Large coupling in a FitzHugh-Nagumo neural network: quantitative and strong convergence results

Alain Blaustein

LARGE COUPLING IN A FITZHUGH-NAGUMO NEURAL NETWORK: QUANTITATIVE AND STRONG CONVERGENCE RESULTS

Alain Blaustein
Université Paul Sabatier, Institut de Mathématiques de Toulouse,
118 route de Narbonne - F-31062 TOULOUSE Cedex 9, France;
email: alain.blaustein@math.univ-toulouse.fr

Abstract. We consider a mesoscopic model for a spatially extended FitzHugh-Nagumo neural network and prove that in the regime where short-range interactions dominate, the probability density of the potential throughout the network concentrates into a Dirac distribution whose center of mass solves the classical non-local reaction-diffusion FitzHugh-Nagumo system. In order to refine our comprehension of this regime, we focus on the blow-up profile of this concentration phenomenon. Our main purpose here consists in deriving two quantitative and strong convergence estimates proving that the profile is Gaussian: the first one in a L^1 functional framework and the second in a weighted L^2 functional setting. We develop original relative entropy techniques to prove the first result whereas our second result relies on propagation of regularity.

Keywords: Diffusive limit, relative entropy, FitzHugh-Nagumo, neural network.

Mathematics Subject Classification (2010): 35B40 – 35Q92 – 82C32 – 92B20

Contents

1. Introduction

1.1. **Physical model and motivations.** Over the last century, mathematical models were built in order to describe biological neural activity, laying the groundwork for computational neuroscience. We mention the pioneer work A. Hodgkin and A. Huxley [21] who derived a precise model for the voltage dynamics of a nerve cell submitted to an external input. However a general and precise description of cerebral activity seems out of reach, due to the number of neurons, the complexity of their behavior and intricate interactions. Therefore, numerous simplified models arose from neuroscience

Date: October 12, 2022.

over the last decade allowing to recover some of the behaviors observed in regimes or situations of interest. They may usually be interpreted as the mean-field limit of stochastic microscopic models. We mention integrate-and-fire neural networks [4, 7, 5], time-elapsed neuronal models [9, 8, 10, 30] and also McKean-Vlasov models including voltage and conductance variables [6, 31]. In this article, we study a FitzHugh-Nagumo neural field represented by its distribution $\mu(t, \boldsymbol{x}, \boldsymbol{u})$ depending on time t, position $\boldsymbol{x} \in K$ with K a compact set of $\mathbb{R}^d$, and $\boldsymbol{u} = (v, w) \in \mathbb{R}^2$ where v stands for the membrane potential and w is an adaptation variable. The distribution μ is normalized by the total density $\rho_0(\boldsymbol{x})$ of neuron at position $\boldsymbol{x}$. Therefore μ is a non-negative function taken in $\mathscr{C}^0\left(\mathbb{R}^+ \times K\,,\, L^1\left(\mathbb{R}^2\right)\right)$ which verifies

$$\int_{\mathbb{R}^2} \mu(t, \boldsymbol{x}, \boldsymbol{u})\, \mathrm{d}\boldsymbol{u} \,=\, 1, \quad \forall\, (t, \boldsymbol{x}) \,\in\, \mathbb{R}^+ \times K\,,$$

and which solves the following McKean-Vlasov equation (see [11, 12, 27] for other instances of such model)

$$\partial_t\, \mu \,+\, \partial_v\left(\left(N(v) - w - \mathcal{K}_\Phi[\rho_0\, \mu]\right) \mu\right) \,+\, \partial_w\left(A\,(v, w)\, \mu\right) \,-\, \partial_v^2 \mu \,=\, 0\,,$$

where the non-linear term $\mathcal{K}_\Phi[\rho_0\, \mu]\, \mu$ is induced by non-local electrostatic interactions: we suppose that neurons interact through Ohm's law and that the conductance between two neurons is given by an interaction kernel $\Phi : K^2 \to \mathbb{R}$ which depends on their position, this yields

$$\mathcal{K}_\Phi[\rho_0\, \mu](t, \boldsymbol{x}, v) \,=\, \int_{K \times \mathbb{R}^2} \Phi(\boldsymbol{x}, \boldsymbol{x}')\, (v - v')\, \rho_0(\boldsymbol{x}') \mu(t,\, \boldsymbol{x}', \boldsymbol{u}') \mathrm{d}\boldsymbol{x}'\, \mathrm{d}\boldsymbol{u}'\,.$$

Moreover, the other terms are associated to the individual behavior of each neuron, given in our context by the model built by R. FitzHugh and J. Nagumo in [17, 28]. They consist in a drift $N \in \mathscr{C}^2\,(\mathbb{R})$, which is a confining non-linearity: setting $\omega(v) \,=\, N(v)/v$ we suppose

$$\begin{cases} \limsup\limits_{|v| \,\to\, +\infty} \omega(v) \,=\, -\infty\,, & \text{(1.1a)} \\[2ex] \sup\limits_{|v| \,\geq\, 1} \dfrac{|\omega(v)|}{|v|^{p-1}} \,<\, +\infty\,, & \text{(1.1b)} \end{cases}$$

for some $p \,\geq\, 2$. For instance, these assumptions are met by the original model proposed by R. FitzHugh and J. Nagumo, where N is a cubic non-linearity

$$N(v) \,=\, v \,-\, v^3\,.$$

On top of that A drives the dynamics of the adaptation variable, it is given by

$$A(v, w) \,=\, a\, v \,-\, b\, w \,+\, c\,,$$

where $a,\, c \in \mathbb{R}$ and $b > 0$. We also add a diffusion term with respect to v in order to take into account random fluctuations of the voltage. This type of model has been rigorously derived as the mean field limit of microscopic model in [1, 3, 12, 27, 25]. Well posedeness of the latter equation is well known and we do not discuss it here. We refer to [2, Theorem 2.3] for a precise discussion over that matter.

1.2. **Regime of strong short-range interactions.** In this article we consider a situation where Φ decomposes as follows

$$\Phi(\boldsymbol{x}, \boldsymbol{x}') \,=\, \Psi(\boldsymbol{x}, \boldsymbol{x}') \,+\, \frac{1}{\varepsilon}\, \delta_0(\boldsymbol{x} - \boldsymbol{x}')\,,$$

where the Dirac mass δ_0 accounts for short-range interactions whereas the interaction kernel Ψ models long-range interactions: it is "smoother" then δ_0 since we suppose $\Psi \,\in\, \mathscr{C}^0\left(K_{\boldsymbol{x}},\, L^1\left(K_{\boldsymbol{x}'}\right)\right)$ and assume the following bounds to hold

$$\sup_{\boldsymbol{x} \in K} \int_K \left|\Psi(\boldsymbol{x}', \boldsymbol{x})\right| \,+\, \left|\Psi(\boldsymbol{x}, \boldsymbol{x}')\right|^r \,\mathrm{d}\boldsymbol{x}' \,<\, +\infty\,, \tag{1.2}$$

for some $r > 1$ (we denote r' its conjugate: $r' = (r-1)/r$). We point out that our assumptions on Ψ are quite general, this is in line with other works which put great efforts in order to consider general interactions [22]. The scaling parameter ε represents the magnitude of short-range interactions; we focus on the regime where they dominate, that is, when $\varepsilon \ll 1$. From these assumptions, the equation on μ can be rewritten as

$$\partial_t\, \mu^\varepsilon \,+\, \partial_v\left(\left(N(v) - w - \mathcal{K}_\Psi[\rho_0^\varepsilon\, \mu^\varepsilon]\right) \mu^\varepsilon\right) \,+\, \partial_w\left(A\,(v, w)\, \mu^\varepsilon\right) \,-\, \partial_v^2 \mu^\varepsilon \,=\, \frac{\rho_0^\varepsilon}{\varepsilon}\, \partial_v\left(\left(v - \mathcal{V}^\varepsilon\right) \mu^\varepsilon\right), \tag{1.3}$$

where the averaged voltage and adaptation variables $\mathcal{U}^\varepsilon = (\mathcal{V}^\varepsilon, \mathcal{W}^\varepsilon)$ at a spatial location $\boldsymbol{x}$ are defined as

$$\left\{ \begin{array}{lll} \mathcal{V}^\varepsilon(t,\boldsymbol{x}) & = & \displaystyle\int_{\mathbb{R}^2} v\ \mu^\varepsilon(t,\boldsymbol{x},\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,, \\ \mathcal{W}^\varepsilon(t,\boldsymbol{x}) & = & \displaystyle\int_{\mathbb{R}^2} w\ \mu^\varepsilon(t,\boldsymbol{x},\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,. \end{array} \right. \tag{1.4}$$

Previous works already went through the analysis of the asymptotic $\varepsilon \ll 1$ and it was proven that in this regime the voltage distribution undergoes a concentration phenomenon. We mention [11] which investigates this subject in a deterministic setting using relative entropy methods and also [33] where the author study this model in a spatially homogeneous framework following a Hamilton-Jacobi approach. These works conclude that as ε vanishes, μ^ε converges as follows

$$\mu^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) \underset{\varepsilon\to 0}{\longrightarrow} \delta_{\mathcal{V}(t,\boldsymbol{x})}(v) \otimes \bar{\mu}(t,\boldsymbol{x},w)\,,$$

where the couple $(\mathcal{V},\bar{\mu})$ solves

$$\left\{ \begin{array}{l} \partial_t \mathcal{V} \,=\, N(\mathcal{V}) \,-\, \mathcal{W} \,-\, \mathcal{L}_{\rho_0}[\,\mathcal{V}\,]\,, \\ \partial_t\,\bar{\mu} \,+\, \partial_w\,(A(\mathcal{V}\,,\,w)\,\bar{\mu}) \,=\, 0\,, \end{array} \right. \tag{1.5}$$

with

$$\mathcal{W} \,=\, \int_{\mathbb{R}} w\,\bar{\mu}\,(t\,,\,\boldsymbol{x}\,,\,w)\,\mathrm{d}w\,,$$

and where $\mathcal{L}_{\rho_0}[\,\mathcal{V}\,]$ is a non local operator given by

$$\mathcal{L}_{\rho_0}[\,\mathcal{V}\,] = \mathcal{V}\,\Psi *_r \rho_0 \,-\, \Psi *_r (\rho_0 \mathcal{V})\,,$$

where $*_r$ is a shorthand notation for the convolution on the right side of any function g with Ψ

$$\Psi *_r g(\boldsymbol{x}) \,=\, \int_K \Psi(\boldsymbol{x},\boldsymbol{x}')\,g(\boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}'\,.$$

Then the concentration profile of μ^ε around $\delta_{\mathcal{V}}$ was investigated in [2]. The strategy consists in considering the following re-scaled version ν^ε of μ^ε

$$\mu^\varepsilon\,(t,\,\boldsymbol{x},\,\boldsymbol{u}) \,=\, \frac{1}{\theta^\varepsilon}\,\nu^\varepsilon\left(t,\,\boldsymbol{x},\,\frac{v-\mathcal{V}^\varepsilon}{\theta^\varepsilon},\,w-\mathcal{W}^\varepsilon\right),$$

where θ^ε shall be interpreted as the concentration rate of μ^ε around its mean value $\mathcal{V}^\varepsilon$. Under the scaling $\theta^\varepsilon = \sqrt{\varepsilon}$, it is proven that

$$\nu^\varepsilon \underset{\varepsilon\to 0}{\sim} \nu := \mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}\,, \tag{1.6}$$

where $\bar{\nu}$ solves the following linear transport equation

$$\partial_t\,\bar{\nu} \,-\, b\,\partial_w\,(w\,\bar{\nu}) \,=\, 0\,, \tag{1.7}$$

and where the Maxwellian $\mathcal{M}_{\rho_0^\varepsilon}$ is defined as

$$\mathcal{M}_{\rho_0^\varepsilon(\boldsymbol{x})}(v) = \sqrt{\frac{\rho_0^\varepsilon(\boldsymbol{x})}{2\pi}}\,\exp\left(-\,\rho_0^\varepsilon(\boldsymbol{x})\,\frac{|v|^2}{2}\right).$$

