemotaxis, R. Macnab and D. Koshland [34] experimentally confirmed that the driving cellular mechanism for sensing gradients is memory. This observation was corroborated and completed by a phenomenological model in [12]. In this article H. Berg *et al.* explained cell memory through a simple adaptation process. Adaptation arise from the interplay between fast internal variables which rapidly adapt, and therefore may be identified, to $M(t, \boldsymbol{x})$, and a slower internal variable $m(t) \in \mathbb{R}$ which relaxes to $M(t, \boldsymbol{x})$ after longer periods of time. In their simple model, the internal variable $m(t)$ of each cell satisfies the following equation

$$\frac{\mathrm{d}}{\mathrm{d} t} m(t) = -\frac{1}{\varepsilon}\left(m(t) - M(t, \boldsymbol{x})\right),$$

where the parameter $\varepsilon > 0$ represents the time delay of adaptation. The response of the cell is then proportional to the derivative of $m(t)$. The latter model is of course an approximation of the full dynamics and by now, extensive experimental investigations led to a fine understanding of the signaling pathways from chemoreceptors to flagellar-motor: identification of proteins involved in the process [13, 27, 39, 2], ultra-sensitivity mechanisms [21, 43, 42] and robustness through adaptation [4, 46]. Comprehensive models for internal cell dynamics have been subsequently proposed and experimentally tested [33, 32]. In particular, we know that the slower variable $m(t)$ accounts for the methylation level of chemoreceptors whereas faster dynamics rely on phosporylation of the protein chain involved in the signaling pathways. In this article, we will stick to the initial model proposed in [12] for simplicity and because it is sufficient for our purpose. Incorporating the latter internal dynamics into the kinetic description of the run-and-tumble process leads to the following equation

$$(1.4) \qquad \frac{\partial p^\varepsilon}{\partial t} + \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} p^\varepsilon - \frac{1}{\varepsilon} \partial_m\left[(m - M) p^\varepsilon\right] - \varepsilon \partial_m^2 p^\varepsilon = Q^\varepsilon[p^\varepsilon],$$

where $p^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, m)$ is the probability distribution of bacteria at time $t \geq 0$, position $\boldsymbol{x} \in \mathbb{R}^d$, with velocity $\boldsymbol{v} \in V$ and methylation level $m \in \mathbb{R}$. The parameter $\varepsilon > 0$ quantifies the ratio between adaptation and observation time scales. According to our previous observations and unlike $Q[\bar{p}]$, the tumbling operator Q^ε does not depend on $D_t M$, but on the time derivative of the methylation level $m(t)$ instead, that is

$$(1.5) \qquad Q^\varepsilon[p^\varepsilon] = \int_V \Lambda\left(\frac{m-M}{\varepsilon}, \boldsymbol{v}, \boldsymbol{v}'\right) p^\varepsilon\left(t, \boldsymbol{x}, \boldsymbol{v}', m\right) - \Lambda\left(\frac{m-M}{\varepsilon}, \boldsymbol{v}', \boldsymbol{v}\right) p^\varepsilon\left(t, \boldsymbol{x}, \boldsymbol{v}, m\right) \mathrm{d}\boldsymbol{v}'.$$

In this article, we rigorously and quantitatively prove that over large observation periods compared to the adaptation time scale, that is, in the $\varepsilon \to 0$ regime, solutions to (1.4)-(1.5) converge towards solutions to (1.1)-(1.3). This result constitutes a mathematical illustration of how adaptation may enable bacteria to feel gradients and thus respond to external signal.

Before proceeding to the core of this article, we give an overview of related mathematical investigations. The kinetic run-and-tumble model (1.1)-(1.3) has drawn lots of attention from the mathematical community as it captures biologically relevant phenomena. One of the most striking being the existence of chemotactic waves, which have been studied by J. Saragosti *et al.* in [38], on a modified Keller-Segel type model obtained in the macroscopic limit of (1.1)-(1.3). The existence of chemotactic waves was then extended to the full kinetic model by V. Calvez in [18]. In the kinetic case, V. Calvez *et al.* [19] numerically observed that several waves with different velocities may coexist, highlighting how intricate the situation may be. In addition to traveling waves, N. Bournaveas and V. Calvez [15] proved that, under spherically symmetric assumptions, and reminiscent of Keller-Segel models, (1.1)-(1.3) features finite time blow-up above a certain mass threshold, when it is coupled with an elliptic equation on the concentration of chemical. Below the mass threshold, the coupled nonlinear system admits global in time solutions, as proved under various assumptions on the tumbling kernel [20, 30, 14, 16]. Macroscopic Keller-Segel limits of the kinetic model have been extensively investigated [3, 35, 20, 38] and, in the hyperbolic scaling, singular limits were derived by F. James and N. Vauchelet [31].

Let us now review the literature on the model with internal variables (1.4)-(1.5). First, R. Erban and H. Othmer integrated more detailed internal dynamics to the kinetic model, in the case where the tumbling kernel is a linear function of the internal state, in dimension one [24] and then in higher dimensions [25]. Authors also took into account external forces [26]. This approach was then extended by C. Xue and H. Othmer in order to include nonlinear dependence with respect to internal variables and time dependent external signal [45]. In these articles, authors explain how to obtain macroscopic equations in parabolic and hyperbolic scaling. The longtime and macroscopic regimes of the model were also rigorously analyzed by Y. Dolak and C. Schmeiser in [23], when the turning kernel smoothly and only depends on $D_t M$. Similarly, macroscopic Keller-Segel equations were derived from a one dimensional model by G. Si *et al.* in [41]. Finally, the mechanics that relate the internal behavior and the resulting macroscopic dynamics were further analyzed by C. Xue in [44]. At this stage in the literature, there was no investigation regarding the relation between the kinetic model with internal variables (1.4)-(1.5) and its run-and-tumble counterpart (1.1)-(1.3). B. Perthame *et al.* filled this gap in [37] and recovered (1.1)-(1.3) from (1.4)-(1.5) in the limit $\varepsilon \to 0$. The present article completes the former, as we prove strong convergence with explicit rates. We also relax the assumption on the tumbling kernel Λ, merely requiring integrability and moment assumptions.

Let us now proceed to the core of this article. First, we provide formal insights on the asymptotic behavior of p^ε in the regime $\varepsilon \to 0$. This will supply the reader with an overview of the upcoming analysis. We consider non-negative and normalized solutions to (1.4)-(1.5), that is $p_0^\varepsilon := p^\varepsilon(t = 0) \in L^1\left(\mathbb{R}^d \times V \times \mathbb{R}\right)$ for all $\varepsilon > 0$ with

$$p_0^\varepsilon \geq 0 \quad \text{a.e. and} \quad \int_{\mathbb{R}^d \times V \times \mathbb{R}} p_0^\varepsilon(\boldsymbol{x}, \boldsymbol{v}, m) \, \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}m = 1.$$

Since equation (1.4) is conservative, the latter property is then satisfied for all time $t \geq 0$. Considering the leading order in ε in (1.4), we find

$$(m - M) \, p^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, m) \underset{\varepsilon \to 0}{\sim} 0,$$

which, due to our normalization hypothesis, means that p^ε concentrates into a Dirac distribution as ε vanishes. To desingularize the limit, we rescale p^ε as follows

$$(1.6) \qquad p^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, m) = \frac{1}{\varepsilon} q^\varepsilon \left(t, \boldsymbol{x}, \boldsymbol{v}, \frac{1}{\varepsilon} (m - N(t, \boldsymbol{x}, \boldsymbol{v})) \right),$$

for some function N to be determined. The equation on the re-scaled distribution q^ε is derived performing the following change of variable in (1.4)

$$(1.7) \qquad y \longleftarrow \frac{1}{\varepsilon} (m - N(t, \boldsymbol{x}, \boldsymbol{v})).$$

We find that q^ε solves

$$\frac{\partial q^\varepsilon}{\partial t} + \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} q^\varepsilon - \frac{1}{\varepsilon} \partial_y \left[\left(y + D_t N + \frac{N - M}{\varepsilon} \right) q^\varepsilon + \partial_y q^\varepsilon \right] = \tilde{Q}^\varepsilon [q^\varepsilon],$$

where D_t is defined in (1.2). Still with the view of desingularizing the limit as $\varepsilon \to 0$, we choose N as the solution to

$$(1.8) \qquad \begin{cases} D_t N = \dfrac{1}{\varepsilon} (M - N), \\[2mm] N(0, \boldsymbol{x}, \boldsymbol{v}) = M(0, \boldsymbol{x}), \end{cases}$$

and therefore deduce that q^ε solves

$$(1.9) \qquad \frac{\partial q^\varepsilon}{\partial t} + \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} q^\varepsilon - \frac{1}{\varepsilon} \partial_y [y \, q^\varepsilon + \partial_y q^\varepsilon] = \tilde{Q}^\varepsilon [q^\varepsilon],$$

where the re-scaled tumbling operator $\tilde{Q}^\varepsilon$ is given by

$$\tilde{Q}^\varepsilon [q^\varepsilon] = \int_V \Lambda \left(y' - D_t N(t, \boldsymbol{x}, \boldsymbol{v}'), \boldsymbol{v}, \boldsymbol{v}' \right) q^\varepsilon \left(t, \boldsymbol{x}, \boldsymbol{v}', y' \right)$$

$$- \Lambda \left(y - D_t N(t, \boldsymbol{x}, \boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v} \right) q^\varepsilon (t, \boldsymbol{x}, \boldsymbol{v}, y) \, \mathrm{d} \boldsymbol{v}',$$

and where we used the short-hand notation

$$y' = y + \left(N(t, \boldsymbol{x}, \boldsymbol{v}) - N(t, \boldsymbol{x}, \boldsymbol{v}') \right) / \varepsilon.$$

Thanks to the change of variable (1.7), the limit of q^ε is no longer singular. Indeed, the leading term in (1.9) as $\varepsilon \to 0$ is given by the Fokker-Planck operator $\partial_y [y \, q^\varepsilon + \partial_y q^\varepsilon]$ whose kernel is generated by the Maxwellian distribution defined as

$$(1.10) \qquad \mathcal{M}(y) = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} |y|^2 \right), \quad \forall \, y \in \mathbb{R}.$$

Hence, in the regime of fast adaptation, we expect

$$q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y) \underset{\varepsilon \to 0}{\sim} \bar{q}^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}) \mathcal{M}(y),$$

where $\bar{q}^\varepsilon$ denotes the marginal of q^ε with respect to y, that is

$$(1.11) \qquad \bar{q}^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}) = \int_{\mathbb{R}} q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y) \mathrm{d} y.$$

To determine the limiting dynamics of $\bar{q}^\varepsilon$, we formally replace q^ε with $\bar{q}^\varepsilon \mathcal{M}$ in (1.9) and close the equation by integrating with respect to y. We find that $\bar{q}^\varepsilon$ converges to a solution of (1.1)-(1.3) with tumbling kernel

$$(1.12) \qquad \bar{\Lambda}(m, \boldsymbol{v}, \boldsymbol{v}') = \int_{\mathbb{R}} \Lambda \left(y - m, \boldsymbol{v}, \boldsymbol{v}' \right) \mathcal{M}(y) \mathrm{d} y.$$

We emphasize that the latter identity explicitly relates the tumbling kernel in the model with internal dynamics (1.4)-(1.5) and the resulting kernel in the run-and-tumble model (1.1)-(1.3). Indeed, gathering these considerations, we deduce

$$q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y) \underset{\varepsilon \to 0}{\longrightarrow} \bar{p}(t, \boldsymbol{x}, \boldsymbol{v}) \mathcal{M}(y),$$

and then, inverting the change of variable (1.7) to translate the latter result on p^ε, we derive

$$p^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, m) \xrightarrow[\varepsilon \to 0]{} \bar{p}\, \mathcal{M}_{\varepsilon, N}(t, \boldsymbol{x}, \boldsymbol{v}, m)\,,$$

where $\bar{p}$ solves (1.1)-(1.3) with tumbling kernel (1.12) and where $\mathcal{M}_{\varepsilon, N}$ is defined for all $(t, \boldsymbol{x}, \boldsymbol{v}, m) \in \mathbb{R}^+ \times \mathbb{R}^d \times V \times \mathbb{R}$ as

$$(1.13) \qquad \mathcal{M}_{\varepsilon, N}(t, \boldsymbol{x}, \boldsymbol{v}, m) \;=\; \frac{1}{\varepsilon}\, \mathcal{M}\left(\frac{m - N(t, \boldsymbol{x}, \boldsymbol{v})}{\varepsilon}\right)\,,$$

with N given in (1.8). In Theorem 2.1 below, we rigorously prove this result. The asymptotic analysis of (1.4)-(1.5) was first carried in [37], in which B. Perthame *et al.* obtained the weak convergence of p^ε in probability spaces. In this article, we complete their result by deriving strong convergence in $L^1\left(\mathbb{R}^d \times V \times \mathbb{R}\right)$ and by providing explicit convergence rates.