The latter convergence translates on μ^ε as follows

$$\mu^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) \underset{\varepsilon\to 0}{\sim} \mathcal{M}_{\rho_0\,|\theta^\varepsilon|^{-2}}\,(v \,-\, \mathcal{V}) \otimes \bar{\mu}(t,\boldsymbol{x},\boldsymbol{u})\,, \tag{1.8}$$

with $\theta^\varepsilon = \sqrt{\varepsilon}$. More precisely, it was proven in [2] that (1.8) occurs up to an error of order ε in the sense of weak convergence in some probability space. Our goal here is to strengthen the results obtained in [2] by providing strong convergence estimates for (1.8). The general strategy is to prove (1.6) on ν^ε in strong topology and then to deduce (1.8) on μ^ε. The main difficulties to achieve this is twofold. On the one hand, since the norms associated to strong topology are usually not scaling invariant, the time homogeneous scaling $\theta^\varepsilon = \sqrt{\varepsilon}$ comes down to considering well-prepared initial conditions. Therefore we find an appropriate scaling θ^ε which enables to treat general initial condition. On the other hand, the proof is made challenging by the cross terms between v and w in (1.3). This issue is analogous to the difficulty induced by the free transport operator in the context of kinetic theory [14, 19, 20]. In our context, we propagate regularity in order to overcome this difficulty and

obtain error estimates.

This article is organized as follows. We start with Section 2, in which we carry out an heuristic in order to derive the appropriate scaling θ^ε and then state our two main results. We first provide convergence estimates for μ^ε in a L^1 setting, which is the natural space to consider for such type of conservative problem (see Theorem 2.2). This result is a direct consequence of the convergence of the re-scaled distribution ν^ε (see Theorem 2.1) which is obtained in Section 3. Then we propose convergence estimates in a weighted L^2 setting (see Theorem 2.4). Once again this result is the consequence of the convergence of ν^ε (see Theorem 2.3) provided in Section 4. We emphasize that the latter result allows us to recover the optimal convergence rates obtained in [2] and to achieve pointwise convergence estimates with respect to time. This analysis is in line with [27], which focuses on the regime of weak interactions between neurons (this corresponds to the asymptotic $\varepsilon \to +\infty$ in equation (1.3)).

2. Heuristic and main results

As mentioned before the time homogeneous scaling $\theta^\varepsilon = \sqrt{\varepsilon}$ comes down to considering well-prepared initial conditions. We seek for a stronger result which also applies for ill-prepared initial conditions. To overcome this difficulty, our strategy consists in adding the following constraint on the concentration rate

$$\theta^\varepsilon\,(t\,=\,0)\,=\,1\,.$$

In this setting, the equation on ν^ε is obtained performing the following change of variable

$$(t\,,\,v\,,\,w)\mapsto\left(t\,,\,\frac{v-\mathcal{V}^\varepsilon}{\theta^\epsilon}\,,\,w-\mathcal{W}^\varepsilon\right) \tag{2.1}$$

in equation (1.3). Following computations detailed in [2], it yields

$$\partial_t\,\nu^\varepsilon\,+\,\mathrm{div}_{\boldsymbol{u}}\,[\,\mathbf{b}_0^\varepsilon\,\nu^\varepsilon]\,=\,\frac{1}{(\theta^\varepsilon)^2}\,\partial_v\left[\left(\frac{1}{2}\,\frac{\mathrm{d}}{\mathrm{d}t}\,|\theta^\varepsilon|^2+\frac{\rho_0^\varepsilon}{\varepsilon}\,|\theta^\varepsilon|^2\right)v\,\nu^\varepsilon+\partial_v\nu^\varepsilon\right]\,, \tag{2.2}$$

where $\mathbf{b}_0^\varepsilon$ is given by

$$\mathbf{b}_0^\varepsilon\,(t,\boldsymbol{x},\boldsymbol{u})\,=\,\begin{pmatrix}(\theta^\varepsilon)^{-1}\,B_0^\varepsilon\,(t,\boldsymbol{x},\theta^\varepsilon\,v,w)\\ \\ A_0\,(\theta^\varepsilon v,w)\end{pmatrix} \tag{2.3}$$

and B_0^ε is defined as

$$B_0^\varepsilon(t\,,\,\boldsymbol{x}\,,\,\boldsymbol{u})\,=\,N(\mathcal{V}^\varepsilon\,+\,v)\,-\,N(\mathcal{V}^\varepsilon)\,-\,w\,-\,v\,\Psi\ast_r\,\rho_0^\varepsilon(\boldsymbol{x})\,-\,\mathcal{E}\,(\mu^\varepsilon)\,,$$

with $\mathcal{E}\,(\mu^\varepsilon)$ the following error term

$$\mathcal{E}\,(\mu^\varepsilon\,(t,\boldsymbol{x},\,\cdot\,))\,=\,\int_{\mathbb{R}^2}N(v)\,\mu^\varepsilon\,(t,\boldsymbol{x},\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,-\,N\,(\mathcal{V}^\varepsilon\,(t,\boldsymbol{x}))\,, \tag{2.4}$$

and where A_0 is the linear version of A

$$A_0(\boldsymbol{u})=A(\boldsymbol{u})-A(\mathbf{0})\,.$$

Since we expect concentration with Gaussian profile $\mathcal{M}_{\rho_0^\varepsilon}$ as ε vanishes, and keeping the leading order in (2.2), we find that θ^ε should verify

$$\begin{cases}\dfrac{1}{2}\,\dfrac{\mathrm{d}}{\mathrm{d}t}\,|\theta^\varepsilon|^2\,+\,\dfrac{\rho_0^\varepsilon}{\varepsilon}\,|\theta^\varepsilon|^2\,=\,\rho_0^\varepsilon\,,\\ \theta^\varepsilon\,(t\,=\,0)\,=\,1\,,\end{cases}$$

whose solution is given by the following explicit formula

$$\theta^\varepsilon(t,\,\boldsymbol{x})^2\,=\,\varepsilon\,(1\,-\,\exp\,(-(\,2\,\rho_0^\varepsilon(\boldsymbol{x})\,t\,)\,/\,\varepsilon))\,+\,\exp\,(-(\,2\,\rho_0^\varepsilon(\boldsymbol{x})\,t\,)\,/\,\varepsilon)\,. \tag{2.5}$$

Therefore, we obtain a time dependent θ^ε, which is of order $\sqrt{\varepsilon}$, up to an exponentially decaying correction to authorize ill-prepared initial conditions. With this choice, the equation on ν^ε rewrites:

$$\partial_t\,\nu^\varepsilon\,+\,\mathrm{div}_{\boldsymbol{u}}\,[\,\mathbf{b}_0^\varepsilon\,\nu^\varepsilon\,]\,=\,\frac{1}{|\theta^\varepsilon|^2}\,\mathcal{F}_{\rho_0^\varepsilon}\,[\,\nu^\varepsilon\,]\,, \tag{2.6}$$

where $\mathbf{b}_0^\varepsilon$ is given by (2.3) and the Fokker-Planck operator is defined as

$$\mathcal{F}_{\rho_0^\varepsilon}[\,\nu^\varepsilon\,] \,=\, \partial_v\,[\,\rho_0^\varepsilon\, v\,\nu^\varepsilon + \partial_v\,\nu^\varepsilon\,]\,.$$

Let us now precise our assumptions on the initial data. We suppose the following uniform boundedness condition on the spatial distribution ρ_0^ε

$$\rho_0^\varepsilon \in \mathscr{C}^0\,(\,K\,) \quad \text{and} \quad m_* \leq \rho_0^\varepsilon \leq 1/m_*\,, \tag{2.7}$$

as well as moment assumptions on the initial data

$$\begin{cases} \displaystyle\sup_{\boldsymbol{x}\in K}\int_{\mathbb{R}^2} |\boldsymbol{u}|^{2p}\mu_0^\varepsilon\,(\boldsymbol{x},\boldsymbol{u})\;\mathrm{d}\boldsymbol{u} \,\leq\, m_p\,, & \text{(2.8a)}\\[2ex] \displaystyle\int_{K\times\mathbb{R}^2} |\boldsymbol{u}|^{2pr'}\,\rho_0^\varepsilon(\boldsymbol{x})\,\mu_0^\varepsilon\,(\boldsymbol{x},\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{x} \,\leq\, \overline{m}_p\,, & \text{(2.8b)} \end{cases}$$

where p and r' are given in (1.1b) and (1.2), for constants $m_*, m_p, \overline{m}_p$ uniform with respect to ε. All along our analysis, we denote by τ_{w_0} the translation by w_0 with respect to the w-variable, for any given w_0 in $\mathbb{R}$

$$\tau_{w_0}\,\nu\,(t,\boldsymbol{x},v,w) \,=\, \nu\,(t,\boldsymbol{x},v,w+w_0)\;.$$

Since well posedeness of the mean field equation (1.3) and the limiting model (1.5) is well known, we do not discuss it here and refer to [2, Theorems 2.3 and 2.6] for a precise discussion on that matter. To apply these results we suppose the following assumptions which are not uniform with respect to ε on μ_0^ε

$$\begin{cases} \displaystyle\sup_{\boldsymbol{x}\in K}\int_{\mathbb{R}^2} e^{|\boldsymbol{u}|^2/2}\,\mu_0^\varepsilon(\boldsymbol{x},\boldsymbol{u})\mathrm{d}\boldsymbol{u} \,<\, +\infty\,,\\[2ex] \displaystyle\sup_{\boldsymbol{x}\in K}\Big\|\nabla_{\boldsymbol{u}}\sqrt{\mu_0^\varepsilon}(\boldsymbol{x},\cdot)\Big\|_{L^2(\mathbb{R}^2)} \,<\, +\infty\,, \end{cases} \tag{2.9}$$

and for the limiting problem (1.5), we suppose

$$(\mathcal{V}_0\,,\,\bar{\mu}_0) \in \mathscr{C}^0\,(K)\times\mathscr{C}^0\,\big(K\,,\,L^1\,(\mathbb{R})\big)\;. \tag{2.10}$$

2.1. L^1 **convergence result.** In the following result, we provide explicit convergence rates for ν^ε towards the asymptotic concentration profile of the neural network's distribution μ^ε in the regime of strong interactions in the following L^1 setting

$$L^\infty_{\boldsymbol{x}}L^1_{\boldsymbol{u}} \,:=\, L^\infty\,\big(K\,,\,L^1\,(\mathbb{R}^2)\big)\;.$$

We prove that the profile of concentration with respect to v is Gaussian and we also characterize the limiting distribution with respect to the adaptation variable w. We denote by H the Boltzmann entropy, defined for all function $\mu:\mathbb{R}^2\to\mathbb{R}^+$ as follows

$$H\,[\,\mu\,] \,=\, \int_{\mathbb{R}^2}\mu\,\ln\,(\mu)\,\mathrm{d}\boldsymbol{u}\,.$$