2. Assumptions and main result

Before stating our main result, we specify our setting and the assumptions on the data of the problem.

Let us first point out that the existence theory has been extensively investigated in the case where (1.1)-(1.3) is coupled with an equation on the effective signal $M(t, \boldsymbol{x})$ [20, 30, 14, 16]. In our case, the external signal is given, which makes equations (1.1)-(1.3) and (1.4)-(1.5) linear. Therefore, existence and uniqueness of solutions is well established. Furthermore, L^∞ *a priori* bounds have been derived in [37] for (1.1)-(1.3). In what follows, we will not discuss well-posedeness issues and consider strong solutions to both (1.1)-(1.3) and (1.4)-(1.5) in order to focus on our main concern, that is, the regime $\varepsilon \to 0$.

We now specify our assumptions on the data of the problem. First, we suppose that the effective signal M is a given function verifying

$$(2.1) \qquad (M : (t, \boldsymbol{x}) \longmapsto M(t, \boldsymbol{x})) \in W^{2,\infty}\left(\mathbb{R}^+ \times \mathbb{R}^d\right)\,,$$

and we use the notation $M_0 := M(t = 0)$. Second, we suppose that the tumbling kernel Λ takes positive values and satisfies the following general assumptions

$$(2.2a) \qquad \begin{cases} \displaystyle\sup_{(\boldsymbol{v},m)\in V\times\mathbb{R}} \left(\int_V \left(1 + |\boldsymbol{v}'|^2\right) \Lambda\left(m, \boldsymbol{v}', \boldsymbol{v}\right) \mathrm{d}\boldsymbol{v}'\right) < +\infty\,, \\[4mm] \end{cases}$$

$$(2.2b) \qquad \phantom{\begin{cases}\end{cases}} \displaystyle\sup_{(\boldsymbol{v},m)\in V\times\mathbb{R}} \left(\int_V \Lambda\left(m, \boldsymbol{v}, \boldsymbol{v}'\right) \mathrm{d}\boldsymbol{v}'\right) < +\infty\,.$$

Regarding the sequence of initial conditions $(p_0^\varepsilon)_{\varepsilon>0}$, we assume that it satisfies the following moment and entropy conditions

$$(2.3) \qquad \sup_{\varepsilon>0} \left(\int_{\mathbb{R}^d\times V\times\mathbb{R}} \left(|\log\left(p_0^\varepsilon(\boldsymbol{x}, \boldsymbol{v}, m)\right)| + |\boldsymbol{x}| + |\boldsymbol{v}|^2 + |m|^2\right) p_0^\varepsilon(\boldsymbol{x}, \boldsymbol{v}, m)\, \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}m\right) < +\infty\,.$$

In addition, we suppose the initial data to be well-prepared in the following sense

$$(2.4) \qquad \sup_{\varepsilon>0} \left(\varepsilon^{-2} \int_{\mathbb{R}^d\times V\times\mathbb{R}} |m - M_0(\boldsymbol{x})|^2\, p_0^\varepsilon(\boldsymbol{x}, \boldsymbol{v}, m)\, \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}m\right) < +\infty\,.$$

The latter assumption will be thoroughly discussed below our main result Theorem 2.1. To conclude with the assumption on $(p_0^\varepsilon)_{\varepsilon>0}$, we assume that the marginals with respect to m converge to the initial data $\bar{p}_0$ of the solution $\bar{p}$ to (1.1)-(1.3), that is

$$(2.5) \qquad \sup_{\varepsilon>0} \left(\varepsilon^{-2} \|\bar{p}_0 - \bar{p}_0^\varepsilon\|_{L^1\left(\mathbb{R}^d\times V\right)}\right) < +\infty \quad \text{where} \quad \bar{p}_0^\varepsilon(\boldsymbol{x}, \boldsymbol{v}) = \int_\mathbb{R} p_0^\varepsilon(\boldsymbol{x}, \boldsymbol{v}, m)\mathrm{d}m\,.$$

Regarding the initial data of the limiting model, we suppose that it satisfies

$$(2.6) \qquad \int_{\mathbb{R}^d\times V} |\boldsymbol{v}|^2\, \bar{p}_0(\boldsymbol{x}, \boldsymbol{v})\, \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v} < +\infty\,.$$

Our main result establishes a rigorous and quantitative link between equations (1.1)-(1.3) and (1.4)-(1.5), providing a mathematical description of the mechanisms allowing bacteria to feel and respond to environmental gradients. More precisely, we prove that the marginal $\bar{p}^\varepsilon$ of p^ε converges towards the solution $\bar{p}$ to (1.1)-(1.3) in $L^1\left(\mathbb{R}^d \times V \times \mathbb{R}\right)$ with pointwise and polynomial control in time. Furthermore, we characterize the blow-up profile in m by proving that the methylation level distribution converges in L^1 towards a Maxwellian distribution with vanishing variance and whose center is explicitly determined. Our estimates are quantitative since we obtain explicit convergence rates with respect to the scaling parameter ε. To conclude, the time dependence is also explicit: we achieve global in time estimates with polynomial control, avoiding exponential blow-up. The first part of our result features linear growth, ensuring that our estimate stays valid for large time periods

Theorem 2.1. *Under assumptions (2.1) on M, (2.2a)-(2.2b) on Λ, (2.3)-(2.5) on $(p_0^\varepsilon)_{\varepsilon>0}$ and (2.6) on $\bar{p}_0$, consider a sequence of strong solutions $(p^\varepsilon)_{\varepsilon>0}$ to (1.4)-(1.5) with initial data $(p_0^\varepsilon)_{\varepsilon>0}$ and a strong solution $\bar{p}$ to (1.1)-(1.3) with initial data $\bar{p}_0$ and tumbling kernel $\bar{\Lambda}$ given in (1.12). There exist two positive constants ε_0 and C such that for all ε between 0 and ε_0 and all time $t \geq 0$, it holds*

$$\left\| p^\varepsilon - \bar{p}^\varepsilon \mathcal{M}_{\varepsilon,N} \right\|_{L^2\left([0,t], L^1\left(\mathbb{R}^d \times V \times \mathbb{R}\right)\right)} \leq C\sqrt{\varepsilon}\left(t + 1 + |\log(\varepsilon)|^{\frac{1}{2}}\right),$$

and

$$\|\bar{p}^\varepsilon(t) - \bar{p}(t)\|_{L^1\left(\mathbb{R}^d \times V \times \mathbb{R}\right)} \leq C\left(\sqrt{\varepsilon t}\left(t + 1 + |\log(\varepsilon)|^{\frac{1}{2}}\right) + \varepsilon^2\right),$$

where $\mathcal{M}_{\varepsilon,N}$ is defined in (1.13) and where the solution N to (1.8) satisfies

$$|D_t M(t,\boldsymbol{x}) - D_t N(t,\boldsymbol{x},\boldsymbol{v})| \leq C\left(\varepsilon + e^{-\frac{t}{\varepsilon}}\right)\left(1 + |\boldsymbol{v}|^2\right), \quad \forall(\boldsymbol{x},\boldsymbol{v}) \in \mathbb{R}^d \times V,$$

with D_t is defined in (1.2). In these estimates, the constant $C > 0$ can be computed explicitly and only depends on the constants in assumptions (2.1)-(2.6).

The main difficulty to obtain Theorem 2.1 consists in proving that the rescaled distribution q^ε introduced in (1.9) converges towards $\bar{p}\,\mathcal{M}$. This proof is divided into two main steps, corresponding to the formal computations developed in the introduction. First, we prove that the entropy dissipation along the flow of (1.9) controls the L^1 convergence of q^ε towards its Maxwellian local equilibrium $\bar{q}^\varepsilon\,\mathcal{M}$. Then, we deduce from the first step the convergence of $\bar{q}^\varepsilon$ towards the solution $\bar{p}$ to (1.1)-(1.3). This step relies on a fine decomposition of the error between $\bar{q}^\varepsilon$ and $\bar{p}$. More precisely, we take advantage of the dissipative structure of equation (1.1) by identifying signed contributions in the time fluctuations of the L^1 distance that separates $\bar{q}^\varepsilon$ and $\bar{p}$.

Before proceeding to the proof, let us comment on Theorem 2.1. Its main feature lies in the explicit convergence rates with respect to the adaptation/observation ratio ε, completed by the strong convergence in L^1. On the biological point of view, this may give insights regarding the range of validity for the model (1.1)-(1.3). Furthermore, pointwise convergence in time is achieved for $\bar{p}^\varepsilon$ and our estimates are global in time with at most polynomial growth. This also provides insights regarding the time scales over which the internal behavior of bacteria may have measurable contributions that the kinetic run-and-tumble model (1.1)-(1.3) cannot account for. We also emphasize that very mild assumptions are taken on the tumbling kernel Λ compared to what is usually done in the literature. To conclude, we point out that the constants in our estimates are explicitly computable in terms of the bounds assumed from assumption (2.1) through (2.6). In particular, we do not require regularity on the solution p^ε to carry out our proof.

The main restriction of Theorem 2.1 lies in the well-prepardness assumption (2.5). The latter may be removed at the cost of additional technicalities. To do so, the main ingredient lies in considering the following time dependent scaling instead of (1.7)

$$y \longleftarrow \frac{1}{\theta^\varepsilon(t)}\left(m - N(t,\boldsymbol{x},\boldsymbol{v})\right),$$

where θ^ε is explicitly given by $\theta^\varepsilon(t) = \varepsilon + (1-\varepsilon)e^{-2t/\varepsilon}$. The subsequent equation on q^ε remains unchanged except for we substitute ε by $\theta^\varepsilon(t)$. This strategy was successfully applied in [9] in the context of concentration phenomena for neural networks. Due to the very similar structure with the

present model, the arguments developed in [9] should adapt without further difficulty in our context. However, this path includes technicalities that we choose to avoid for the sake of clarity.

3. A priori estimates

In this section, we derive uniform in ε moment estimates for the solution q^ε to (1.9) and prove that the solution N to (1.8) converges to the effective signal M.

The following Lemma ensures that the solution N to (1.8) is globally Lipschitz with respect to $\boldsymbol{v}$, uniformly with respect to ε and also that N converges to M as ε vanishes. The proof relies on an explicit formula for N in terms of the given external signal M.

Lemma 3.1. *Under assumption* (2.1), *the solution* N *to* (1.8) *satisfies*

$$(3.1) \qquad \left|N(t,\boldsymbol{x},\boldsymbol{v}) - N(t,\boldsymbol{x},\boldsymbol{v}')\right| \leq \|\nabla_{\boldsymbol{x}} M\|_{L^\infty}\, \varepsilon \left|\boldsymbol{v} - \boldsymbol{v}'\right|,$$

and

$$(3.2) \qquad |D_t M(t,\boldsymbol{x}) - D_t N(t,\boldsymbol{x},\boldsymbol{v})| \ \leq \ 4\,\|M\|_{W^{2,\infty}} \left(\varepsilon + e^{-\frac{t}{\varepsilon}}\right)\left(1 + |\boldsymbol{v}|^2\right),$$

for all $(t,\boldsymbol{x},\boldsymbol{v},\boldsymbol{v}') \in \mathbb{R}^+ \times \mathbb{R}^d \times V^2$, *with* D_t *is defined in* (1.2).

Proof. We consider $(t,\boldsymbol{x},\boldsymbol{v},\boldsymbol{v}') \in \mathbb{R}^+ \times \mathbb{R}^d \times V^2$. First, we point out that since N solves (1.8), it is given by the following explicit formula

$$(3.3) \qquad N(t,\boldsymbol{x},\boldsymbol{v}) = M(0,\boldsymbol{x} - t\boldsymbol{v})\, e^{-\frac{t}{\varepsilon}} + \frac{1}{\varepsilon} \int_0^t M(s, \boldsymbol{x} + (s-t)\boldsymbol{v})\, e^{\frac{s-t}{\varepsilon}} \mathrm{d}s\,.$$

Let us first derive (3.1). According to (3.3), we have

$$N(t,\boldsymbol{x},\boldsymbol{v}) - N(t,\boldsymbol{x},\boldsymbol{v}') = \left(M(0,\boldsymbol{x} - t\boldsymbol{v}) - M(0,\boldsymbol{x} - t\boldsymbol{v}')\right) e^{-\frac{t}{\varepsilon}}$$

$$+ \frac{1}{\varepsilon} \int_0^t \left(M(s,\boldsymbol{x} + (s-t)\boldsymbol{v}) - M(s,\boldsymbol{x} + (s-t)\boldsymbol{v}')\right) e^{\frac{s-t}{\varepsilon}} \mathrm{d}s\,.$$

Taking the absolute value on both sides in the latter relation and applying the triangular inequality, this yields:

$$\left|N(t,\boldsymbol{x},\boldsymbol{v}) - N(t,\boldsymbol{x},\boldsymbol{v}')\right| \leq \left|M(0,\boldsymbol{x} - t\boldsymbol{v}) - M(0,\boldsymbol{x} - t\boldsymbol{v}')\right| e^{-\frac{t}{\varepsilon}}$$

$$+ \frac{1}{\varepsilon} \int_0^t \left|M(s,\boldsymbol{x} + (s-t)\boldsymbol{v}) - M(s,\boldsymbol{x} + (s-t)\boldsymbol{v}')\right| e^{\frac{s-t}{\varepsilon}} \mathrm{d}s\,.$$

According to (2.1), M is uniformly Lipschitz. Hence, we deduce

$$\left|N(t,\boldsymbol{x},\boldsymbol{v}) - N(t,\boldsymbol{x},\boldsymbol{v}')\right| \leq C\left|\boldsymbol{v} - \boldsymbol{v}'\right| t\, e^{-\frac{t}{\varepsilon}} + \frac{C}{\varepsilon}\left|\boldsymbol{v} - \boldsymbol{v}'\right| \int_0^t (t - s)\, e^{\frac{s-t}{\varepsilon}} \mathrm{d}s\,,$$

where $C = \|\nabla_{\boldsymbol{x}} M\|_{L^\infty}$. In the latter right-hand side, we estimate $t\,e^{-t/\varepsilon}$ with ε and the time integral thanks to the following relation

$$\int_0^t (t - s)e^{\frac{s-t}{\varepsilon}}\, \mathrm{d}s \ = \ \int_0^t \sigma\, e^{-\frac{\sigma}{\varepsilon}}\, \mathrm{d}\sigma \ = \ \varepsilon^2 - \left(\varepsilon\, t\, e^{-\frac{t}{\varepsilon}} + \varepsilon^2 e^{-\frac{t}{\varepsilon}}\right) \leq \varepsilon^2,$$

where the first equality is obtained through the change of variable $\sigma \leftarrow t - s$ whereas the last equality is obtained integrating by part with respect to σ. It yields the expected result

$$\left|N(t,\boldsymbol{x},\boldsymbol{v}) - N(t,\boldsymbol{x},\boldsymbol{v}')\right| \leq C\varepsilon \left|\boldsymbol{v} - \boldsymbol{v}'\right|.$$

We now estimate $D_t M - D_t N$ to prove (3.2). First, we rewrite N by integrating by part twice the time integral in (3.3) and find

$$N(t,\boldsymbol{x},\boldsymbol{v}) = M(t,\boldsymbol{x}) - \varepsilon\, D_t M(t,\boldsymbol{x}) + \varepsilon\, D_t M(0,\boldsymbol{x} - t\boldsymbol{v})\, e^{-\frac{t}{\varepsilon}} + \varepsilon \int_0^t D_t^2 M(s, \boldsymbol{x} + (s-t)\boldsymbol{v})\, e^{\frac{s-t}{\varepsilon}} \mathrm{d}s\,.$$

Then, we divide the latter relation by ε and replace $(M-N)/\varepsilon$ by $D_t N$ according to (1.8), this yields

$$D_t M(t, \boldsymbol{x}) - D_t N(t, \boldsymbol{x}, \boldsymbol{v}) = D_t M(0, \boldsymbol{x} - t\boldsymbol{v}) e^{-\frac{t}{\varepsilon}} + \int_0^t D_t^2 M(s, \boldsymbol{x} + (s-t)\boldsymbol{v}) e^{\frac{s-t}{\varepsilon}} \, \mathrm{d}s \, .$$

Taking the absolute value on both sides and using assumption (2.1) to bound the derivatives of M, we obtain the result

$$|D_t M(t, \boldsymbol{x}) - D_t N(t, \boldsymbol{x}, \boldsymbol{v})| \leq 4 \|M\|_{W^{2,\infty}} \left(\varepsilon + e^{-\frac{t}{\varepsilon}} \right) \left(1 + |\boldsymbol{v}|^2 \right) \, .$$

$\square$

In the next proposition, we provide uniform in ε moment estimates with respect to $\boldsymbol{x}$, $\boldsymbol{v}$ and y for the solution q^ε to (1.9). To obtain this result, we first propagate moments with respect to $\boldsymbol{v}$ using the integrability assumption (2.2a) on the tumbling kernel Λ, then deduce propagation of moments with respect to $\boldsymbol{x}$ using standard techniques for transport equations. Moments with respect to y are also estimated thanks assumption (2.2a) coupled with the estimates for N in Lemma 3.1 and the bound on moments with respect to $\boldsymbol{v}$. For later purposes, we use the moment estimates on $\boldsymbol{v}$ and y to prove that $\tilde{Q}^\varepsilon [q^\varepsilon]$ also has second order moments with respect to y. To conclude, we obtain propagation of moments in $\boldsymbol{v}$ for the solution $\bar{p}$ to (1.1)-(1.3) thanks to similar arguments.

Proposition 3.2. *Under assumptions* (2.2a) *on* Λ, (2.1) *on* M *and* (2.3) *on* $(p_0^\varepsilon)_{\varepsilon>0}$ *consider a sequence of strong solutions* $(p^\varepsilon)_{\varepsilon>0}$ *to* (1.4)-(1.5) *with initial data* $(p_0^\varepsilon)_{\varepsilon>0}$.
There exists a positive constant C *such that for all* $\varepsilon > 0$ *and all time* $t \geq 0$, *the function* q^ε *defined by* (1.6) *satisfies*

$$(3.4) \qquad \int_{\mathbb{R}^d \times V \times \mathbb{R}} |\boldsymbol{v}|^2 \, q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y) \, \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y \; < \; C \, (t+1) \, ,$$

and

$$(3.5) \qquad \int_{\mathbb{R}^d \times V \times \mathbb{R}} |\boldsymbol{x}| \, q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y) \, \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y \; < \; C \left(t^{\frac{3}{2}} + 1 \right) \, ,$$

and there exists $\varepsilon_0 > 0$ *such that for all* ε *between* 0 *and* ε_0, *it holds*

$$(3.6) \qquad \int_{\mathbb{R}^d \times V \times \mathbb{R}} |y|^2 q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y) \, \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y \; \leq \; C \left(\varepsilon^{-2} e^{-\frac{t}{\varepsilon}} + \varepsilon \, t + 1 \right) \, ,$$

and

$$(3.7) \qquad \int |y|^2 \tilde{Q}^\varepsilon [q^\varepsilon] (t, \boldsymbol{x}, \boldsymbol{v}, y) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y \; \leq \; C \left(t^{\frac{1}{2}} \varepsilon^{-1} e^{-\frac{t}{2\varepsilon}} + \varepsilon^{-1} e^{-\frac{t}{2\varepsilon}} + t + 1 \right) \, .$$

Furthermore, under assumption (2.6) *on* $\bar{p}_0$, *consider a strong solution* $\bar{p}$ *to* (1.1)-(1.3) *with initial data* $\bar{p}_0$. *There exists a positive constant* C *such that*

$$(3.8) \qquad \int_{\mathbb{R}^d \times V} |\boldsymbol{v}|^2 \, \bar{p}(t, \boldsymbol{x}, \boldsymbol{v}) \, \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v} \; < \; C \, (t+1) \, , \quad \forall \, t \in \mathbb{R}^+ \, .$$

Proof. We first prove (3.4). We denote by $\left\| |\boldsymbol{v}|^2 p^\varepsilon(t) \right\|_{L^1}$ the second order moments of q^ε with respect to $\boldsymbol{v}$, that is

$$\left\| |\boldsymbol{v}|^2 q^\varepsilon(t) \right\|_{L^1} \; = \; \int_{\mathbb{R}^d \times V \times \mathbb{R}} |\boldsymbol{v}|^2 \, q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y) \, \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y \, .$$

To estimate $\left\| |\boldsymbol{v}|^2 q^\varepsilon \right\|_{L^1}$, we will compute its time derivative. However, we first notice that q^ε and the solution p^ε to (1.4) have the same second order moments with respect to $\boldsymbol{v}$. Indeed, inverting (1.7), we obtain

$$\left\| |\boldsymbol{v}|^2 q^\varepsilon(t) \right\|_{L^1} \; = \; \int_{\mathbb{R}^d \times V \times \mathbb{R}} |\boldsymbol{v}|^2 \, p^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, m) \, \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}m \, .$$

Therefore the time derivative of $\left\| |\boldsymbol{v}|^2 q^\varepsilon \right\|_{L^1}$ is obtained multiplying equation (1.4) by $|\boldsymbol{v}|^2$ and integrating with respect to $(\boldsymbol{x}, \boldsymbol{v}, m) \in \mathbb{R}^d \times V \times \mathbb{R}$, it yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\| |\boldsymbol{v}|^2 q^\varepsilon(t) \right\|_{L^1} \; = \; \int |\boldsymbol{v}|^2 \left(\partial_m \left[\frac{m-M}{\varepsilon} p^\varepsilon + \varepsilon \, \partial_m p^\varepsilon \right] - \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} p^\varepsilon + Q^\varepsilon [p^\varepsilon] \right) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}m \, .$$

Let us simplify the latter right-hand side: both contributions of the transport and the Fokker-Planck operators cancel since it holds

$$\int |\boldsymbol{v}|^2 \, \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} p^\varepsilon \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}m \; = \; \int \nabla_{\boldsymbol{x}} \cdot \left(|\boldsymbol{v}|^2 \boldsymbol{v} \, p^\varepsilon \right) \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}m \; = \; 0 \,,$$

and

$$\int |\boldsymbol{v}|^2 \, \partial_m \left[\frac{m-M}{\varepsilon} p^\varepsilon + \varepsilon \, \partial_m p^\varepsilon \right] \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}m = \int \partial_m \left[|\boldsymbol{v}|^2 \left(\frac{m-M}{\varepsilon} p^\varepsilon + \varepsilon \, \partial_m p^\varepsilon \right) \right] \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}m \; = \; 0 \,.$$

Hence, the time derivative of $\left\| |\boldsymbol{v}|^2 q^\varepsilon \right\|_{L^1}$ rewrites

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\| |\boldsymbol{v}|^2 q^\varepsilon(t) \right\|_{L^1} \; = \; \int |\boldsymbol{v}|^2 \, \Lambda \left(\frac{1}{\varepsilon} \left(m - M(t,\boldsymbol{x}) \right), \boldsymbol{v}, \boldsymbol{v}' \right) p^\varepsilon \left(t, \boldsymbol{x}, \boldsymbol{v}', m \right) \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}\boldsymbol{v}' \mathrm{d}m$$

$$- \int |\boldsymbol{v}|^2 \, \Lambda \left(\frac{1}{\varepsilon} \left(m - M(t,\boldsymbol{x}) \right), \boldsymbol{v}', \boldsymbol{v} \right) p^\varepsilon \left(t, \boldsymbol{x}, \boldsymbol{v}, m \right) \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}\boldsymbol{v}' \mathrm{d}m \,.$$

Since the second term in the latter right-hand side is negative, we bound it by 0 and hence deduce

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\| |\boldsymbol{v}|^2 q^\varepsilon(t) \right\|_{L^1} \; \leq \; \int |\boldsymbol{v}|^2 \, \Lambda \left(\frac{1}{\varepsilon} \left(m - M(t,\boldsymbol{x}) \right), \boldsymbol{v}, \boldsymbol{v}' \right) p^\varepsilon \left(t, \boldsymbol{x}, \boldsymbol{v}', m \right) \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}\boldsymbol{v}' \mathrm{d}m \,.$$

Then, we estimate the supremum of $\int |\boldsymbol{v}|^2 \, \Lambda \left(\frac{1}{\varepsilon} \left(m - M(t,\boldsymbol{x}) \right), \boldsymbol{v}, \boldsymbol{v}' \right) \mathrm{d}\boldsymbol{v}$ over all $(\boldsymbol{x}, \boldsymbol{v}', m) \in \mathbb{R}^d \times V \times \mathbb{R}$ thanks to assumption (2.2a). Thanks to conservation of mass for (1.4), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\| |\boldsymbol{v}|^2 q^\varepsilon(t) \right\|_{L^1} \; \leq \; C \,.$$

We integrate the latter inequality with respect to t and use assumption (2.3) to bound $\left\| |\boldsymbol{v}|^2 q_0^\varepsilon \right\|_{L^1}$, which yields estimate (3.4):

$$\left\| |\boldsymbol{v}|^2 q^\varepsilon(t) \right\|_{L^1} \; \leq \; C(t+1) \,.$$

The proof of estimate (3.8) for moments in $\boldsymbol{v}$ of $\bar{p}$ follows the same lines as the one detailed above for q^ε. Hence we do not detail it here.