Theorem 2.1. *Under assumptions (1.1a)-(1.1b) on the drift N, assumption (1.2) on the interaction kernel Ψ, consider the unique sequence of solutions $(\mu^\varepsilon)_{\varepsilon>0}$ to (1.3) with initial conditions satisfying assumptions (2.7)-(2.9) and the solution $\bar{\nu}$ to equation (1.7) with an initial condition $\bar{\nu}_0$ such that*

$$\bar{\nu}_0 \in L^\infty\,\big(K\,,\,W^{2,1}\,(\mathbb{R})\big)\;,\quad \textit{and}\quad \sup_{\boldsymbol{x}\in K}\int_{\mathbb{R}}|w\,\partial_w\,\bar{\nu}_0(\boldsymbol{x},w)|\,\mathrm{d}w<+\infty\,. \tag{2.11}$$

On top of that, suppose that there exists a positive constant m_1 such that it holds

$$\sup_{\varepsilon>0}\,\|\,\nu_0^\varepsilon \,-\, \tau_{w_0}\,\nu_0^\varepsilon\,\|_{L^\infty_{\boldsymbol{x}}L^1_{\boldsymbol{u}}} \,\leq\, m_1\,|w_0|\,,\qquad \forall\, w_0\in\mathbb{R}\,, \tag{2.12}$$

and suppose that there exists a positive constant m_2 such that

$$\sup_{\varepsilon>0}\,\|\,H\,[\,\nu_0^\varepsilon\,]\,\|_{L^\infty(K)} \,\leq\, m_2^2\,. \tag{2.13}$$

Then, there exists a positive constant C independent of ε such that for all ε less than 1, it holds

$$\int_0^t \Big\|\,\nu^\varepsilon \,-\, \mathcal{M}_{\rho_0^\varepsilon}\otimes\bar{\nu}\,\Big\|_{L^\infty_{\boldsymbol{x}}L^1_{\boldsymbol{u}}}(s)\,\mathrm{d}s \,\leq\, 2\sqrt{2}\,t\,\|\,\bar{\nu}_0^\varepsilon \,-\, \bar{\nu}_0\|^{1/2}_{L^\infty_{\boldsymbol{x}}L^1_w} \,+\, \sqrt{\varepsilon}\,\Big(4\,\sqrt{t}\,m_2 \,+\, C\,e^{2\,b\,t}\Big)\;,$$

for all time $t \geq 0$. In particular, under the compatibility assumption

$$\| \bar{\nu}_0^\varepsilon - \bar{\nu}_0 \|_{L^\infty_{\boldsymbol{x}} L^1_w}^{1/2} \underset{\varepsilon \to 0}{=} O\left(\sqrt{\varepsilon}\right) ,$$

it holds

$$\sup_{t \in \mathbb{R}^+} \left[e^{-2\,b\,t} \int_0^t \left\| \nu^\varepsilon(s) - \mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}(s) \right\|_{L^\infty_{\boldsymbol{x}} L^1_{\boldsymbol{u}}} \mathrm{d}s \right] \underset{\varepsilon \to 0}{=} O\left(\sqrt{\varepsilon}\right) .$$

In this result, the constant C only depends on m_1, m_*, m_p and $\overline{m}_p$ (see assumptions (2.7)-(2.8b)) and the data of the problem $\bar{\nu}_0$, N, Ψ and A_0.

The proof of this result is divided into two steps. First we prove that ν^ε converges towards the following local equilibrium of the Fokker-Planck operator

$$\mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}^\varepsilon ,$$

where $\bar{\nu}^\varepsilon$ is the marginal of ν^ε with respect to the re-scaled adaptation variable

$$\bar{\nu}^\varepsilon(t\,,\,\boldsymbol{x}\,,\,w) = \int_{\mathbb{R}} \nu^\varepsilon(t\,,\,\boldsymbol{x}\,,\,\boldsymbol{u})\,\mathrm{d}v ,$$

and solves the following equation, obtained after integrating equation (2.6) with respect to v

$$\partial_t \bar{\nu}^\varepsilon - b\,\partial_w\left(w\,\bar{\nu}^\varepsilon\right) = -a\,\theta^\varepsilon\,\partial_w \int_{\mathbb{R}} v\,\nu^\varepsilon(t\,,\,\boldsymbol{x}\,,\,\boldsymbol{u})\,\mathrm{d}v . \tag{2.14}$$

The argument relies on a rather classical relative entropy. However, the analysis becomes more intricate when it comes to the convergence of the marginal $\bar{\nu}^\varepsilon$. As already mentioned, the proof of convergence is made challenging by cross terms between v and w in equation (2.6) inducing in equation (2.14) the following term which involves derivatives of ν^ε

$$\partial_w \int_{\mathbb{R}} v\,\nu^\varepsilon(t\,,\,\boldsymbol{x}\,,\,\boldsymbol{u})\,\mathrm{d}v .$$

To overcome this difficulty, we perform a change of variable which cancels the latter source term and then conclude by proving a uniform equicontinuity estimate.

We now interpret Theorem 2.1 on the solution μ^ε to equation (1.3) in the regime of strong interactions

Theorem 2.2. *Under the assumptions of Theorem 2.1 consider the unique sequence of solutions $(\mu^\varepsilon)_{\varepsilon>0}$ to (1.3) as well as the unique solution $(\mathcal{V}\,,\,\bar{\mu})$ to equation (1.5) with an initial condition $\bar{\mu}_0$ which fulfills assumption (2.10)-(2.11). On top of that, we suppose the following compatibility assumption to be fulfilled*

$$\|\mathcal{U}_0 - \mathcal{U}_0^\varepsilon\|_{L^\infty(K)} + \|\rho_0 - \rho_0^\varepsilon\|_{L^\infty(K)} + \|\bar{\mu}_0^\varepsilon - \bar{\mu}_0\|_{L^\infty_{\boldsymbol{x}} L^1_w}^{1/2} \underset{\varepsilon \to 0}{=} O\left(\sqrt{\varepsilon}\right) . \tag{2.15}$$

There exists $(C\,,\,\varepsilon_0) \in \left(\mathbb{R}_+^*\right)^2$ such that for all ε less than ε_0, it holds

$$\int_0^t \|\mu^\varepsilon - \mu\|_{L^\infty_{\boldsymbol{x}} L^1_{\boldsymbol{u}}}(s)\,\mathrm{d}s \leq C\,e^{C\,t}\,\sqrt{\varepsilon} , \quad \forall\, t \in \mathbb{R}^+ ,$$

where the limit μ is given by

$$\mu := \mathcal{M}_{\rho_0\,|\theta^\varepsilon|^{-2}}\,(v - \mathcal{V}) \otimes \bar{\mu} .$$

In this result, the constant C and ε_0 only depend on the implicit constant in assumption (2.15), on the constants m_1, m_2 m_*, m_p and $\overline{m}_p$ (see assumptions (2.12)-(2.13) and (2.7)-(2.8b)) and on the data of the problem $\bar{\mu}_0$, N, Ψ and A_0.

Proof. Since the norm $\|\cdot\|_{L^1(\mathbb{R}^2)}$ is unchanged by the re-scaling (2.1), this theorem is a straightforward consequence of Theorem 2.1 and Proposition 2.5, which ensures the convergence of the macroscopic quantities $(\mathcal{V}^\varepsilon, \mathcal{W}^\varepsilon)$. □

On the one hand we obtain L^1 in time convergence result which is a consequence of the relative entropy estimate between ν^ε and $\mathcal{M}_{\rho_0^\varepsilon}$. This is somehow similar to what is obtained in various classical kinetic models. Let us mention for instance the diffusive limit for the Vlasov-Poisson-Fokker-Planck [14, 19, 26]. On the other hand, we obtain the convergence rate $O(\sqrt{\varepsilon})$ instead of the optimal convergence rate, which should be $O(\varepsilon)$ as rigorously proven for weak convergence metrics (see [2], Theorem 2.7). This is due to the fact that we use the Csizár-Kullback inequality to close our L^1 convergence estimates. Therefore it seems quite unlikely to recover the optimal convergence rate in a L^1 setting. This motivates our next work, on the L^2 convergence (Theorems 2.3 and 2.4), in which pointwise in time convergence is achieved.

2.2. **Weighted L^2 convergence result.** In this section, we provide a pair of result analog to Theorems 2.1 and 2.2 this time in a weighted L^2 setting. Since our approach relies on propagating w-derivatives, we introduce the following functional framework

$$\mathscr{H}^k\left(m^\varepsilon\right) \,=\, L^\infty\left(K_{\boldsymbol{x}}, H^k_w\left(m^\varepsilon_{\boldsymbol{x}}\right)\right)\,,$$

equipped with the norm

$$\|\,\nu\,\|_{\mathscr{H}^k(m^\varepsilon)} \,=\, \sup_{\boldsymbol{x}\in K}\left\{\|\,\nu\left(\boldsymbol{x}\,,\,\cdot\,\right)\|_{H^k_w(m^\varepsilon_{\boldsymbol{x}})}\right\}\,,$$

where $H^k_w(m^\varepsilon_{\boldsymbol{x}})$ is the weighted Sobolev space whose norm is given by

$$\|\,\nu\,\|^2_{H^k_w(m^\varepsilon_{\boldsymbol{x}})} \,=\, \sum_{l\,\leq\,k}\int_{\mathbb{R}^2}\left|\partial^l_w\,\nu(\boldsymbol{u})\right|^2\,m^\varepsilon_{\boldsymbol{x}}(\boldsymbol{u})\mathrm{d}\boldsymbol{u}\,,$$

and where the weight $m^\varepsilon_{\boldsymbol{x}}$ is given by

$$m^\varepsilon_{\boldsymbol{x}}(\boldsymbol{u}) \,=\, \frac{2\,\pi}{\sqrt{\rho^\varepsilon_0\,\kappa}}\,\exp\left(\frac{1}{2}\left(\,\rho^\varepsilon_0(\boldsymbol{x})\,|v|^2\,+\,\kappa\,|w|^2\,\right)\right)\,, \tag{2.16}$$

for some exponent $\kappa>0$ which will be prescribed later. We also introduce the associated weight with respect to the adaptation variable

$$\overline{m}(w) \,=\, \sqrt{\frac{2\,\pi}{\kappa}}\,\exp\left(\frac{\kappa}{2}\,|w|^2\right)\,.$$

and denote by $\mathscr{H}^k\left(\overline{m}\right)$ the corresponding functional space associated to the marginal $\bar{\nu}$, depending only on $(\boldsymbol{x},w)\in K\times\mathbb{R}$.

Hence, the following result tackles the convergence of ν^ε in the L^2 weighted setting

Theorem 2.3. *Under assumptions (1.1a)-(1.1b) on the drift N and the additional assumption*

$$\sup_{|v|\,\geq\,1}\left(v^2\,\omega(v)\,-\,C_0\,N'(v)\right)\,<\,+\infty\,, \tag{2.17}$$

for all positive constant $C_0>0$, supposing assumption (1.2) on the interaction kernel Ψ, consider the unique sequence of solutions $(\mu^\varepsilon)_{\varepsilon\,>\,0}$ to (1.3) with initial conditions satisfying assumptions (2.7)-(2.9) and the solution $\bar{\nu}$ to equation (1.7) with an initial condition $\bar{\nu}_0$. On top of that, consider an exponent κ which verifies the condition

$$\kappa\,\in\,\left(\frac{1}{2\,b}\,,\,+\infty\right)\,, \tag{2.18}$$

and consider a rate α_* lying in $\left(0\,,\,1-(2b\kappa)^{-1}\right)$. There exists a positive constant C independent of ε such that for all ε between 0 and 1 the following results hold true

(1) consider k in $\{0\,,\,1\}$ and suppose that the sequence $(\nu^\varepsilon_0)_{\varepsilon\,>\,0}$ verifies

$$\sup_{\varepsilon\,>\,0}\|\,\nu^\varepsilon_0\,\|_{\mathscr{H}^{k+1}(m^\varepsilon)}\,<\,+\infty\,, \tag{2.19}$$

and that $\bar{\nu}_0$ verifies

$$\bar{\nu}_0\,\in\,\mathscr{H}^{\,k}(\overline{m})\,. \tag{2.20}$$

Then for all time t in $\mathbb{R}^+$ it holds

$$\| \nu^\varepsilon(t) - \nu(t) \|_{\mathscr{H}^k(m^\varepsilon)} \leq e^{Ct} \left(\| \bar{\nu}_0^\varepsilon - \bar{\nu}_0 \|_{\mathscr{H}^k(\overline{m})} + C \| \nu_0^\varepsilon \|_{\mathscr{H}^{k+1}(m^\varepsilon)} \left(\sqrt{\varepsilon} + \min\left\{ 1 , e^{-\alpha_* \frac{t}{\varepsilon}} \varepsilon^{-\frac{\alpha_*}{2m_*}} \right\} \right) \right) ,$$

where the asymptotic profile ν is given by (1.6);

(2) *suppose assumption (2.19) with index $k = 1$ and assumption (2.20) with index $k = 0$, it holds for all time $t \geq 0$*

$$\| \bar{\nu}^\varepsilon(t) - \bar{\nu}(t) \|_{\mathscr{H}^0(\overline{m})} \leq e^{Ct} \left(\| \bar{\nu}_0^\varepsilon - \bar{\nu}_0 \|_{\mathscr{H}^0(\overline{m})} + C \| \nu_0^\varepsilon \|_{\mathscr{H}^2(m^\varepsilon)} \varepsilon \sqrt{|\ln \varepsilon| + 1} \right) .$$

In this theorem, the positive constant C only depends on κ, α_*, m_*, m_p, $\overline{m}_p$ (see assumptions (2.7), (2.8a) and (2.8b)) and on the data of the problem: N, A_0 and Ψ.

The proof of this result is provided in Section 4 and relies on regularity estimates for the solution ν^ε to equation (2.6). These regularity estimates allow us to bound the source term which appears in the right hand side of equation (2.14).

We now interpret the latter theorem in terms μ^ε. Let us emphasize that since m^ε defined by (2.16) depends on ε through the spatial distribution ρ_0^ε, we introduce weights which do not depend on ε anymore and which are meant to upper and lower bound m^ε. We consider $(\boldsymbol{x}, \boldsymbol{u})$ lying in $K \times \mathbb{R}^2$ and define

$$\begin{cases} m^-_{\boldsymbol{x}}(\boldsymbol{u}) = (\rho_0(\boldsymbol{x})\,\kappa)^{-\frac{1}{2}} \exp\left(\frac{1}{8} \left(\rho_0(\boldsymbol{x}) |v|^2 + \kappa |w|^2 \right) \right) , \\ \overline{m}^-(w) = \kappa^{-\frac{1}{2}} \exp\left(\frac{\kappa}{8} |w|^2 \right) , \\ m^+_{\boldsymbol{x}}(\boldsymbol{u}) = (\rho_0(\boldsymbol{x})\,\kappa)^{-\frac{1}{2}} \exp\left(2 \left(\rho_0(\boldsymbol{x}) |v|^2 + \kappa |w|^2 \right) \right) , \\ \overline{m}^+(w) = \kappa^{-\frac{1}{2}} \exp\left(2\,\kappa |w|^2 \right) . \end{cases}$$

With these notations, our result reads as follows

Theorem 2.4. *Under the assumptions of Theorem 2.3 consider the unique sequence of solutions $(\mu^\varepsilon)_{\varepsilon>0}$ to (1.3) and the solution $(\mathcal{V}, \bar{\mu})$ to (1.5) with initial condition $(\mathcal{V}, \bar{\mu})$ satisfying (2.10). The following results hold true*

(1) *Consider k in $\{0, 1\}$ and suppose*

$$\sup_{\varepsilon>0} \| \mu_0^\varepsilon \|_{\mathscr{H}^{k+1}(m^+)} < +\infty , \tag{2.21}$$

as well as the following compatibility assumption

$$\| \mathcal{U}_0 - \mathcal{U}_0^\varepsilon \|_{L^\infty(K)} + \| \rho_0 - \rho_0^\varepsilon \|_{L^\infty(K)} + \| \bar{\mu}_0^\varepsilon - \bar{\mu}_0 \|_{\mathscr{H}^k(\overline{m}^+)} \underset{\varepsilon\to 0}{=} O(\varepsilon) . \tag{2.22}$$

On top of that, suppose that there exists a constant C such that

$$\sup_{\varepsilon>0} \| \bar{\mu}_0 - \tau_{w_0} \bar{\mu}_0 \|_{\mathcal{H}^k(\overline{m}^+)} \leq C |w_0| , \quad \forall w_0 \in \mathbb{R} . \tag{2.23}$$

Then for all integer i and under the constraint $\alpha_* < \min\{ m_*/2 , 1 - (2b\kappa)^{-1} \}$, there exists $(C_i , \varepsilon_0) \in (\mathbb{R}_+^*)^2$ such that for all ε less than ε_0, it holds for all $t \in \mathbb{R}_+$,

$$\| (v - \mathcal{V})^i (\mu^\varepsilon - \mu)(t) \|_{\mathcal{H}^k(m^-)} \leq C_i e^{C_i (t + \varepsilon e^{C_i t})} \left(\varepsilon^{\frac{i}{2} + \frac{1}{4}} + e^{-\alpha_* \frac{t}{\varepsilon}} \varepsilon^{-\frac{1}{2}} \right) ,$$

where the limit μ is given by

$$\mu = \mathcal{M}_{\rho_0 |\theta^\varepsilon|^{-2}} (v - \mathcal{V}) \otimes \bar{\mu} .$$

(2) *Suppose assumption (2.21) with $k = 1$, assumption (2.23) with $k = 0$ and*

$$\| \mathcal{U}_0 - \mathcal{U}_0^\varepsilon \|_{L^\infty(K)} + \| \rho_0 - \rho_0^\varepsilon \|_{L^\infty(K)} + \| \bar{\mu}_0^\varepsilon - \bar{\mu}_0 \|_{\mathscr{H}^0(\overline{m}^+)} \underset{\varepsilon\to 0}{=} O\left(\varepsilon \sqrt{|\ln \varepsilon|} \right) .$$

There exists $(C , \varepsilon_0) \in (\mathbb{R}_+^*)^2$ such that for all ε less than ε_0, it holds

$$\| \bar{\mu}^\varepsilon(t) - \bar{\mu}(t) \|_{\mathscr{H}^0(\overline{m}^-)} \leq C e^{Ct} \varepsilon \sqrt{|\ln \varepsilon|} , \quad \forall t \in \mathbb{R}_+ .$$

This result is a straightforward consequence of Theorem 2.3 and the convergence estimates for the macroscopic quantities given by item (1) Proposition 2.5. We postpone the proof to Section 4.3 and make a few comments. On the one hand, we achieve pointwise in time convergence estimates, which is an improvement in comparison to our result in the L^1 setting. This is made possible thanks to the regularity results obtained for ν^ε, which we were not able to obtain in the L^1 setting. On the other hand, we recover the optimal convergence rate for the marginal $\bar{\mu}^\varepsilon$ of μ^ε towards the limit $\bar{\mu}$, up to a logarithmic correction. The logarithmic correction arises due to the fact that we do not consider well prepared initial data (see Proposition 4.10 for more details). In the statement (1), we prove convergence with rate $O(\varepsilon^i)$ for all i. This is specific to the structure of the weighted L^2 at play in this result.

2.3. **Useful estimates.** Before going to the proof of our main results, we remind here uniform estimates with respect to ε, already established in [2], for the moments of μ^ε and for the relative energy given by

$$\begin{cases} M_q[\mu^\varepsilon](t,\boldsymbol{x}) := \displaystyle\int_{\mathbb{R}^2} |\boldsymbol{u}|^q \mu^\varepsilon(t,\boldsymbol{x},\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,, \\ D_q[\mu^\varepsilon](t,\boldsymbol{x}) := \displaystyle\int_{\mathbb{R}^2} |v-\mathcal{V}^\varepsilon(t,\boldsymbol{x})|^q \mu^\varepsilon(t,\boldsymbol{x},\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,, \end{cases}$$

where $q \geq 2$.

Proposition 2.5. *Under assumptions (1.1a)-(1.1b) on the drift N, (1.2) on Ψ, (2.7)-(2.8b) on the initial conditions μ_0^ε consider the unique solutions μ^ε and $(\mathcal{V},\bar{\mu})$ to (1.3) and (1.5). Furthermore, define the initial macroscopic error as*

$$\mathcal{E}_{\mathrm{mac}} = \|\mathcal{U}_0 - \mathcal{U}_0^\varepsilon\|_{L^\infty(K)} + \|\rho_0 - \rho_0^\varepsilon\|_{L^\infty(K)}\,.$$

There exists $(C,\varepsilon_0) \in (\mathbb{R}_*^+)^2$ such that

(1) for all $\varepsilon \leq \varepsilon_0$, it holds

$$\|\mathcal{U}(t) - \mathcal{U}^\varepsilon(t)\|_{L^\infty(K)} \leq C \min\left(e^{Ct}(\mathcal{E}_{\mathrm{mac}}+\varepsilon),1\right), \qquad \forall t \in \mathbb{R}^+,$$

where $\mathcal{U}^\varepsilon$ and $\mathcal{U}$ are respectively given by (1.4) and (1.5).

(2) For all $\varepsilon > 0$ and all q in $[2,2p]$ it holds

$$M_q[\mu^\varepsilon](t,\boldsymbol{x}) \leq C, \qquad \forall (t,\boldsymbol{x}) \in \mathbb{R}^+ \times K,$$

where exponent p is given in assumption (1.1b). In particular, $\mathcal{U}^\varepsilon$ is uniformly bounded with respect to both $(t,\boldsymbol{x}) \in \mathbb{R}^+ \times K$ and ε.

(3) For all $\varepsilon > 0$ and all q in $[2,2p]$ it holds

$$D_q[\mu^\varepsilon](t,\boldsymbol{x}) \leq C\left[\exp\left(-q\,m_*\frac{t}{\varepsilon}\right) + \varepsilon^{\frac{q}{2}}\right], \quad \forall (t,\boldsymbol{x}) \in \mathbb{R}^+ \times K\,.$$

(4) For all $\varepsilon > 0$ we have

$$|\mathcal{E}(\mu^\varepsilon(t,\boldsymbol{x},\cdot))| \leq C\left[\exp\left(-2\,m_*\frac{t}{\varepsilon}\right) + \varepsilon\right], \quad \forall (t,\boldsymbol{x}) \in \mathbb{R}^+ \times K\,,$$

where $\mathcal{E}$ is defined by (2.4).

The proof of this result can be found in [2]. More precisely, we refer to [2, Proposition 4.4] for the proof of (1), [2, Proposition 3.1] for the proof of (2) , [2, Proposition 3.3] for the proof of (3) and [2, Proposition 3.5] for the proof of (4).