We now proceed to the proof of (3.5). We denote by $\left\| |\boldsymbol{x}| \, q^\varepsilon(t) \right\|_{L^1}$ the first order moments of q^ε with respect to $\boldsymbol{x}$, that is

$$\left\| |\boldsymbol{x}| \, q^\varepsilon(t) \right\|_{L^1} \; = \; \int_{\mathbb{R}^d \times V \times \mathbb{R}} |\boldsymbol{x}| \, q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y) \, \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}y \,.$$

To estimate $\left\| |\boldsymbol{x}| \, q^\varepsilon \right\|_{L^1}$, we compute its time derivative. Again, q^ε and the solution p^ε to (1.4) have same moments with respect to $\boldsymbol{x}$. Indeed, inverting (1.7), we obtain

$$\left\| |\boldsymbol{x}| \, q^\varepsilon(t) \right\|_{L^1} \; = \; \int_{\mathbb{R}^d \times V \times \mathbb{R}} |\boldsymbol{x}| \, p^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, m) \, \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}m \,.$$

Therefore the time derivative of $\left\| |\boldsymbol{x}| \, q^\varepsilon \right\|_{L^1}$ is obtained multiplying equation (1.4) by $|\boldsymbol{x}|$ and integrating with respect $(\boldsymbol{x}, \boldsymbol{v}, m) \in \mathbb{R}^d \times V \times \mathbb{R}$, it yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\| |\boldsymbol{x}| \, q^\varepsilon(t) \right\|_{L^1} \; = \; \int |\boldsymbol{x}| \left(\partial_m \left[\frac{m-M}{\varepsilon} p^\varepsilon + \varepsilon \, \partial_m p^\varepsilon \right] - \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} p^\varepsilon + Q^\varepsilon \left[p^\varepsilon \right] \right) \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}m \,.$$

To simplify the latter right-hand side, we first notice that, as for moments with respect to the velocity variable, the contribution the Fokker-Planck operator cancels. The same holds for the contribution of the tumbling operator since $\int Q^\varepsilon \left[p^\varepsilon \right] \mathrm{d}\boldsymbol{v} = 0$. Hence, the latter relation rewrites

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\| |\boldsymbol{x}| \, q^\varepsilon(t) \right\|_{L^1} \; = \; - \int |\boldsymbol{x}| \, \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} p^\varepsilon \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}m \,.$$

After integrating by part with respect to $\boldsymbol{x}$ in the latter integral, applying the Cauchy-Schwarz inequality and using mass conservation for (1.4), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\| |\boldsymbol{x}| \, q^\varepsilon(t) \right\|_{L^1} \; \leq \; \left(\int |\boldsymbol{v}|^2 p^\varepsilon \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}m \right)^{\frac{1}{2}} \,.$$

We bound the latter right-hand side thanks to (3.4) and deduce

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\| |\boldsymbol{x}| \, q^\varepsilon(t) \right\|_{L^1} \leq C \, (1+t)^{\frac{1}{2}} \, .$$

Estimate (3.5) is obtained integrating the latter with respect to t and applying assumption (2.3) to bound $\left\| |\boldsymbol{x}| \, q_0^\varepsilon \right\|_{L^1}$.

We now derive (3.6). We denote by $\left\| |y|^2 q^\varepsilon(t) \right\|_{L^1}$ the second order moments of q^ε with respect to y, that is

$$\left\| |y|^2 q^\varepsilon(t) \right\|_{L^1} = \int_{\mathbb{R}^d \times V \times \mathbb{R}} |y|^2 q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y) \, \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}y \, .$$

To estimate $\left\| |y|^2 q^\varepsilon(t) \right\|_{L^1}$, we compute its time derivative, obtained multiplying equation (1.9) by $|y|^2$ and integrating with respect $(\boldsymbol{x}, \boldsymbol{v}, y) \in \mathbb{R}^d \times V \times \mathbb{R}$, it yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\| |y|^2 q^\varepsilon(t) \right\|_{L^1} = \int |y|^2 \left(\frac{1}{\varepsilon} \partial_y \left[y \, q^\varepsilon + \partial_y q^\varepsilon \right] - \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} q^\varepsilon + \tilde{Q}^\varepsilon \left[q^\varepsilon \right] \right) \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}y \, .$$

Let us simplify the latter right-hand side: first, the transport operator has no contribution, that is

$$\int |y|^2 \, \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} q^\varepsilon \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}y = \int \nabla_{\boldsymbol{x}} \cdot \left(|y|^2 \, \boldsymbol{v} \, q^\varepsilon \right) \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}y = 0 \, .$$

Second, we integrate by part with respect to y the contribution associated to the Fokker-Planck operator. Since q^ε as mass 1, it yields

$$\int |y|^2 \, \partial_y \left[y \, q^\varepsilon + \partial_y q^\varepsilon \right] \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}y = -2 \left\| |y|^2 q^\varepsilon(t) \right\|_{L^1} + 2 \, .$$

Therefore the time derivative of $\left\| |y|^2 q^\varepsilon \right\|_{L^1}$ rewrites

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\| |y|^2 q^\varepsilon(t) \right\|_{L^1} = \frac{2}{\varepsilon} \left(1 - \left\| |y|^2 q^\varepsilon(t) \right\|_{L^1} \right) + \mathcal{J} \, ,$$

where $\mathcal{J}$ denotes the contribution of the tumbling operator, that is

(3.9) $$\mathcal{J} = \int |y|^2 \, \tilde{Q}^\varepsilon \left[q^\varepsilon \right] \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}y \, .$$

To estimate $\mathcal{J}$, we rewrite it as $\mathcal{J} = \mathcal{J}_1 - \mathcal{J}_2$ where $\mathcal{J}_1$ and $\mathcal{J}_2$ are given by

$$\begin{cases} \mathcal{J}_1 = \int |y|^2 \, \Lambda \left(y' - D_t N(t, \boldsymbol{x}, \boldsymbol{v}'), \boldsymbol{v}, \boldsymbol{v}' \right) q^\varepsilon \left(t, \boldsymbol{x}, \boldsymbol{v}', y' \right) \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}\boldsymbol{v}' \mathrm{d}y \, , \\[3mm] \mathcal{J}_2 = \int |y|^2 \, \Lambda \left(y - D_t N(t, \boldsymbol{x}, \boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v} \right) q^\varepsilon \left(t, \boldsymbol{x}, \boldsymbol{v}, y \right) \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}\boldsymbol{v}' \mathrm{d}y \, , \end{cases}$$

and where we used the shorthand notation

$$y' = y + \left(N(t, \boldsymbol{x}, \boldsymbol{v}) - N(t, \boldsymbol{x}, \boldsymbol{v}') \right) / \varepsilon \, .$$

In $\mathcal{J}_1$, we perform the change of variable $y \leftarrow y'$ and then invert variables $\boldsymbol{v}$ and $\boldsymbol{v}'$, this yields

$$\mathcal{J}_1 = \int \left| y + \frac{N(t, \boldsymbol{x}, \boldsymbol{v}) - N(t, \boldsymbol{x}, \boldsymbol{v}')}{\varepsilon} \right|^2 \Lambda \left(y - D_t N(t, \boldsymbol{x}, \boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v} \right) q^\varepsilon \left(t, \boldsymbol{x}, \boldsymbol{v}, y \right) \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}\boldsymbol{v}' \mathrm{d}y \, .$$

We rewrite $\mathcal{J}$ substituting $\mathcal{J}_1$ thanks to the latter relation, this yields

$$\mathcal{J} = \int \left(\left| y + \frac{N(t, \boldsymbol{x}, \boldsymbol{v}) - N(t, \boldsymbol{x}, \boldsymbol{v}')}{\varepsilon} \right|^2 - |y|^2 \right) \Lambda \left(y - D_t N(t, \boldsymbol{x}, \boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v} \right) q^\varepsilon \left(t, \boldsymbol{x}, \boldsymbol{v}, y \right) \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}\boldsymbol{v}' \mathrm{d}y.$$

10

To estimate $\mathcal{J}$, we take the absolute value inside the latter integral and bound the difference between squares as follows

$$\left| \left| y + \frac{N(t,\boldsymbol{x},\boldsymbol{v}) - N(t,\boldsymbol{x},\boldsymbol{v}')}{\varepsilon} \right|^2 - |y|^2 \right| \leq 2\,|y|\,\left| \frac{N(t,\boldsymbol{x},\boldsymbol{v}) - N(t,\boldsymbol{x},\boldsymbol{v}')}{\varepsilon} \right| + \left| \frac{N(t,\boldsymbol{x},\boldsymbol{v}) - N(t,\boldsymbol{x},\boldsymbol{v}')}{\varepsilon} \right|^2$$

$$\leq C\left(|y|\,|\boldsymbol{v} - \boldsymbol{v}'| + |\boldsymbol{v} - \boldsymbol{v}'|^2\right),$$

where the last line is obtained thanks to (3.1) in Lemma 3.1 to estimate N. Hence, we obtain

$$\mathcal{J} \leq C \int \left(|y|\,|\boldsymbol{v} - \boldsymbol{v}'| + |\boldsymbol{v} - \boldsymbol{v}'|^2\right) \Lambda\left(y - D_t N(t,\boldsymbol{x},\boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v}\right) q^\varepsilon(t,\boldsymbol{x},\boldsymbol{v},y)\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{v}'\mathrm{d}y.$$

In the latter integral, we apply the Cauchy-Schwarz inequality to estimate the cross product $|y|\,|\boldsymbol{v} - \boldsymbol{v}'|$ and bound $|\boldsymbol{v} - \boldsymbol{v}'|^2$ by $2(|\boldsymbol{v}|^2 + |\boldsymbol{v}'|^2)$, it yields

$$\mathcal{J} \leq C\left(\mathcal{J}_3^{\frac{1}{2}}\,\mathcal{J}_4^{\frac{1}{2}} + \mathcal{J}_4\right),$$

where $\mathcal{J}_3$ and $\mathcal{J}_4$ are defined as

$$\begin{cases} \mathcal{J}_3 = \int |y|^2 \Lambda\left(y - D_t N(t,\boldsymbol{x},\boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v}\right) q^\varepsilon(t,\boldsymbol{x},\boldsymbol{v},y)\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{v}'\mathrm{d}y, \\[2mm] \mathcal{J}_4 = \int \left(|\boldsymbol{v}|^2 + |\boldsymbol{v}'|^2\right) \Lambda\left(y - D_t N(t,\boldsymbol{x},\boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v}\right) q^\varepsilon(t,\boldsymbol{x},\boldsymbol{v},y)\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{v}'\mathrm{d}y. \end{cases}$$