3. Convergence analysis in L^1

In this section, we prove Theorem 2.1 which ensures the convergence of ν^ε towards the asymptotic profile $\mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}$ in a L^1 setting. In order to explain our main argument, we outline the main steps of our approach on a simplified example : the diffusive limit for the kinetic Fokker-Planck equation. We consider the asymptotic limit $\varepsilon \to 0$ of the following linear kinetic Fokker-Planck equation

$$\partial_t f^\varepsilon + \frac{1}{\varepsilon}\boldsymbol{v}\cdot\nabla_{\boldsymbol{x}} f^\varepsilon = \frac{1}{\varepsilon^2}\nabla_{\boldsymbol{v}}\cdot(\boldsymbol{v} f^\varepsilon + \nabla_{\boldsymbol{v}} f^\varepsilon)\,,$$

where $(\boldsymbol{x}, \boldsymbol{v})$ lie in the phase space $\mathbb{R}^d \times \mathbb{R}^d$. In this context, the challenge consists in proving that as ε vanishes, it holds

$$f^\varepsilon\,(t\,,\,\boldsymbol{x}\,,\,\boldsymbol{v}) \underset{\varepsilon\to 0}{\sim} \mathcal{M}(\boldsymbol{v}) \otimes \rho(t,\boldsymbol{x})\,,$$

where ρ is a solution to the heat equation

$$\partial_t\,\rho \,=\, \Delta_{\boldsymbol{x}}\,\rho\,,$$

and where $\mathcal{M}$ stands for the standard Maxwellian. Relying on a rather classical relative entropy estimate, it is possible to prove that f^ε converges to the following local equilibrium of the Fokker-Planck operator

$$\mathcal{M} \otimes \rho^\varepsilon\,,$$

where the spatial density of particles ρ^ε is defined by

$$\rho^\varepsilon \,=\, \int_{\mathbb{R}^d} f^\varepsilon \,\mathrm{d}\boldsymbol{v}\,.$$

Then, the difficulty lies in proving that the spatial density of particles ρ^ε converges to ρ. The convergence analysis is made intricate by the transport operator, which keeps us from obtaining a closed equation on ρ^ε

$$\partial_t\,\rho^\varepsilon \,+\, \frac{1}{\varepsilon}\,\nabla_{\boldsymbol{x}} \cdot \int_{\mathbb{R}^d} \boldsymbol{v}\, f^\varepsilon \,\mathrm{d}\,\boldsymbol{v} \,=\, 0\,.$$

To overcome this difficulty, our strategy consists in considering the following re-scaled quantity

$$\pi^\varepsilon(t,\boldsymbol{x}) \,=\, \int_{\mathbb{R}^d} f^\varepsilon\,(t, \boldsymbol{x} - \varepsilon\,\boldsymbol{v}, \boldsymbol{v})\;\mathrm{d}\boldsymbol{v}\,.$$

On this simplified example, the advantage of considering π^ε instead of ρ^ε is straightforward as it turns out that π^ε is an exact solution of the limiting equation. Indeed, changing variables in the equation on f^ε and integrating with respect to $\boldsymbol{v}$, we obtain

$$\partial_t\,\pi^\varepsilon \,=\, \Delta_{\boldsymbol{x}}\,\pi^\varepsilon\,.$$

Therefore, the convergence analysis comes down to proving that π^ε is close to ρ^ε. It is possible to achieve this final step taking advantage of the following estimate

$$\|\,\rho^\epsilon \,-\, \pi^\varepsilon\,\|_{L^1\left(\mathbb{R}^d\right)} \,\leq\, \mathcal{A} \,+\, \mathcal{B}\,,$$

where $\mathcal{A}$ and $\mathcal{B}$ are defined as follows

$$\begin{cases} \mathcal{A} \,=\, \|\,\mathcal{M} \otimes \tau_{-\varepsilon\,\boldsymbol{v}}\,\rho^\varepsilon \,-\, \tau_{-\varepsilon\,\boldsymbol{v}}\,f^\varepsilon\,\|_{L^1\left(\mathbb{R}^{2d}\right)}\,, \\ \mathcal{B} \,=\, \displaystyle\int_{\mathbb{R}^d} \mathcal{M}(\tilde{\boldsymbol{v}})\;\|\,f^\varepsilon \,-\, \tau_{-\varepsilon\,\tilde{\boldsymbol{v}}}\,f^\varepsilon\,\|_{L^1\left(\mathbb{R}^{2d}\right)}\;\mathrm{d}\tilde{\boldsymbol{v}}\,, \end{cases}$$

and where $\tau_{\boldsymbol{x}_0}$ stands for the translation of vector $\boldsymbol{x}_0$ with respect to the $\boldsymbol{x}$-variable. To estimate $\mathcal{A}$, we use the first step, which ensures that f^ε is close to $\mathcal{M} \otimes \rho^\varepsilon$. Then, to estimate $\mathcal{B}$, it is sufficient to prove equicontinuity estimates for f^ε, that is

$$\|\,f^\varepsilon \,-\, \tau_{\boldsymbol{x}_0}\,f^\varepsilon\,\|_{L^1\left(\mathbb{R}^{2d}\right)} \,\lesssim\, |\,\boldsymbol{x}_0\,|\,.$$

In the forthcoming analysis, we adapt the latter arguments in our context.

3.1. ***A priori* estimates.** The main object of this section consists in deriving equicontinuity estimates for the sequence of solutions $(\nu^\varepsilon)_{\varepsilon>0}$ to equation (2.6). To obtain this result, we make use of the following key result

Lemma 3.1. *Consider δ in $\{0\,,\,1\}$ and smooth solutions f and g to the following equations*

$$\begin{cases} \partial_t\,f \,+\, \mathrm{div}_{\boldsymbol{y}}\,[\,\mathbf{a}\,(t\,,\,\boldsymbol{y}\,,\,\xi)\;f\,] \,+\, \lambda(t)\,\mathrm{div}_\xi\,[\,(\mathbf{b}_1 \,+\, \mathbf{b}_3)\,(t\,,\,\boldsymbol{y}\,,\,\xi)\;f\,] \,=\, \lambda(t)^2\,\Delta_\xi\,f\,, \\ \partial_t\,g \,+\, \mathrm{div}_{\boldsymbol{y}}\,[\,\mathbf{a}\,(t\,,\,\boldsymbol{y}\,,\,\xi)\;g\,] \,+\, \lambda(t)\,\mathrm{div}_\xi\,[\,\mathbf{b}_2\,(t\,,\,\boldsymbol{y}\,,\,\xi)\;g\,] \,=\, \delta\,\lambda(t)^2\,\Delta_\xi\,g\,. \end{cases}$$

set on the phase space $(t, \boldsymbol{y}, \xi) \in \mathbb{R}^+ \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, with $d_1 \geq 0$ and $d_2 \geq 1$, where

$$\left(\mathbf{a} \,:\, \mathbb{R}_+ \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \,\to\, \mathbb{R}^{d_1}\right) \;\text{ and }\; \left(\mathbf{b}_i \,:\, \mathbb{R}_+ \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \,\to\, \mathbb{R}^{d_2}\right)\,, \quad i \in \{1\,,\,2\,,\,3\}\,,$$

are given vector fields and where λ is a positive valued function. Suppose that f and g have positive values and are normalized as follows

$$\int_{\mathbb{R}^{d_1+d_2}} f \,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\xi \;=\; \int_{\mathbb{R}^{d_1+d_2}} g \,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\xi \;=\; 1\,.$$

Then it holds for all time $t \geq 0$

$$\|\, f(t) \,-\, g(t) \,\|_{L^1(\mathbb{R}^{d_1+d_2})} \;\leq\; 2\sqrt{2}\left(\|\, f_0 \,-\, g_0 \,\|^{\frac{1}{2}}_{L^1(\mathbb{R}^{d_1+d_2})} \,+\, \left(\int_0^t \mathcal{R}(s)\,\mathrm{d}s\right)^{\frac{1}{2}}\right), \tag{3.1}$$

where $\mathcal{R}$ is defined as

$$\mathcal{R}(t) \;=\; \int_{\mathbb{R}^{d_1+d_2}} \left(\frac{1}{4}\,|\,\mathbf{b}_1 \,-\, \mathbf{b}_2\,|^2 f \,+\, \lambda(t)\,|\,\mathrm{div}_\xi\,[\,\mathbf{b}_3\, g \,+\, (\delta \,-\, 1)\,\lambda\,\nabla_\xi\, g\,]\,|\right)(t\,,\,\boldsymbol{y}\,,\,\xi)\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\xi\,.$$

We postpone the proof of this result to Appendix A. Thanks to the latter lemma, we prove the following equicontinuity estimate for solutions to (2.6)

Proposition 3.2. *Consider a sequence $(\nu^\varepsilon)_{\varepsilon>0}$ of smooth solutions to equation (2.6) whose initial conditions meet assumption (2.12). There exists a positive constant C independent of ε such that for all $\varepsilon > 0$, it holds*

$$\|\,\nu^\varepsilon(t\,,\,\boldsymbol{x}) \,-\, \tau_{w_0}\,\nu^\varepsilon(t\,,\,\boldsymbol{x})\,\|_{L^1(\mathbb{R}^2)} \;\leq\; C\,e^{b\,t}\left(\left|\,e^{b\,t}\,w_0\,\right| \,+\, \left|\,e^{b\,t}\,w_0\,\right|^{\frac{1}{2}}\right), \qquad \forall\,(w_0\,,\,\boldsymbol{x}) \in \mathbb{R}\times K\,,$$

where C is explicitly given by

$$C \;=\; \sqrt{\max\left(8\,m_1\,,\,1\,/\,b\right)}\,,$$

with m_1 defined in assumption (2.12).

Proof. We fix some $\boldsymbol{x}$ in K and some positive ε. Then we consider some w_0 in $\mathbb{R}$ and the following re-scaled version f of ν^ε

$$f\,(t\,,\,w\,,\,v) \;=\; e^{-b\,t}\,\nu^\varepsilon\left(t\,,\,\boldsymbol{x}\,,\,v\,,\,e^{-b\,t}\,w\right).$$

We compute the equation solved by f performing the change of variable

$$w \;\mapsto\; e^{-b\,t}\,w \tag{3.2}$$

in equation (2.6), this yields

$$\partial_t\, f \,+\, \partial_w\left[e^{b\,t}\,A_0\,(\theta^\varepsilon\,v\,,\,0)\,f\right] \,+\, \frac{1}{\theta^\varepsilon}\,\partial_v\left[\,B_0^\varepsilon\left(t\,,\,\boldsymbol{x}\,,\,\theta^\varepsilon\,v\,,\,e^{-b\,t}\,w\right)f\right] \;=\; \frac{1}{|\theta^\varepsilon|^2}\,\mathcal{F}_{\rho_0^\varepsilon}\,[\,f\,]\,,$$

where A_0 and B_0^ε are given by (2.3). On top of that, we define $g := \tau_{w_0}\,f$, which solves the following equation

$$\partial_t\, g \,+\, \partial_w\left[e^{b\,t}\,A_0\,(\theta^\varepsilon\,v\,,\,0)\,g\right] \,+\, \frac{1}{\theta^\varepsilon}\,\partial_v\left[\,B_0^\varepsilon\left(t\,,\,\boldsymbol{x}\,,\,\theta^\varepsilon\,v\,,\,e^{-b\,t}\,(w+w_0)\right)g\right] \;=\; \frac{1}{|\theta^\varepsilon|^2}\,\mathcal{F}_{\rho_0^\varepsilon}\,[\,g\,]\,.$$

Thanks to the change of variable (3.2), we can apply Lemma 3.1 to f and g with the following parameters

$$\begin{cases} (\,\delta\,,\,\lambda\,) \;=\; (\,1\,,\,1\,/\,\theta^\varepsilon\,)\,, \\ \mathbf{a}\,(t\,,\,w\,,\,v) \;=\; e^{b\,t}\,A_0\,(\theta^\varepsilon\,v\,,\,0)\,, \\ \mathbf{b}_1\,(t\,,\,w\,,\,v) \;=\; B_0^\varepsilon\left(t\,,\,\boldsymbol{x}\,,\,\theta^\varepsilon\,v\,,\,e^{-b\,t}\,w\right) \,-\, \dfrac{1}{\theta^\varepsilon}\,v\,, \\ (\,\mathbf{b}_2\,,\,\mathbf{b}_3\,) \;=\; (\,\tau_{w_0}\,\mathbf{b}_1\,,\,0\,)\,. \end{cases}$$

According to (3.1) in Lemma 3.1, it holds

$$\|\, f(t) \,-\, g(t)\,\|_{L^1(\mathbb{R}^2)} \;\leq\; 2\sqrt{2}\,\|\, f_0 \,-\, g_0\,\|^{\frac{1}{2}}_{L^1(\mathbb{R}^2)} \,+\, \left(\frac{1 \,-\, e^{-2\,b\,t}}{b}\right)^{\frac{1}{2}}|w_0|\,, \quad \forall\, t \in \mathbb{R}^+\,.$$

Therefore, according to assumption (2.12), we obtain the result after inverting the change of variable (3.2) and taking the supremum over all $\boldsymbol{x}$ in K.

□

We conclude this section providing regularity estimates for the limiting distribution $\bar{\nu}$ with respect to the adaptation variable, which solves (1.7). The proof for this result is mainly computational since we have an explicit formula for the solutions to equation (1.7).

Proposition 3.3. *Consider some $\bar{\nu}_0$ satisfying assumption (2.11). The solution $\bar{\nu}$ to equation (1.7) with initial condition $\bar{\nu}_0$ verifies*

$$\| \bar{\nu}(t) \|_{L^\infty(K\,,W^{2\,,1}(\mathbb{R}))} \leq \exp(2\,b\,t)\, \| \bar{\nu}_0 \|_{L^\infty(K\,,W^{2\,,1}(\mathbb{R}))}\,, \quad \forall\, t \in \mathbb{R}^+\,,$$

and

$$\| w\, \partial_w\, \bar{\nu}(t) \|_{L^\infty(K\,,\,L^1(\mathbb{R}))} = \| w\, \partial_w\, \bar{\nu}_0 \|_{L^\infty(K\,,\,L^1(\mathbb{R}))}\,, \quad \forall\, t \in \mathbb{R}^+\,.$$

Proof. Since $\bar{\nu}$ solves (1.7), it is given by the following formula

$$\bar{\nu}(t,\, \boldsymbol{x}\,,\, w) \;=\; e^{bt}\, \bar{\nu}_0 \left(\boldsymbol{x}\,,\, e^{bt}\, w \right)\,, \quad \forall\, (t, \boldsymbol{x}) \,\in\, \mathbb{R}^+ \times K\,.$$

Consequently, we easily obtain the expected result. □

We are now ready to prove the first convergence result on ν^ε.

3.2. **Proof of Theorem 2.1.** The proof is divided in three steps. First, we prove that the solution ν^ε to (2.6) converges to the local equilibrium

$$\mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}^\varepsilon\,,$$

by applying a rather classical entropy estimate. Then, as in the example developed at the beginning of Section 3, we introduce an intermediate quantity $\bar{g}^\varepsilon$, which converges to the solution $\bar{\nu}$ to equation (1.7). At last, we prove that $\bar{g}^\varepsilon$ is close to $\bar{\nu}^\varepsilon$ thanks to the equicontinuity estimate given in Proposition 3.2 and therefore conclude that the marginal $\bar{\nu}^\varepsilon$ converges towards $\bar{\nu}$.

3.2.1. *Convergence of ν^ε towards $\mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}^\varepsilon$: relative entropy estimate.* This step relies on a rather classical relative entropy estimate for the solution ν^ε to equation (2.6): we investigate the time evolution of the relative entropy

$$H\left[\, \nu^\varepsilon \,|\, \mathcal{M}_{\rho_0^\varepsilon} \,\right] \;=\; \int_{\mathbb{R}^2} \nu^\varepsilon\, \ln\left(\frac{\nu^\varepsilon}{\mathcal{M}_{\rho_0^\varepsilon}} \right) \mathrm{d}\boldsymbol{u}\,,$$

along the trajectories of equation (2.6).

Proposition 3.4. *Under assumptions (1.1a)-(1.1b) on the drift N and (1.2) on the interaction kernel Ψ, consider a sequence of solutions $(\mu^\varepsilon)_{\varepsilon>0}$ to (1.3) with initial conditions satisfying assumptions (2.7)-(2.8b) and (2.13). Then for all $\varepsilon \leq 1$, it holds*

$$\int_0^t \left\| \nu^\varepsilon \,-\, \mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}^\varepsilon \right\|_{L^\infty_{\boldsymbol{x}} L^1_{\boldsymbol{u}}} (s)\, \mathrm{d}s \;\leq\; \sqrt{\varepsilon} \left(2\, m_2\, \sqrt{t} \;+\; C\,(t+1) \right)\,, \quad \forall\, t \,\geq\, 0\,,$$

where m_2 is given in assumption (2.13). In this result, the constant C only depends on m_*, m_p and $\overline{m}_p$ (see assumptions (2.7)-(2.8b)) and the data of the problem N, Ψ and A_0.

Proof. All along this proof, we choose some $\boldsymbol{x}$ lying in K and we omit the dependence with respect to $(t\,,\,\boldsymbol{x})$ when the context is clear. We compute the time derivative of $H\left[\, \nu^\varepsilon \,|\, \mathcal{M}_{\rho_0^\varepsilon} \,\right]$ multiplying equation (2.6) by $\ln\left(\nu^\varepsilon \,/\, \mathcal{M}_{\rho_0^\varepsilon} \right)$. After integrating by part the stiffer term, it yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\, H\left[\, \nu^\varepsilon \,|\, \mathcal{M}_{\rho_0^\varepsilon} \,\right] \;+\; \frac{1}{|\theta^\varepsilon|^2}\, I\left[\, \nu^\varepsilon \,|\, \mathcal{M}_{\rho_0^\varepsilon} \,\right] \;=\; \mathcal{A}\,,$$

where the Fisher information is defined by

$$I\left[\, \nu^\varepsilon \,|\, \mathcal{M}_{\rho_0^\varepsilon} \,\right] \;:=\; \int_{\mathbb{R}^2} \left| \partial_v \ln\left(\frac{\nu^\varepsilon}{\mathcal{M}_{\rho_0^\varepsilon}} \right) \right|^2 \nu^\varepsilon\, \mathrm{d}\boldsymbol{u}\,,$$

and where $\mathcal{A}$ is given by

$$\mathcal{A} \;=\; -\int_{\mathbb{R}^2} \mathrm{div}_{\boldsymbol{u}} \left[\, \mathbf{b}_0^\varepsilon\, \nu^\varepsilon \,\right] \ln\left(\frac{\nu^\varepsilon}{\mathcal{M}_{\rho_0^\varepsilon}} \right) \mathrm{d}\boldsymbol{u}\,.$$

After an integration by part, $\mathcal{A}$ rewrites as follows

$$\mathcal{A} \;=\; \frac{1}{\theta^\varepsilon} \int_{\mathbb{R}^2} B_0^\varepsilon\,(\,t\,,\,\boldsymbol{x}\,,\,\theta^\varepsilon\, v,\, w\,)\; \partial_v \left[\ln\left(\frac{\nu^\varepsilon}{\mathcal{M}_{\rho_0^\varepsilon}} \right) \right] \nu^\varepsilon\; \mathrm{d}\boldsymbol{u} \;+\; b\,,$$

where B_0^ε is given by (2.3). According to items (2) and (4) in Proposition 2.5 , $\mathcal{V}^\varepsilon$ and $\mathcal{E}(\mu^\varepsilon)$ are uniformly bounded with respect to both $(t, \boldsymbol{x}) \in \mathbb{R}^+ \times K$ and $\varepsilon > 0$. On top of that, according to assumptions (1.2) and (2.7) on Ψ and ρ_0^ε, $\Psi *_r \rho_0^\varepsilon$ is uniformly bounded with respect to both $\boldsymbol{x} \in K$ and $\varepsilon > 0$. Consequently, applying Young's inequality, assumption (1.1b) and since N is locally Lipschitz, we obtain

$$\mathcal{A} \leq \frac{1}{2\,|\theta^\varepsilon|^2}\, I\left[\,\nu^\varepsilon \,|\, \mathcal{M}_{\rho_0^\varepsilon}\,\right] + C\left(1 + \int_{\mathbb{R}^2} \left(\,|\,\theta^\varepsilon\, v\,|^{2p} + w^2\,\right) \nu^\varepsilon \,\mathrm{d}\boldsymbol{u}\right),$$

for some positive constant C only depending on m_*, m_p, $\overline{m}_p$ and the data of the problem: N, A and Ψ. Then we invert the change of variable (2.1) in the integral in the right-hand side of the latter inequality and apply item (2) in Proposition 2.5. In the end, it yields

$$\mathcal{A} \leq \frac{1}{2\,|\theta^\varepsilon|^2}\, I\left[\,\nu^\varepsilon \,|\, \mathcal{M}_{\rho_0^\varepsilon}\,\right] + C\,.$$

Consequently, we end up with following differential inequality

$$\frac{\mathrm{d}}{\mathrm{d}t}\, H\left[\,\nu^\varepsilon \,|\, \mathcal{M}_{\rho_0^\varepsilon}\,\right] + \frac{1}{2\,|\theta^\varepsilon|^2}\, I\left[\,\nu^\varepsilon \,|\, \mathcal{M}_{\rho_0^\varepsilon}\,\right] \leq C\,,$$

Then we substitute I with H in the latter inequality according to the logarithmic Sobolev inequality for Gaussian measures which reads as follows (see [16])

$$2\, H\left[\,\nu^\varepsilon \,|\, \mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}^\varepsilon\,\right] \leq I\left[\,\nu^\varepsilon \,|\, \mathcal{M}_{\rho_0^\varepsilon}\,\right],$$

and we integrate between 0 and t to get

$$\int_0^t \frac{1}{|\theta^\varepsilon(s)|^2}\, H\left[\,\nu^\varepsilon \,|\, \mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}^\varepsilon\,\right] \mathrm{d}s \leq H\left[\,\nu_0^\varepsilon \,|\, \mathcal{M}_{\rho_0^\varepsilon}\,\right] + C\,t\,.$$

Furthermore, according to the explicit formula (2.5) for θ^ε it holds

$$\frac{1}{2\,\varepsilon} \leq |\theta^\varepsilon(s)|^{-2}\,,$$

as long as ε is less than 1 and s is greater than T^ε, where T^ε is given

$$T^\varepsilon = \frac{\varepsilon}{2\,m_*}\,|\ln(\varepsilon)|\,.$$

Consequently, we obtain

$$\int_{T^\varepsilon}^t H\left[\,\nu^\varepsilon(s\,,\boldsymbol{x}) \,|\, \mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}^\varepsilon(s\,,\boldsymbol{x})\,\right] \mathrm{d}s \leq 2\,\varepsilon\, H\left[\,\nu_0^\varepsilon(\boldsymbol{x})\,\right] + C\,\varepsilon\,(t+1)\,.$$