In both $\mathcal{J}_3$ and $\mathcal{J}_4$, we integrate Λ with respect to $\boldsymbol{v}'$ and estimate $\int (1+|\boldsymbol{v}'|^2)\Lambda\left(y - D_t N(t,\boldsymbol{x},\boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v}\right)\mathrm{d}\boldsymbol{v}'$ by its supremum over all $(\boldsymbol{x},\boldsymbol{v},y) \in \mathbb{R}^d \times V \times \mathbb{R}$, which is finite according to assumption (2.2a). Since the total mass of q^ε is conserved, we deduce

$$\mathcal{J} \leq C\left(\left(\int |y|^2 q^\varepsilon \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y\right)^{\frac{1}{2}} \left(\int |\boldsymbol{v}|^2 q^\varepsilon \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y + 1\right)^{\frac{1}{2}} + \int |\boldsymbol{v}|^2 q^\varepsilon \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y + 1\right).$$

We bound moments with respect to $\boldsymbol{v}$ thanks to (3.4) in Proposition 3.2 and deduce the following estimate for $\mathcal{J}$

$$(3.10) \qquad \mathcal{J} \leq C\left((t+1)^{\frac{1}{2}}\left(\int |y|^2 q^\varepsilon \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y\right)^{\frac{1}{2}} + t + 1\right).$$

After applying Young inequality, we obtain

$$\mathcal{J} \leq C\left(\left\||y|^2 q^\varepsilon(t)\right\|_{L^1} + t + 1\right).$$

Therefore, we get the following differential inequality for $\left\||y|^2 q^\varepsilon\right\|_{L^1}$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\||y|^2 q^\varepsilon(t)\right\|_{L^1} \leq \left(C - \frac{2}{\varepsilon}\right)\left\||y|^2 q^\varepsilon(t)\right\|_{L^1} + \frac{1}{\varepsilon} + C(t+1).$$

We multiply this inequality by $\exp\left((2/\varepsilon - C)\,t\right)$ and integrate with respect to the time variable, which yields

$$\left\||y|^2 q^\varepsilon(t)\right\|_{L^1} \leq \left\||y|^2 q_0^\varepsilon\right\|_{L^1} e^{(C - \frac{2}{\varepsilon})t} + \frac{C\varepsilon + 1}{2 - C\varepsilon}\left(1 - e^{(C - \frac{2}{\varepsilon})t}\right) + C\int_0^t s\,e^{(C - \frac{2}{\varepsilon})(t - s)}\,\mathrm{d}s.$$

We compute the time integral on the right-hand side thanks to an integration by part and obtain

$$\left\||y|^2 q^\varepsilon(t)\right\|_{L^1} \leq \left\||y|^2 q_0^\varepsilon\right\|_{L^1} e^{(C - \frac{2}{\varepsilon})t} + \frac{C\varepsilon + 1}{2 - C\varepsilon}\left(1 - e^{(C - \frac{2}{\varepsilon})t}\right)$$

$$+ \frac{C\varepsilon\,t}{2 - C\varepsilon} - \frac{C\varepsilon^2}{(2 - C\varepsilon)^2}\left(1 - e^{(C - \frac{2}{\varepsilon})t}\right).$$

Taking $\varepsilon \leq 1/C$, we obtain

$$\left\||y|^2 q^\varepsilon(t)\right\|_{L^1} \leq \left\||y|^2 q_0^\varepsilon\right\|_{L^1} e^{-\frac{t}{\varepsilon}} + 2 + C\varepsilon\,t.$$

To estimate $\left\||y|^2 q_0^\varepsilon\right\|_{L^1}$, we invert the change of variable (1.7) with $t = 0$ and obtain

$$\left\||y|^2 q^\varepsilon(t)\right\|_{L^1} \leq e^{-\frac{t}{\varepsilon}} \int_{\mathbb{R}^d \times V \times \mathbb{R}} \left|\frac{m - M(0, \boldsymbol{x})}{\varepsilon}\right|^2 p_0^\varepsilon(\boldsymbol{x}, \boldsymbol{v}, m) \, \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}m + 2 + C\varepsilon t.$$

We estimate the integral on the right-hand side thanks to assumption (2.3) to bound the moments of p^ε with respect to m, and assumption (2.1) to bound $M(0, \boldsymbol{x})$. This yields the expected result

$$\int_{\mathbb{R}^d \times V \times \mathbb{R}} |y|^2 q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y) \, \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y \leq C \left(\varepsilon^{-2} e^{-\frac{t}{\varepsilon}} + \varepsilon t + 1\right).$$

To conclude this proof, we derive estimate (3.7) as a straightforward consequence of our previous computations. Indeed, according to (3.10), it holds

$$\mathcal{J} \leq C \left((t+1)^{\frac{1}{2}} \left(\int |y|^2 q^\varepsilon \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y\right)^{\frac{1}{2}} + t + 1\right),$$

where $\mathcal{J}$ is defined by (3.9). In the latter integral, we bound $\int |y|^2 q^\varepsilon \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y$ thanks to (3.6) in Proposition 3.2, which yields

$$\mathcal{J} \leq C \left((t+1)^{\frac{1}{2}} \left(\varepsilon^{-2} e^{-\frac{t}{\varepsilon}} + \varepsilon t + 1\right)^{\frac{1}{2}} + t + 1\right).$$

Then, simple computations yield the result

$$\mathcal{J} \leq C \left(t^{\frac{1}{2}} \varepsilon^{-1} e^{-\frac{t}{2\varepsilon}} + \varepsilon^{-1} e^{-\frac{t}{2\varepsilon}} + t + 1\right).$$

$\square$

4. Convergence of the re-scaled distribution

This section is dedicated to the convergence of q^ε defined by (1.6) toward $\bar{p}\mathcal{M}$, where $\bar{p}$ is the solution to (1.1)-(1.3) and $\mathcal{M}$ stands for the standard Maxwellian and is given in (1.10). We proceed in two steps: first, we prove the convergence of q^ε towards its Maxwellian local equilibrium, that is

$$q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y) \underset{\varepsilon \to 0}{\sim} \bar{q}^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v})\mathcal{M}(y),$$

where the marginal $\bar{q}^\varepsilon$ is given by (1.11). Second, we use the latter convergence to deduce the convergence of $\bar{q}^\varepsilon$ toward $\bar{p}$.

To derive the convergence of q^ε towards its Maxwellian local equilibrium, we estimate the following relative energy along the trajectories of equation (1.9)

$$E[q^\varepsilon] = \int_{\mathbb{R}^d \times V \times \mathbb{R}} q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y) \log\left(\frac{q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y)}{\mathcal{M}(y)}\right) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y,$$

where $\mathcal{M}$ is given by (1.10). We prove that the dissipation rate of this quantity controls the distance between q^ε and $\bar{q}^\varepsilon \mathcal{M}$. We obtain the following convergence estimate:

Proposition 4.1. *Under assumptions* (2.2a)-(2.2b) *on* Λ, (2.1) *on* M *and* (2.3) *on* $(p_0^\varepsilon)_{\varepsilon>0}$ *consider a sequence of strong solutions* $(p^\varepsilon)_{\varepsilon>0}$ *to* (1.4)-(1.5) *with initial data* $(p_0^\varepsilon)_{\varepsilon>0}$.
There exist two positive constants C *and* ε_0 *such that for all* ε *between* 0 *and* ε_0, *it holds*

$$\int_0^t \| q^\varepsilon(s) - \bar{q}^\varepsilon(s)\mathcal{M} \|_{L^1(\mathbb{R}^d \times V \times \mathbb{R})}^2 \, \mathrm{d}s \leq \varepsilon E[q_0^\varepsilon] + C\varepsilon\left(t^2 + 1\right), \quad \forall t \in \mathbb{R}^+.$$

Proof. To compute the time derivative of the relative energy $E[q^\varepsilon]$, we multiply equation (1.9) by $\log(q^\varepsilon/\mathcal{M})$ and integrate with respect to $\boldsymbol{x}$, $\boldsymbol{v}$ and y. Since mass is conserved along the trajectories of (1.9), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} E[q^\varepsilon] = \int \log\left(\frac{q^\varepsilon}{\mathcal{M}}\right) \left(\frac{1}{\varepsilon} \partial_y \left[y \, q^\varepsilon + \partial_y q^\varepsilon\right] - \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} q^\varepsilon + \tilde{Q}^\varepsilon[q^\varepsilon]\right) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y.$$

We first simplify the latter right-hand side: since $\mathcal{M}$ does not depend on $\boldsymbol{x}$, the transport operator has no contribution, that is

$$\int \log\left(\frac{q^\varepsilon}{\mathcal{M}}\right) \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} q^\varepsilon \,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y = 0\,.$$

Furthermore, the Fokker-Planck operator has a signed contribution. Indeed, an integration by part yields

$$\int \log\left(\frac{q^\varepsilon}{\mathcal{M}}\right) \partial_y \left[y\, q^\varepsilon + \partial_y q^\varepsilon\right] \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y = -I[q^\varepsilon]\,,$$

where the Fisher information is defined as follows

$$I[q^\varepsilon] = \int \left|\partial_y \log\left(\frac{q^\varepsilon}{\mathcal{M}}\right)\right|^2 q^\varepsilon \,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y\,.$$

Consequent to the latter two remarks, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} E[q^\varepsilon] + \frac{1}{\varepsilon} I[q^\varepsilon] = \int \log\left(\frac{q^\varepsilon}{\mathcal{M}}\right) \tilde{Q}^\varepsilon\,[q^\varepsilon] \,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y\,.$$

To estimate the contribution of the tumbling operator, we decompose it as follows

$$\int \log\left(\frac{q^\varepsilon}{\mathcal{M}}\right) \tilde{Q}^\varepsilon\,[q^\varepsilon] \,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y = \frac{1}{2}\mathcal{J} + \mathcal{K}\,,$$

where $\mathcal{J}$ and $\mathcal{K}$ are given by

$$\begin{cases} \mathcal{J} = \displaystyle\int |y|^2\, \tilde{Q}^\varepsilon\,[q^\varepsilon] \,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y\,, \\[2mm] \mathcal{K} = \displaystyle\int \log\left(q^\varepsilon\right) \tilde{Q}^\varepsilon\,[q^\varepsilon] \,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y\,. \end{cases}$$

We use dissipation properties of the tumbling kernel to estimate $\mathcal{K}$, whereas $\mathcal{J}$ was already estimated in Proposition 3.2, estimate (3.7).

To estimate $\mathcal{K}$, we decompose $\tilde{Q}^\varepsilon$, defined in (1.9), into gain and loss terms, that is

$$\mathcal{K} = \mathcal{K}_1 - \mathcal{K}_2\,,$$

where $\mathcal{K}_1$ and $\mathcal{K}_2$ are given by

$$\begin{cases} \mathcal{K}_1 = \displaystyle\int \Lambda\left(y' - D_t N(t,\boldsymbol{x},\boldsymbol{v}'),\boldsymbol{v},\boldsymbol{v}'\right) q^\varepsilon\left(t,\boldsymbol{x},\boldsymbol{v}',y'\right) \log\left(q^\varepsilon\left(t,\boldsymbol{x},\boldsymbol{v},y\right)\right)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{v}'\mathrm{d}y\,, \\[2mm] \mathcal{K}_2 = \displaystyle\int \Lambda\left(y - D_t N(t,\boldsymbol{x},\boldsymbol{v}),\boldsymbol{v}',\boldsymbol{v}\right) q^\varepsilon\left(t,\boldsymbol{x},\boldsymbol{v},y\right) \log\left(q^\varepsilon\left(t,\boldsymbol{x},\boldsymbol{v},y\right)\right)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{v}'\mathrm{d}y\,, \end{cases}$$

and where we used the shorthand notation

$$y' = y + \left(N(t,\boldsymbol{x},\boldsymbol{v}) - N(t,\boldsymbol{x},\boldsymbol{v}')\right)/\varepsilon\,.$$

In $\mathcal{K}_2$, we perform the change of variable $y \leftarrow y + (N(t,\boldsymbol{x},\boldsymbol{v}) - N(t,\boldsymbol{x},\boldsymbol{v}'))/\varepsilon$ and then, we invert $\boldsymbol{v}$ and $\boldsymbol{v}'$, this yields

$$\mathcal{K}_2 = \int \Lambda\left(y' - D_t N(t,\boldsymbol{x},\boldsymbol{v}'),\boldsymbol{v},\boldsymbol{v}'\right) q^\varepsilon\left(t,\boldsymbol{x},\boldsymbol{v}',y'\right) \log\left(q^\varepsilon\left(t,\boldsymbol{x},\boldsymbol{v}',y'\right)\right)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{v}'\mathrm{d}y\,.$$

According to the latter relation, $\mathcal{K}$ rewrites as follows

$$\mathcal{K} = \int \Lambda\left(y' - D_t N(t,\boldsymbol{x},\boldsymbol{v}'),\boldsymbol{v},\boldsymbol{v}'\right) q^\varepsilon\left(t,\boldsymbol{x},\boldsymbol{v}',y'\right) \log\left(\frac{q^\varepsilon\left(t,\boldsymbol{x},\boldsymbol{v},y\right)}{q^\varepsilon\left(t,\boldsymbol{x},\boldsymbol{v}',y'\right)}\right)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{v}'\mathrm{d}y\,.$$

We estimate the logarithm in the latter right-hand side thanks to the inequality $\log\left(x\right) \leq x$, this yields

$$\mathcal{K} \leq \int \Lambda\left(y - D_t N(t,\boldsymbol{x},\boldsymbol{v}),\boldsymbol{v},\boldsymbol{v}'\right) q^\varepsilon\left(t,\boldsymbol{x},\boldsymbol{v},y\right) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{v}'\mathrm{d}y\,,$$

where we used that $y' - D_t N(t, \boldsymbol{x}, \boldsymbol{v}') = y - D_t N(t, \boldsymbol{x}, \boldsymbol{v})$. Thanks to assumption (2.2b), we estimate $\int \Lambda (y - D_t N(t, \boldsymbol{x}, \boldsymbol{v}), \boldsymbol{v}, \boldsymbol{v}') \, \mathrm{d}\boldsymbol{v}'$ by its supremum over all $(\boldsymbol{x}, \boldsymbol{v}, y) \in \mathbb{R}^d \times V \times \mathbb{R}$. Since the mass of q^ε is conserved, this yields

$$\mathcal{K} \leq C.$$

To estimate $\mathcal{J}$, we apply (3.7), which ensures

$$\mathcal{J} \leq C \left(t^{\frac{1}{2}} \varepsilon^{-1} e^{-\frac{t}{2\varepsilon}} + \varepsilon^{-1} e^{-\frac{t}{2\varepsilon}} + t + 1 \right).$$

Gathering estimates for $\mathcal{J}$ and $\mathcal{K}$, we deduce

$$\frac{\mathrm{d}}{\mathrm{d}t} E[q^\varepsilon] + \frac{1}{\varepsilon} I[q^\varepsilon] \leq C \left(t^{\frac{1}{2}} \varepsilon^{-1} e^{-\frac{t}{2\varepsilon}} + \varepsilon^{-1} e^{-\frac{t}{2\varepsilon}} + t + 1 \right).$$

Integrating the latter inequality between 0 and t, we obtain

$$\frac{1}{\varepsilon} \int_0^t I[q^\varepsilon(s)] \, \mathrm{d}s \leq E[q_0^\varepsilon] - E[q^\varepsilon(t)] + C \left(t^2 + 1 \right),$$

where we used

$$\varepsilon^{-1} \int_0^t s^{\frac{1}{2}} e^{-\frac{s}{2\varepsilon}} \, \mathrm{d}s \leq \sqrt{\varepsilon} \int_0^{+\infty} \sigma^{\frac{1}{2}} e^{-\frac{\sigma}{2}} \, \mathrm{d}s \leq C.$$

On the one hand, the Fisher information $I[q^\varepsilon(s)]$ controls the L^1-norm between $q^\varepsilon(s)$ and $\bar{q}^\varepsilon(s)\mathcal{M}$ since we have

$$\| q^\varepsilon(s) - \bar{q}^\varepsilon(s) \mathcal{M} \|^2_{L^1 \left(\mathbb{R}^d \times V \times \mathbb{R} \right)} \leq 2 \int_{\mathbb{R}^d \times V \times \mathbb{R}} q^\varepsilon(s) \log \left(\frac{q^\varepsilon(s)}{\bar{q}^\varepsilon(s) \mathcal{M}} \right) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y \leq I \left[q^\varepsilon(s) \right],$$

where first and second inequalities correspond respectively to Csizár-Kullback and Gaussian logarithmic Sobolev (see [28]) inequalities. On the other hand, we have

$$- E[q^\varepsilon(t)] = - \int q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y) \left(\log \left(\frac{q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y)}{\mathcal{M}(y)\mathcal{M}(|\boldsymbol{v}|)\mathcal{M}(|\boldsymbol{x}|^{\frac{1}{2}})} \right) - \frac{1}{2}|\boldsymbol{v}|^2 - \frac{1}{2}|\boldsymbol{x}| - \log(2\pi) \right) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y$$

$$\leq \int q^\varepsilon(t, \boldsymbol{x}, \boldsymbol{v}, y) \left(\log \left(\int \mathcal{M}(|\boldsymbol{v}|)\mathcal{M}(|\boldsymbol{x}|^{\frac{1}{2}}) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v} \right) + \frac{1}{2}|\boldsymbol{v}|^2 + \frac{1}{2}|\boldsymbol{x}| + \log(2\pi) \right) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y$$

$$\leq C \left(t^{\frac{3}{2}} + 1 \right),$$

where we used Jensen inequality with the measure $\mathcal{M}(y)\mathcal{M}(|\boldsymbol{v}|)\mathcal{M}(|\boldsymbol{x}|^{\frac{1}{2}})$ to obtain the second line and estimates (3.4)-(3.5) to bound moments at the last line. Gathering these two estimates, we obtain the result

$$\int_0^t \| q^\varepsilon(s) - \bar{q}^\varepsilon(s) \mathcal{M} \|^2_{L^1 \left(\mathbb{R}^d \times V \times \mathbb{R} \right)} \mathrm{d}s \leq \varepsilon E[q_0^\varepsilon] + C \varepsilon \left(t^2 + 1 \right).$$

$\square$

To complete this section, we estimate the distance between $\bar{q}^\varepsilon$ given in (1.11) and the solution $\bar{p}$ to limiting model (1.1)-(1.3). The key point consists in taking advantage of the dissipation properties of the limiting tumbling kernel Q, which has a signed contribution to the error between $\bar{q}^\varepsilon$ and $\bar{p}$. Then, remaining contributions are error terms that we estimate thanks to our previous convergence estimates for N towards M and q^ε towards $\bar{q}^\varepsilon \mathcal{M}$.

Proposition 4.2. *Under assumptions* (2.2a)-(2.2b) *on* Λ, (2.1) *on* M, (2.3) *on* $(p_0^\varepsilon)_{\varepsilon > 0}$ *and* (2.6) *on* $\bar{p}_0$, *consider a sequence of strong solutions* $(p^\varepsilon)_{\varepsilon > 0}$ *to* (1.4)-(1.5) *with initial data* $(p_0^\varepsilon)_{\varepsilon > 0}$ *and a strong solution* $\bar{p}$ *to* (1.1)-(1.3) *with initial data* $\bar{p}_0$.
There exist two positive constants C *and* ε_0 *such that for all* ε *between 0 and* ε_0, *it holds*

$$\| \bar{p}(t) - \bar{q}^\varepsilon(t) \|_{L^1} \leq \| \bar{p}_0 - \bar{q}_0^\varepsilon \|_{L^1} + C \left(\sqrt{\varepsilon \, t} \, E[q_0^\varepsilon]^{\frac{1}{2}} + \sqrt{\varepsilon \, t} \, (t + 1) + \varepsilon^2 \right), \quad \forall t \in \mathbb{R}^+.$$

Proof. To simplify notations, we omit the dependence with respect to $(t, \boldsymbol{x}) \in \mathbb{R}^+ \times \mathbb{R}^d$ when the context is clear. To estimate $\|\bar{p}(t) - \bar{q}^\varepsilon(t)\|_{L^1(\mathbb{R}^d \times V)}$, we first compute the equation solved by $\bar{q}^\varepsilon$, which is obtained integrating (1.9) with respect to $y \in \mathbb{R}$, that is

(4.1)
$$\frac{\partial \bar{q}^\varepsilon}{\partial t} + \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} \bar{q}^\varepsilon = \int_{\mathbb{R}} \tilde{Q}^\varepsilon [q^\varepsilon] \, \mathrm{d}y \,,$$

where

$$\int_{\mathbb{R}} \tilde{Q}^\varepsilon [q^\varepsilon] \, \mathrm{d}y = \int \Lambda \left(y' - D_t N(\boldsymbol{v}'), \boldsymbol{v}, \boldsymbol{v}' \right) q^\varepsilon \left(\boldsymbol{v}', y' \right) - \Lambda \left(y - D_t N(\boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v} \right) q^\varepsilon \left(\boldsymbol{v}, y \right) \mathrm{d}\boldsymbol{v}' \mathrm{d}y \,,$$

and where we used the shorthand notation

$$y' = y + \left(N(t, \boldsymbol{x}, \boldsymbol{v}) - N(t, \boldsymbol{x}, \boldsymbol{v}') \right) / \varepsilon \,.$$

We perform the change of variable $y \leftarrow y + \left(N(t, \boldsymbol{x}, \boldsymbol{v}) - N(t, \boldsymbol{x}, \boldsymbol{v}') \right) / \varepsilon$ in the first term of the latter right-hand side and deduce

$$\int_{\mathbb{R}} \tilde{Q}^\varepsilon [q^\varepsilon] \, \mathrm{d}y = \int \Lambda \left(y - D_t N(\boldsymbol{v}'), \boldsymbol{v}, \boldsymbol{v}' \right) q^\varepsilon \left(\boldsymbol{v}', y \right) - \Lambda \left(y - D_t N(\boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v} \right) q^\varepsilon \left(\boldsymbol{v}, y \right) \mathrm{d}\boldsymbol{v}' \mathrm{d}y \,.$$

Then, we compute the time derivative of $\|\bar{p}(t) - \bar{q}^\varepsilon(t)\|_{L^1(\mathbb{R}^d \times V)}$ by multiplying the difference between (1.1) and (4.1) by $\mathrm{sgn}(\bar{p}(t) - \bar{q}^\varepsilon(t))$ and integrating with respect to $(\boldsymbol{x}, \boldsymbol{v}) \in \mathbb{R}^d \times V$; we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} \|\bar{p} - \bar{q}^\varepsilon\|_{L^1(\mathbb{R}^d \times V)} = \int_{\mathbb{R}^d \times V} \mathrm{sgn}(\bar{p} - \bar{q}^\varepsilon) \left(Q[\bar{p}] - \int_{\mathbb{R}} \tilde{Q}^\varepsilon [q^\varepsilon] \, \mathrm{d}y \right) \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \,.$$

We decompose the right-hand side as follows

$$\frac{\mathrm{d}}{\mathrm{d}t} \|\bar{p} - \bar{q}^\varepsilon\|_{L^1(\mathbb{R}^d \times V)} = \mathcal{L}_{11} - \mathcal{L}_{12} + \mathcal{L}_{21} - \mathcal{L}_{22} + \mathcal{L}_{31} - \mathcal{L}_{32} \,,$$

and carefully detail each contribution. The first term $\mathcal{L}_{11} - \mathcal{L}_{12}$ corresponds to the error due to the difference between the external signal M and N defined in (1.8), that is

$$\begin{cases} \mathcal{L}_{11} = \int \mathrm{sgn}(\bar{p} - \bar{q}^\varepsilon)(\boldsymbol{v}) \left(\Lambda \left(y - D_t M(\boldsymbol{v}'), \boldsymbol{v}, \boldsymbol{v}' \right) - \Lambda \left(y - D_t N(\boldsymbol{v}'), \boldsymbol{v}, \boldsymbol{v}' \right) \right) \bar{p} \left(\boldsymbol{v}' \right) \mathcal{M}(y) \, \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}\boldsymbol{v}' \mathrm{d}y \,, \\[2mm] \mathcal{L}_{12} = \int \mathrm{sgn}(\bar{p} - \bar{q}^\varepsilon)(\boldsymbol{v}) \left(\Lambda \left(y - D_t M(\boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v} \right) - \Lambda \left(y - D_t N(\boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v} \right) \right) \bar{p} \left(\boldsymbol{v} \right) \mathcal{M}(y) \, \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}\boldsymbol{v}' \mathrm{d}y \,, \end{cases}$$

we estimate $\mathcal{L}_{11}$ and $\mathcal{L}_{12}$ thanks to the convergence of N towards M ensured in Lemma 3.1. The second term $\mathcal{L}_{21} - \mathcal{L}_{22}$ corresponds to the propagation of the L^1-norm for the limiting model (1.1), that is