Then we substitute the relative entropy with the L^1 norm according to Csizar-Kullback inequality

$$\left\|\,\nu^\varepsilon - \mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}^\varepsilon\,\right\|^2_{L^1(\mathbb{R}^2)} \leq 2\, H\left[\,\nu^\varepsilon \,|\, \mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}^\varepsilon\,\right],$$

and take the supremum over all $\boldsymbol{x}$ in K. After taking the square root, it yields

$$\int_{T^\varepsilon}^t \left\|\,\nu^\varepsilon(s) - \mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}^\varepsilon\,\right\|_{L^\infty_{\boldsymbol{x}} L^1_{\boldsymbol{u}}} \mathrm{d}s \leq 2\sqrt{\varepsilon\, t}\, m_2 + C\sqrt{\varepsilon}\,(t+1)\,, \quad \forall\, t \geq 0\,,$$

where m_2 is given in (2.13). To conclude, we notice that since equation (2.6) is conservative, it holds

$$\int_0^{T^\varepsilon} \left\|\,\nu^\varepsilon(s) - \mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}^\varepsilon\,\right\|_{L^\infty_{\boldsymbol{x}} L^1_{\boldsymbol{u}}} \mathrm{d}s \leq 2\,T^\varepsilon \leq C\sqrt{\varepsilon}\,.$$

□

3.2.2. *Convergence of $\bar{g}^\varepsilon$ towards $\bar{\nu}$.* As in the example developed at the beginning of this section, we consider the following re-scaled version g^ε of ν^ε

$$\nu^\varepsilon\,(t\,,\,\boldsymbol{x}\,,\,v\,,\,w)\;=\;g^\varepsilon\,(t\,,\,\boldsymbol{x}\,,\,v\,,\,w\,+\,\gamma^\varepsilon(t,\boldsymbol{x})\,v)\;,$$

where γ^ε is given by

$$\gamma^\varepsilon\,(t\,,\,\boldsymbol{x})\;=\;\frac{a\,\varepsilon}{\rho_0^\varepsilon(\boldsymbol{x})}\,\theta^\varepsilon\,(t\,,\,\boldsymbol{x})\;.$$

Operating the following change of variable in equation (2.6)

$$(t\,,\,v\,,\,w)\mapsto(t\,,\,v\,,\,w\,+\,\gamma^\varepsilon\,v)\;, \tag{3.3}$$

and integrating the equation with respect to v, the equation on the marginal $\bar{g}^\varepsilon$ of g^ε defined as

$$\bar{g}^\varepsilon\,(t\,,\,\boldsymbol{x}\,,\,w)\;=\;\int_{\mathbb{R}}g^\varepsilon\,(t\,,\,\boldsymbol{x}\,,\,v\,,\,w)\,\mathrm{d}v\,,$$

reads as follows

$$\partial_t\,\bar{g}^\varepsilon\;+\;\frac{a\,\varepsilon}{\rho_0^\varepsilon}\,\partial_w\left[\int_{\mathbb{R}}(B_0^\varepsilon\,(t,\,\boldsymbol{x},\,\theta^\varepsilon\,v,\,w-\gamma^\varepsilon\,v)+b\,\theta^\varepsilon\,v)\,g^\varepsilon\,\mathrm{d}v\right]\;-\;\left(\frac{a\,\varepsilon}{\rho_0^\varepsilon}\right)^2\,\partial_w^2\,\bar{g}^\varepsilon\;=\;\partial_w\,[\,b\,w\,\bar{g}^\varepsilon\,]\;, \tag{3.4}$$

where B_0^ε is defined by (2.3). As in our example, the equation on $\bar{g}^\varepsilon$ is consistent with the limiting equation (1.7) as ε vanishes, this enables to prove that $\bar{g}^\varepsilon$ converges towards $\bar{\nu}$

Proposition 3.5. *Under assumptions (1.1a)-(1.1b) on the drift N and (1.2) on the interaction kernel Ψ, consider a sequence of solutions $(\mu^\varepsilon)_{\varepsilon>0}$ to (1.3) with initial conditions satisfying assumption (2.7)-(2.8b) and the solution $\bar{\nu}$ to equation (1.7) with initial condition $\bar{\nu}_0$ satisfying assumption (2.11). There exists a positive constant C independent of ε such that for all ε less than 1, it holds*

$$\|\,\bar{g}^\varepsilon(t)\;-\;\bar{\nu}(t)\|_{L^\infty_{\boldsymbol{x}}L^1_w}\;\leq\;2\,\sqrt{2}\,\|\,\bar{g}_0^\varepsilon\;-\;\bar{\nu}_0\|^{\frac{1}{2}}_{L^\infty_{\boldsymbol{x}}L^1_w}\;+\;C\,e^{b\,t}\,\sqrt{\varepsilon}\,,\quad\forall\,t\in\mathbb{R}^+\,.$$

In this result, the constant C only depends on m_*, m_p and $\overline{m}_p$ (see assumptions (2.7)-(2.8b)) and the data of the problem $\bar{\nu}_0$, N, Ψ and A_0.

Proof. All along this proof, we choose some $\boldsymbol{x}$ lying in K and some positive ε; we omit the dependence with respect to $(t,\boldsymbol{x})$ when the context is clear. Since $\bar{\nu}$ and $\bar{g}^\varepsilon$ solve respectively equations (1.7) and (3.4), Lemma 3.1 applies with the following parameters

$$\left\{\begin{array}{l}(\,d_1\,,\,d_2\,,\,\delta\,,\,\lambda\,)\;=\;(\,0\,,\,1\,,\,0\,,\,a\,\varepsilon\,/\,\rho_0^\varepsilon\,)\;,\\[2mm]\mathbf{b}_2\,(t\,,\,w)\;=\;-\,\dfrac{\rho_0^\varepsilon}{a\,\varepsilon}\,b\,w\,,\\[2mm]\mathbf{b}_1\,(t\,,\,w)\;=\;\mathbf{b}_2\,(t\,,\,w)\;+\;\displaystyle\int_{\mathbb{R}}(B_0^\varepsilon\,(t,\,\boldsymbol{x},\,\theta^\varepsilon\,v,\,-\gamma^\varepsilon\,v)+b\,\theta^\varepsilon\,v)\,\frac{g^\varepsilon}{\bar{g}^\varepsilon}\,\mathrm{d}v\;,\\[2mm]\mathbf{b}_3\,(t\,,\,w)\;=\;-\,w\,,\end{array}\right.$$

where B_0^ε is given by (2.3). According to (3.1) in Lemma 3.1, it holds

$$\|\,\bar{g}^\varepsilon(t)\;-\;\bar{\nu}(t)\,\|_{L^1(\mathbb{R})}\;\leq\;2\,\sqrt{2}\left(\|\,\bar{g}_0^\varepsilon\;-\;\bar{\nu}_0\,\|^{\frac{1}{2}}_{L^1(\mathbb{R})}\;+\;\left(\int_0^t\mathcal{R}_1(s)\;+\;\mathcal{R}_2(s)\,\mathrm{d}s\right)^{\frac{1}{2}}\right)\,,\quad\forall\,t\in\mathbb{R}^+\,,$$

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

$$\left\{\begin{array}{l}\mathcal{R}_1(t)\;=\;\dfrac{1}{4}\displaystyle\int_{\mathbb{R}}\left|\int_{\mathbb{R}}(B_0^\varepsilon\,(t\,,\,\boldsymbol{x}\,,\,\theta^\varepsilon\,v\,,\,-\gamma^\varepsilon\,v)\;+\;b\,\theta^\varepsilon\,v\,)\,\frac{g^\varepsilon}{\bar{g}^\varepsilon}\,\mathrm{d}v\right|^2\,\bar{g}^\varepsilon\mathrm{d}w\,,\\[2mm]\mathcal{R}_2(t)\;=\;\dfrac{a\,\varepsilon}{\rho_0^\varepsilon}\displaystyle\int_{\mathbb{R}}|\,\partial_w\,[\,w\,\bar{\nu}\,]|\;+\;\frac{a\,\varepsilon}{\rho_0^\varepsilon}\,\left|\,\partial_w^2\,\bar{\nu}\,\right|\,\mathrm{d}w\,.\end{array}\right.$$

We estimate $\mathcal{R}_1$ according to Jensen's inequality

$$\mathcal{R}_1(t)\;=\;\frac{1}{4}\int_{\mathbb{R}^2}|B_0^\varepsilon\,(t\,,\,\boldsymbol{x}\,,\,\theta^\varepsilon\,v\,,\,-\gamma^\varepsilon\,v)\;+\;b\,\theta^\varepsilon\,v|^2\,\bar{g}^\varepsilon\,\mathrm{d}v\,\mathrm{d}w\,.$$

Then we bound B_0^ε: on the one hand $\mathcal{V}^\varepsilon$ is uniformly bounded with respect to both $(t,\boldsymbol{x})\in\mathbb{R}^+\times K$ and $\varepsilon\,>\,0$ according (2) in Proposition 2.5, on the other hand according to assumptions (1.2) and

(2.7) on Ψ and ρ_0^ε, $\Psi *_r \rho_0^\varepsilon$ is uniformly bounded with respect to both $\boldsymbol{x} \in K$ and $\varepsilon > 0$. Consequently, applying Young's inequality, assumption (1.1b) and since N is locally Lipschitz, we obtain

$$\mathcal{R}_1(t) \leq C \int_{\mathbb{R}^2} \left(|\, \theta^\varepsilon\, v\,|^{2p} + |\, \theta^\varepsilon\, v\,|^2 + |\, \mathcal{E}\,(\mu^\varepsilon)\,|^2 \right) g^\varepsilon \mathrm{d}\boldsymbol{u}\,,$$

as long as ε is less than 1 to ensure that γ^ε given by (3.3) is less than $a\ \theta^\varepsilon\,/\,m_*$. We invert the changes of variable (3.3) and (2.1) and apply items (3) and (4) in Proposition 2.5, this yields

$$\mathcal{R}_1(t) \leq C \left(e^{-2\rho_0^\varepsilon(\boldsymbol{x})t/\varepsilon} + \varepsilon \right).$$

Then to estimate $\mathcal{R}_2$, we apply Proposition 3.3 to bound $\bar{\nu}$ and assumption (2.7) to lower bound ρ_0^ε, it yields

$$\mathcal{R}_2(t) \leq C\,\varepsilon\, e^{2\,b\,t}\,.$$

We gather the former computations and take the supremum over all $\boldsymbol{x}$ in K: it yields the expected result. □

3.2.3. *Convergence of $\bar{\nu}^\varepsilon$ towards $\bar{\nu}$.* In this section, we gather the result from the last steps to deduce that $\bar{\nu}^\varepsilon$ converges towards $\bar{\nu}$.

Proposition 3.6. *Under the assumptions of Theorem 2.1, there exists a positive constant C independent of ε such that for all ε less than 1, it holds*

$$\int_0^t \|\, \bar{\nu}^\varepsilon - \bar{\nu}\,\|_{L^\infty_{\boldsymbol{x}} L^1_w}\, \mathrm{d}s \leq 2\sqrt{2}\,t\, \|\, \bar{\nu}_0^\varepsilon - \bar{\nu}_0\,\|^{\frac{1}{2}}_{L^\infty_{\boldsymbol{x}} L^1_w} + 2\sqrt{\varepsilon\,t}\,m_2 + C\,e^{2\,b\,t}\sqrt{\varepsilon}\,, \quad \forall\, t \in \mathbb{R}^+.$$

In this result, the constant C only depends on m_1, m_*, m_p and $\overline{m}_p$ (see assumptions (2.7)-(2.8b)) and the data of the problem $\bar{\nu}_0$, N, Ψ and A_0.