$$\begin{cases} \mathcal{L}_{21} = \int \mathrm{sgn}(\bar{p} - \bar{q}^\varepsilon)(\boldsymbol{v}) \Lambda \left(y - D_t N(\boldsymbol{v}'), \boldsymbol{v}, \boldsymbol{v}' \right) \left(\bar{p} - \bar{q}^\varepsilon \right) \left(\boldsymbol{v}' \right) \mathcal{M}(y) \, \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}\boldsymbol{v}' \mathrm{d}y \,, \\[2mm] \mathcal{L}_{22} = \int \mathrm{sgn}(\bar{p} - \bar{q}^\varepsilon)(\boldsymbol{v}) \Lambda \left(y - D_t N(\boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v} \right) \left(\bar{p} - \bar{q}^\varepsilon \right) \left(\boldsymbol{v} \right) \mathcal{M}(y) \, \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}\boldsymbol{v}' \mathrm{d}y \,, \end{cases}$$

we prove that $\mathcal{L}_{21} - \mathcal{L}_{22}$ is non-positive. The last term $\mathcal{L}_{31} - \mathcal{L}_{32}$ corresponds to the error due to the difference between q^ε and $\bar{q}^\varepsilon \mathcal{M}$, that is

$$\begin{cases} \mathcal{L}_{31} = \int \mathrm{sgn}(\bar{p} - \bar{q}^\varepsilon)(\boldsymbol{v}) \Lambda \left(y - D_t N(\boldsymbol{v}'), \boldsymbol{v}, \boldsymbol{v}' \right) \left(\bar{q}^\varepsilon \mathcal{M} - q^\varepsilon \right) \left(\boldsymbol{v}', y \right) \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}\boldsymbol{v}' \mathrm{d}y \,, \\[2mm] \mathcal{L}_{32} = \int \mathrm{sgn}(\bar{p} - \bar{q}^\varepsilon)(\boldsymbol{v}) \Lambda \left(y - D_t N(\boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v} \right) \left(\bar{q}^\varepsilon \mathcal{M} - q^\varepsilon \right) \left(\boldsymbol{v}, y \right) \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} \mathrm{d}\boldsymbol{v}' \mathrm{d}y \,, \end{cases}$$

we estimate $\mathcal{L}_{31} - \mathcal{L}_{32}$ thanks to the convergence of q^ε toward $\bar{q}^\varepsilon \mathcal{M}$ ensured by Proposition 4.1.

Taking the absolute value inside the integral in $\mathcal{L}_{21}$ we obtain

$$\mathcal{L}_{21} - \mathcal{L}_{22} \leq \int \Lambda\left(y - D_t N(\boldsymbol{v}'), \boldsymbol{v}, \boldsymbol{v}'\right)|\bar{p} - \bar{q}^\varepsilon|\,\mathcal{M}(\boldsymbol{v}', y)$$

$$- \Lambda\left(y - D_t N(\boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v}\right)|\bar{p} - \bar{q}^\varepsilon|\,\mathcal{M}(\boldsymbol{v}, y)\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{v}'\mathrm{d}y\,,$$

and then notice that the latter right-hand side cancels inverting $\boldsymbol{v}$ and $\boldsymbol{v}'$ in the first term, that is

$$\mathcal{L}_{21} - \mathcal{L}_{22} \leq 0\,.$$

To estimate $\mathcal{L}_{31} - \mathcal{L}_{32}$, we invert $\boldsymbol{v}$ and $\boldsymbol{v}'$ inside $\mathcal{L}_{31}$ and take the absolute value inside both integrals, this yields

$$\mathcal{L}_{31} - \mathcal{L}_{32} \leq 2\int \Lambda\left(y - D_t N(\boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v}\right)|\bar{q}^\varepsilon \mathcal{M} - q^\varepsilon|\,(\boldsymbol{v}, y)\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{v}'\mathrm{d}y\,.$$

Thanks to assumption (2.2a) we estimate $\int \Lambda\left(y - D_t N(t, \boldsymbol{x}, \boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v}\right)\mathrm{d}\boldsymbol{v}'$ by its supremum over all $(\boldsymbol{x}, \boldsymbol{v}, y) \in \mathbb{R}^d \times V \times \mathbb{R}$, this yields

$$\mathcal{L}_{31} - \mathcal{L}_{32} \leq C\,\|\,q^\varepsilon - \bar{q}^\varepsilon\,\mathcal{M}\,\|_{L^1\left(\mathbb{R}^d \times V \times \mathbb{R}\right)}\,.$$

We turn to the last two terms $\mathcal{L}_{11}$ and $\mathcal{L}_{12}$. We perform the change of variable $y \leftarrow y - D_t M(\boldsymbol{v}) + D_t N(\boldsymbol{v})$ in the first term composing $\mathcal{L}_{12}$, this yields

$$\mathcal{L}_{12} = \int \mathrm{sgn}(\bar{p} - \bar{q}^\varepsilon)(\boldsymbol{v})\Lambda\left(y - D_t N(\boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v}\right)\bar{p}(\boldsymbol{v})\left(\mathcal{M}\left(y + D_t M(\boldsymbol{v}) - D_t N(\boldsymbol{v})\right) - \mathcal{M}(y)\right)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{v}'\mathrm{d}y\,.$$

Then, we take the absolute value inside the integral, and thanks to assumption (2.2a), we estimate $\int \Lambda\left(y - D_t N(t, \boldsymbol{x}, \boldsymbol{v}), \boldsymbol{v}', \boldsymbol{v}\right)\mathrm{d}\boldsymbol{v}'$ by its supremum over all $(\boldsymbol{x}, \boldsymbol{v}, y) \in \mathbb{R}^d \times V \times \mathbb{R}$, which yields

$$|\mathcal{L}_{12}| \leq C\int \bar{p}(\boldsymbol{v})\int |\mathcal{M}(y + D_t M(\boldsymbol{v}) - D_t N(\boldsymbol{v})) - \mathcal{M}(y)|\,\mathrm{d}y\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\,.$$

We obtain the same estimate for $\mathcal{L}_{11}$ following the same computations and therefore, we deduce

$$\mathcal{L}_{11} - \mathcal{L}_{12} \leq C\int \bar{p}(\boldsymbol{v})\int |\mathcal{M}\left(y + D_t M(\boldsymbol{v}) - D_t N(\boldsymbol{v})\right) - \mathcal{M}(y)|\,\mathrm{d}y\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\,.$$

To estimate the integral with respect to y in the latter right-hand side, we notice, on the one hand, that since $\mathcal{M}$ is a probability measure, it holds

$$\int |\mathcal{M}\left(y + D_t M(\boldsymbol{v}) - D_t N(\boldsymbol{v})\right) - \mathcal{M}(y)|\,\mathrm{d}y \leq 2\,,$$

for all $(t, \boldsymbol{x}, \boldsymbol{v}) \in \mathbb{R}^+ \times \mathbb{R}^d \times V$. On the other hand, thanks to Csizár-Kullback, we also have

$$\int |\mathcal{M}\left(y + D_t M(\boldsymbol{v}) - D_t N(\boldsymbol{v})\right) - \mathcal{M}(y)|\mathrm{d}y \leq$$

$$\sqrt{2}\left(\int \mathcal{M}(y)\log\left(\frac{\mathcal{M}(y)}{\mathcal{M}\left(y + D_t M(\boldsymbol{v}) - D_t N(\boldsymbol{v})\right)}\right)\mathrm{d}y\right)^{\frac{1}{2}}\,.$$

Hence, after simple computations involving $\mathcal{M}$, we obtain

$$\mathcal{L}_{11} - \mathcal{L}_{12} \leq C\int \bar{p}(\boldsymbol{v})\min\left(1\,, |D_t M(\boldsymbol{v}) - D_t N(\boldsymbol{v})|\right)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\,.$$

We estimate $D_t M - D_t N$ thanks to (3.2) in Lemma 3.1 and apply (3.8) to bound moments of $\bar{p}$ with respect to $\boldsymbol{v}$. This yields

$$\mathcal{L}_{11} - \mathcal{L}_{12} \leq C\min\left(1\,, \left(\varepsilon + e^{-\frac{t}{\varepsilon}}\right)t\right)\,.$$

Gathering our estimates, we deduce

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\bar{p} - \bar{q}^\varepsilon\|_{L^1\left(\mathbb{R}^d \times V\right)} \leq C\left(\|\,q^\varepsilon - \bar{q}^\varepsilon\,\mathcal{M}\,\|_{L^1\left(\mathbb{R}^d \times V \times \mathbb{R}\right)} + \min\left(1\,, \left(\varepsilon + e^{-\frac{t}{\varepsilon}}\right)t\right)\right)\,.$$

Then, we integrate the latter relation between 0 and t and obtain

$$\|\bar{p}(t) - \bar{q}^\varepsilon(t)\|_{L^1} \leq \|\bar{p}_0 - \bar{q}_0^\varepsilon\|_{L^1} + C\left(\int_0^t \|\,q^\varepsilon(s) - \bar{q}^\varepsilon(s)\,\mathcal{M}\,\|_{L^1}\,\mathrm{d}s + \min\left(t\,, \varepsilon\,t^2 + \varepsilon^2\right)\right)\,.$$

Applying Jensen inequality and then Proposition 4.1 to bound the integral in the right-hand side, we obtain

$$\|\bar{p}(t) - \bar{q}^\varepsilon(t)\|_{L^1} \leq \|\bar{p}_0 - \bar{q}_0^\varepsilon\|_{L^1} + C\left(\sqrt{\varepsilon\,t}\,E[q_0^\varepsilon]^{\frac{1}{2}} + \sqrt{\varepsilon\,t}\,(t+1) + \min\left(t\,,\varepsilon\,t^2 + \varepsilon^2\right)\right).$$

Then, we notice that: $\min\left(t\,,\varepsilon\,t^2 + \varepsilon^2\right) \leq t\min\left(1\,,\varepsilon\,t\right) + \varepsilon^2$, and that: $\min\left(1\,,\varepsilon\,t\right) \leq \sqrt{\varepsilon}\,(t+1)/\sqrt{t}$, where to derive the second inequality, we distinguish the cases $t > 1/\varepsilon$ and $t \leq 1/\varepsilon$ separately. Hence, we deduce the expected result

$$\|\bar{p}(t) - \bar{q}^\varepsilon(t)\|_{L^1} \leq \|\bar{p}_0 - \bar{q}_0^\varepsilon\|_{L^1} + C\left(\sqrt{\varepsilon\,t}\,E[q_0^\varepsilon]^{\frac{1}{2}} + \sqrt{\varepsilon\,t}\,(t+1) + \varepsilon^2\right).$$

$\square$

5. Proof of Theorem 2.1

To prove Theorem 2.1, we invert the change of variable (1.7) in the estimates of Propositions 4.1 and 4.2 in order to translate the results from the re-scaled distribution q^ε to the initial quantity p^ε.

Let us start with the first estimate in Theorem 2.1. According to Proposition 4.1, it holds

$$\int_0^t \|\,q^\varepsilon(s) - \bar{q}^\varepsilon(s)\,\mathcal{M}\,\|_{L^1\left(\mathbb{R}^d \times V \times \mathbb{R}\right)}^2 \,\mathrm{d}s \leq \varepsilon \int_{\mathbb{R}^d \times V \times \mathbb{R}} q_0^\varepsilon \log\left(\frac{q_0^\varepsilon}{\mathcal{M}}\right)(\boldsymbol{x}, \boldsymbol{v}, y)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}y + C\,\varepsilon\left(t^2 + 1\right),$$

for all time $t \geq 0$. Inverting (1.7) in both left and right-hand sides, the estimate becomes

$$\int_0^t \|\,p^\varepsilon(s) - \bar{p}^\varepsilon\,\mathcal{M}_{\varepsilon,N}(s)\,\|_{L^1\left(\mathbb{R}^d \times V \times \mathbb{R}\right)}^2 \,\mathrm{d}s \leq \varepsilon \int p_0^\varepsilon \log\left(\frac{p_0^\varepsilon}{\mathcal{M}_{\varepsilon,M_0}}\right)(\boldsymbol{x}, \boldsymbol{v}, m)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}m + C\,\varepsilon\left(t^2 + 1\right),$$

where $\mathcal{M}_{\varepsilon,N}$ is defined by (1.13). To estimate the relative energy on the right-hand side we decompose it as follows

$$\int p_0^\varepsilon \log\left(\frac{p_0^\varepsilon}{\mathcal{M}_{\varepsilon,M_0}}\right)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}m = \int p_0^\varepsilon \log\left(p_0^\varepsilon\right)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}m$$

$$+ \frac{1}{2}\int \left|\frac{m - M_0(\boldsymbol{x})}{\varepsilon}\right|^2 p_0^\varepsilon \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\mathrm{d}m - \log\left(2\pi\right) + \left|\log\left(\varepsilon\right)\right|.$$

Both first and second terms on the right-hand side stay uniformly bounded as ε vanishes thanks to assumptions (2.3) and (2.4) respectively. Hence, we deduce

$$\int_0^t \|\,p^\varepsilon(s) - \bar{p}^\varepsilon\,\mathcal{M}_{\varepsilon,N}(s)\,\|_{L^1\left(\mathbb{R}^d \times V \times \mathbb{R}\right)}^2 \,\mathrm{d}s \leq C\,\varepsilon\left(t^2 + 1 + \left|\log\left(\varepsilon\right)\right|\right), \quad \forall t \in \mathbb{R}^+.$$

Then, the first estimate in Theorem 2.1 is obtained taking the square root on both sides of the latter inequality.

To obtain the second estimate in Theorem 2.1, we invert (1.7) in the estimate of Proposition 4.2. Since computations are very similar to the ones presented above, we do not detail them. To conclude this proof, the last estimate in Theorem 2.1, which ensures the convergence of N towards M, is obtained thanks to (3.2) in Lemma 3.1.

6. Conclusion and perspectives

Through the mean of asymptotic analysis, we investigated a possible description of the internal mechanisms allowing bacteria to feel gradients. The main point of our analysis is that we obtain strong convergence with explicit rates towards the phenomenological run-and-tumble equation, departing from a model incorporating internal variables. In order to capture the complexity of real-life applications, we proposed a robust method, which allow for general tumbling kernel. Numerous perspectives arise from this work.

A natural improvement of our result consists in removing the well-preparedness condition (2.4). The same difficulty was successfully dealt with in the context of concentration phenomena for neural networks [11, 9, 10]. Indeed, the well-preparedness condition was removed in [9] thanks to a time dependent re-scaling of the internal variable. Since (1.4) and the model studied in [9] share a similar

structure, the arguments developed in [9] should adapt in our context.

Another interesting perspective would be to extend our analysis to more general internal dynamics [24, 25]. Nonlinear dynamics should be in the range of our method but in this case, the limiting tumbling kernel $\bar{\Lambda}$ would be given as the convolution product between Λ and the stationary solution to a Fokker-Planck equation with internal dynamics dependent coefficients. This would yield interesting insights on the interplay between Λ, the internal dynamics and the resulting tumbling kernel $\bar{\Lambda}$. It could complement former detailed studies on that matter [44]. Including fast variables could lead to very interesting and biologically relevant questions as it is not clear how scale fast variables compared to ε. Biological experiments [43, 42, 32] provide insights on the multiple time scales involved in the signaling pathways and should therefore constitute the starting point to any further investigation. To conclude on internal dynamics, we could change vanishing diffusion with respect to m in (1.4) to another diffusion process or even no diffusion at all. In the last case, which has been investigated in [37], obtaining convergence rates would require structural changes in our approach.

To conclude, coupling equation (1.4) with an evolution equation for M is of primary interest on both the biological and mathematical point of view. On the mathematical point of view, it would lead to an intricate analysis as (1.4) would become nonlinear. On the biological point of view, this model would hopefully capture complex patterns observed experimentally [17]. Quantifying rigorously these patterns would lead to very interesting developments.

References

[1] J. Adler. Effect of amino acids and oxygen on chemotaxis in Escherichia coli. *J Bacteriol*, 92(1):121–129, Jul 1966.

[2] U. Alon, L. Camarena, M. G. Surette, B. Aguera y Arcas, Y. Liu, S. Leibler, and J. B. Stock. Response regulator output in bacterial chemotaxis. *EMBO J*, 17(15):4238–4248, Aug 1998.

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

[4] N. Barkai and S. Leibler. Robustness in simple biochemical networks. *Nature*, 387(6636):913–917, Jun 1997.

[5] H. C. Berg. How bacteria swim. *Scientific American*, 233(2):36–45, 1975.

[6] H. C. Berg and D. A. Brown. Chemotaxis in escherichia coli analysed by three-dimensional tracking. *Nature*, 239:500–504, 1972.

[7] H. C. Berg and P. M. Tedesco. Transient response to chemotactic stimuli in escherichia coli. *Proc Natl Acad Sci U S A*, 72(8):3235–9, 1975.

[8] M. W. Beyerinck. Ueber atmungsfiguren beweglicher bakterien. *Zentr Bakteriol. Parasite*, 14:827–845, 1893.

[9] A. Blaustein. Large coupling in a FitzHugh-Nagumo neural network: quantitative and strong convergence results. *J. Differential Equations*, 374:218–266, 2023.

[10] A. Blaustein and E. Bouin. Concentration profiles in fitzhugh-nagumo neural networks: A hopf-cole approach, 2022.

[11] A. Blaustein and F. Filbet. Concentration phenomena in FitzHugh-Nagumo equations: a mesoscopic approach. *SIAM J. Math. Anal.*, 55(1):367–404, 2023.

[12] S. M. Block, J. E. Segall, and H. C. Berg. Adaptation kinetics in bacterial chemotaxis. *J Bacteriol*, 154(1):312–323, Apr 1983.

[13] K. A. Borkovich, N. Kaplan, J. F. Hess, and M. I. Simon. Transmembrane signal transduction in bacterial chemotaxis involves ligand-dependent activation of phosphate group transfer. *Proc Natl Acad Sci U S A*, 86(4):1208–1212, Feb 1989.

[14] N. Bournaveas and V. Calvez. Global existence for the kinetic chemotaxis model without pointwise memory effects, and including internal variables. *Kinet. Relat. Models*, 1(1):29–48, 2008.

[15] N. Bournaveas and V. Calvez. Critical mass phenomenon for a chemotaxis kinetic model with spherically symmetric initial data. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 26(5):1871–1895, 2009.

[16] N. Bournaveas, V. Calvez, S. Gutiérrez, and B. Perthame. Global existence for a kinetic model of chemotaxis via dispersion and Strichartz estimates. *Comm. Partial Differential Equations*, 33(1-3):79–95, 2008.

[17] E. O. Budrene and H. C. Berg. Complex patterns formed by motile cells of escherichia coli. *Nature*, 349(6310):630–633, 1991.

[18] V. Calvez. Chemotactic waves of bacteria at the mesoscale. *J. Eur. Math. Soc. (JEMS)*, 22(2):593–668, 2020.

[19] V. Calvez, L. Gosse, and M. Twarogowska. Travelling chemotactic aggregates at mesoscopic scale and bistability. *SIAM J. Appl. Math.*, 77(6):2224–2249, 2017.

[20] F. A. C. C. Chalub, P. A. Markowich, B. Perthame, and C. Schmeiser. Kinetic models for chemotaxis and their drift-diffusion limits. *Monatsh. Math.*, 142(1-2):123–141, 2004.

[21] P. Cluzel, M. Surette, and S. Leibler. An ultrasensitive bacterial motor revealed by monitoring signaling proteins in single cells. *Science*, 287(5458):1652–1655, Mar 2000.

[22] M. Delbrück and W. Reichardt. *System analysis for the light growth reactions of phycomyces*. Princeton University Press, 1956.

[23] Y. Dolak and C. Schmeiser. Kinetic models for chemotaxis: hydrodynamic limits and spatio-temporal mechanisms. *J. Math. Biol.*, 51(6):595–615, 2005.

[24] R. Erban and H. G. Othmer. From individual to collective behavior in bacterial chemotaxis. *SIAM J. Appl. Math.*, 65(2):361–391, 2004/05.

[25] R. Erban and H. G. Othmer. From signal transduction to spatial pattern formation in *e. coli*: a paradigm for multiscale modeling in biology. *Multiscale Model. Simul.*, 3(2):362–394, 2005.

[26] R. Erban and H. G. Othmer. Taxis equations for amoeboid cells. *J. Math. Biol.*, 54(6):847–885, 2007.

[27] J. J. Falke, R. B. Bass, S. L. Butler, S. A. Chervitz, and M. A. Danielson. The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Annu Rev Cell Dev Biol*, 13:457–512, 1997.

[28] M. Fathi, E. Indrei, and M. Ledoux. Quantitative logarithmic Sobolev inequalities and stability estimates. *Discrete Contin. Dyn. Syst.*, 36(12):6835–6853, 2016.

[29] R. Hall and S. Peterson. Trajectories of human granulocytes. *Biophysical Journal*, 25(2, Part 1):365–372, 1979.

[30] H. J. Hwang, K. Kang, and A. Stevens. Global solutions of nonlinear transport equations for chemosensitive movement. *SIAM J. Math. Anal.*, 36(4):1177–1199, 2005.

[31] F. James and N. Vauchelet. Chemotaxis: from kinetic equations to aggregate dynamics. *NoDEA Nonlinear Differential Equations Appl.*, 20(1):101–127, 2013.

[32] L. Jiang, Q. Ouyang, and Y. Tu. Quantitative modeling of Escherichia coli chemotactic motion in environments varying in space and time. *PLoS Comput Biol*, 6(4):e1000735, Apr 2010.

[33] Y. V. Kalinin, L. Jiang, Y. Tu, and M. Wu. Logarithmic sensing in Escherichia coli bacterial chemotaxis. *Biophys J*, 96(6):2439–2448, Mar 2009.

[34] R. M. Macnab and D. E. Koshland. The gradient-sensing mechanism in bacterial chemotaxis. *Proc Natl Acad Sci U S A*, 69(9):2509–2512, Sep 1972.

[35] H. G. Othmer, S. R. Dunbar, and W. Alt. Models of dispersal in biological systems. *J. Math. Biol.*, 26(3):263–298, 1988.

[36] C. S. Patlak. Random walk with persistence and external bias. *Bull. Math. Biophys.*, 15:311–338, 1953.

[37] B. Perthame, M. Tang, and N. Vauchelet. Derivation of the bacterial run-and-tumble kinetic equation from a model with biochemical pathway. *J. Math. Biol.*, 73(5):1161–1178, 2016.

[38] J. Saragosti, V. Calvez, N. Bournaveas, A. Buguin, P. Silberzan, and B. Perthame. Mathematical description of bacterial traveling pulses. *PLoS Comput Biol*, 6(8), Aug 2010.

[39] B. E. Scharf, K. A. Fahrner, L. Turner, and H. C. Berg. Control of direction of flagellar rotation in bacterial chemotaxis. *Proc Natl Acad Sci U S A*, 95(1):201–206, Jan 1998.

[40] J. C. Sherris, N. W. Preston, and J. G. Shoesmith. The influence of oxygen and arginine on the motility of a strain of Pseudomonas sp. *J Gen Microbiol*, 16(1):86–96, Feb 1957.

[41] G. Si, M. Tang, and X. Yang. A pathway-based mean-field model for e. coli chemotaxis: Mathematical derivation and its hyperbolic and parabolic limits. *Multiscale Modeling & Simulation*, 12(2):907–926, 2014.

[42] V. Sourjik. Receptor clustering and signal processing in E. coli chemotaxis. *Trends Microbiol*, 12(12):569–576, Dec 2004.

[43] V. Sourjik and H. C. Berg. Receptor sensitivity in bacterial chemotaxis. *Proc Natl Acad Sci U S A*, 99(1):123–127, Jan 2002.

[44] C. Xue. Macroscopic equations for bacterial chemotaxis: integration of detailed biochemistry of cell signaling. *J. Math. Biol.*, 70(1-2):1–44, 2015.

[45] C. Xue and H. G. Othmer. Multiscale models of taxis-driven patterning in bacterial populations. *SIAM J. Appl. Math.*, 70(1):133–167, 2009.

[46] T. M. Yi, Y. Huang, M. I. Simon, and J. Doyle. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc Natl Acad Sci U S A*, 97(9):4649–4653, Apr 2000.

[47] S. H. Zigmond. Mechanisms of sensing chemical gradients by polymorphonuclear leukocytes. *Nature*, 249:450–452, 1974.