Proof. All along this proof, we omit the dependence with respect to $(t, \boldsymbol{x})$ when the context is clear. We consider the following triangular inequality

$$\|\, \bar{\nu}^\varepsilon - \bar{\nu}\,\|_{L^\infty_{\boldsymbol{x}} L^1_w} \leq \mathcal{A} + \mathcal{B}\,,$$

where $\mathcal{A}$ and $\mathcal{B}$ are given by

$$\begin{cases} \mathcal{A} = \|\, \bar{\nu}^\varepsilon - \bar{g}^\varepsilon\,\|_{L^\infty_{\boldsymbol{x}} L^1_w}\,, \\ \mathcal{B} = \|\, \bar{g}^\varepsilon - \bar{\nu}\,\|_{L^\infty_{\boldsymbol{x}} L^1_w}\,. \end{cases}$$

We estimate $\mathcal{B}$ applying Proposition 3.5, which ensures

$$\mathcal{B} \leq 2\sqrt{2}\,\|\, \bar{g}_0^\varepsilon - \bar{\nu}_0\|^{\frac{1}{2}}_{L^\infty_{\boldsymbol{x}} L^1_w} + C\,e^{b\,t}\sqrt{\varepsilon}\,,$$

at all times t in $\mathbb{R}^+$. Then we notice that inverting the change of variable (3.3), it holds

$$\|\, \bar{g}_0^\varepsilon - \bar{\nu}_0\|_{L^\infty_{\boldsymbol{x}} L^1_w} \leq \|\, \bar{\nu}_0^\varepsilon - \bar{\nu}_0\|_{L^\infty_{\boldsymbol{x}} L^1_w} + \int_{\mathbb{R}} \mathcal{M}(v) \left\| \bar{\nu}_0 - \tau_{(\gamma_0^\varepsilon\, v)}\, \bar{\nu}_0 \right\|_{L^\infty_{\boldsymbol{x}} L^1_w} \mathrm{d}v\,.$$

Therefore, according to the regularity assumption (2.11) on $\bar{\nu}_0$, we obtain for all $\varepsilon \leq 1$.

$$\mathcal{B} \leq 2\sqrt{2}\,\|\, \bar{\nu}_0^\varepsilon - \bar{\nu}_0\|^{\frac{1}{2}}_{L^\infty_{\boldsymbol{x}} L^1_w} + C\,e^{b\,t}\sqrt{\varepsilon}\,,$$

To estimate $\mathcal{A}$, we invert the change of variable (3.3) and apply the triangular inequality. It yields

$$\mathcal{A} \leq \mathcal{A}_1 + \mathcal{A}_2\,,$$

where $\mathcal{A}_1$ and $\mathcal{A}_2$ are defined as follows

$$\begin{cases} \mathcal{A}_1 = \left\| \mathcal{M}_{\rho_0^\varepsilon} \otimes \tau_{-(\gamma^\varepsilon\, v)}\, \bar{\nu}^\varepsilon - \tau_{-(\gamma^\varepsilon\, v)}\, \nu^\varepsilon \right\|_{L^\infty_{\boldsymbol{x}} L^1_{\boldsymbol{u}}}\,, \\ \mathcal{A}_2 = \sup_{\boldsymbol{x} \in K} \displaystyle\int_{\mathbb{R}} \mathcal{M}_{\rho_0^\varepsilon}(v)\, \left\|\, \nu^\varepsilon - \tau_{-(\gamma^\varepsilon\, v)}\, \nu^\varepsilon\,\right\|_{L^1(\mathbb{R}^2)} \mathrm{d}v\,. \end{cases}$$

According to Proposition 3.4, it holds for all $t \geq 0$

$$\int_0^t \mathcal{A}_1(s)\,\mathrm{d}s \leq 2\sqrt{\varepsilon\,t}\,m_2 + C\sqrt{\varepsilon}\,(t+1)\,.$$

To estimate $\mathcal{A}_2$, we apply Proposition 3.2, it yields

$$\mathcal{A}_2 \,=\, C\sqrt{\varepsilon}\, e^{2\,b\,t}\,.$$

We obtain the result gathering the former estimates. □

Theorem 2.1 is obtained taking the sum of the estimates in Propositions 3.6 and 3.4.

4. Convergence analysis in weighted L^2 spaces

In this section, we derive convergence estimates for μ^ε in a weighted L^2 functional framework. We take advantage of the variational structure of L^2 spaces in order to derive uniform regularity estimates for μ^ε. These key estimates are the object of the following section

4.1. *A priori* estimates.

The main purpose of this section is to propagate the $\mathscr{H}^k$-norms along the trajectories of equation (2.6) uniformly with respect to ε. We outline the strategy in the case of the $\mathscr{H}^0$-norm. Its time derivative along the trajectories of equation (2.6) is obtained multiplying (2.6) by $\nu^\varepsilon\, m^\varepsilon$ and integrating with respect to $\boldsymbol{u}$

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\,\|\,\nu^\varepsilon\,\|^2_{L^2(m^\varepsilon)} \,=\, \frac{1}{(\theta^\varepsilon)^2}\,\big\langle\,\mathcal{F}_{\rho_0^\varepsilon}[\,\nu^\varepsilon\,]\,,\,\nu^\varepsilon\,\big\rangle_{L^2(m^\varepsilon)} \,-\, \big\langle\,\mathrm{div}_{\boldsymbol{u}}\,[\,\mathbf{b}_0^\varepsilon\,\nu^\varepsilon\,]\,,\,\nu^\varepsilon\,\big\rangle_{L^2(m^\varepsilon)}\,.$$

We first point out that according to the following lemma, the term associated to the Fokker-Planck operator is dissipative and is consequently a helping term in the upcoming analysis

Lemma 4.1. *For all $\boldsymbol{x}$ in K, it holds*

$$\Big\langle\,\mathcal{F}_{\rho_0^\varepsilon(\boldsymbol{x})}[\,\nu\,]\,,\,\nu\,\Big\rangle_{L^2(m^\varepsilon_{\boldsymbol{x}})} \,=\, -\,\mathcal{D}_{\rho_0^\varepsilon(\boldsymbol{x})}[\,\nu\,] \,\le\, 0\,,$$

for all $\nu \in H^1(m^\varepsilon_{\boldsymbol{x}})$, where the dissipation $\mathcal{D}_{\rho_0^\varepsilon}$ is given by

$$\mathcal{D}_{\rho_0^\varepsilon(\boldsymbol{x})}[\,\nu\,] \,=\, \int_{\mathbb{R}^2} |\partial_v\,(\,\nu\, m^\varepsilon_{\boldsymbol{x}}\,)|^2\,\,(m^\varepsilon_{\boldsymbol{x}})^{-1}\,\mathrm{d}\boldsymbol{u} \,\ge\, 0\,.$$

Proof. The Fokker-Planck operator rewrites as follows

$$\mathcal{F}_{\rho_0^\varepsilon(\boldsymbol{x})}[\,\nu\,] \,=\, \partial_v\Big[\,(m^\varepsilon_{\boldsymbol{x}})^{-1}\,\partial_v\,(\nu\, m^\varepsilon_{\boldsymbol{x}})\,\Big]\,.$$

Consequently, the result is obtained integrating by part $\mathcal{F}_{\rho_0^\varepsilon(\boldsymbol{x})}[\,\nu\,]$ against $\nu\, m^\varepsilon$ with respect to $\boldsymbol{u}$. □

Therefore, the main challenge is to control the contribution of the transport operator $\mathrm{div}_{\boldsymbol{u}}\,\mathbf{b}_0^\varepsilon$ with the dissipation $\mathcal{D}_{\rho_0^\varepsilon}$ brought by the Fokker-Planck operator, which is done in the following lemma.

Lemma 4.2. *Under assumptions (1.1a)-(1.1b) and (2.17) on the drift N, (1.2) on the interaction kernel Ψ, consider a sequence of solutions $(\mu^\varepsilon)_{\varepsilon>0}$ to (1.3) with initial conditions satisfying assumptions (2.7)-(2.8b). Then, there exists a constant $C>0$ such that for all $\varepsilon>0$ and any α greater than $1/(2b\kappa)$, we have*

$$-\,\big\langle\,\mathrm{div}_{\boldsymbol{u}}\,[\,\mathbf{b}_0^\varepsilon\,\nu\,]\,,\,\nu\,\big\rangle_{L^2(m^\varepsilon_{\boldsymbol{x}})} \,\le\, \frac{\alpha}{(\theta^\varepsilon)^2}\,\mathcal{D}_{\rho_0^\varepsilon(\boldsymbol{x})}[\,\nu\,] \,+\, C\,\|\,\nu\,\|^2_{L^2(m^\varepsilon_{\boldsymbol{x}})}\,,$$

for all $(t\,,\,\boldsymbol{x}) \in \mathbb{R}^+\times K$ and all $\nu \in H^1(m^\varepsilon_{\boldsymbol{x}})$.

Before getting into the heart of the proof, we point out that as long as the latter lemma holds with some α less than 1, we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\,\|\,\nu^\varepsilon\,\|^2_{L^2(m^\varepsilon)} \,\le\, C\,\|\,\nu^\varepsilon\,\|^2_{L^2(m^\varepsilon)}\,,$$

which ensures that the $\mathscr{H}^0$-norm is propagated along the curves of (2.6) uniformly with respect to ε. We follow the exact same strategy in order to propagate the $\mathscr{H}^k$-norms when k is not 0: see Proposition 4.5 for more details. On top of that, we emphasize that the constraint (2.18) on κ in Theorem 2.3 arises from the lower bound on α in Lemma 4.2.

Remark 4.3. *Due to the structure of the space $L^2(m^\varepsilon)$, we added the confining assumption (2.17) on the drift N to Theorem 2.3. Our proof of Lemma 4.2 crucially relies on this assumption ; it is the only time that we use it as well.*

Proof. All along this proof, we consider some $\varepsilon > 0$ and some $(t, \boldsymbol{x})$ in $\mathbb{R}_+ \times K$; we omit the dependence with respect to $\boldsymbol{x}$ when the context is clear. Furthermore, we choose some ν in $H^1(m^\varepsilon_{\boldsymbol{x}})$. Since $\mathbf{b}^\varepsilon_{\mathbf{0}}$ is given by (2.3), we have

$$-\left\langle \operatorname{div}_{\boldsymbol{u}}\left[\mathbf{b}^\varepsilon_0\,\nu\right],\nu\right\rangle_{L^2(m^\varepsilon)} = \mathcal{A}_1+\mathcal{A}_2+\mathcal{A}_3\,,$$

where

$$\begin{cases} \mathcal{A}_1 = \dfrac{1}{\theta^\varepsilon}\displaystyle\int_{\mathbb{R}^2}\partial_v\left[w\,\nu\right]\nu\,m^\varepsilon\,\mathrm{d}\boldsymbol{u} - \int_{\mathbb{R}^2}\partial_w\left[A_0\left(\theta^\varepsilon v,\,w\right)\nu\right]\nu\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,,\\ \mathcal{A}_2 = -\dfrac{1}{\theta^\varepsilon}\displaystyle\int_{\mathbb{R}^2}\partial_v\left[\left(N(\mathcal{V}^\varepsilon+\theta^\varepsilon v)-N(\mathcal{V}^\varepsilon)\right)\nu\right]\nu\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,,\\ \mathcal{A}_3 = \displaystyle\int_{\mathbb{R}^2}\partial_v\left[\left(v\,\Psi*_r\rho^\varepsilon_0+\mathcal{E}(\mu^\varepsilon)\,(\theta^\varepsilon)^{-1}\right)\nu\right]\nu\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,. \end{cases}$$

To estimate $\mathcal{A}_1$, we take advantage of the confining properties of A. When it comes to $\mathcal{A}_2$, the estimate relies on the confining properties of the non-linear drift N. The last term $\mathcal{A}_3$ gathers the lower order terms and adds no difficulty.

We start with A_1, which rewrites as follows after exact computations and an integration by part,

$$\mathcal{A}_1 = -\int_{\mathbb{R}^2} w\,\nu\,(m^\varepsilon)^{1/2}\,\frac{1}{\theta^\varepsilon}\,\partial_v\left[\nu\,m^\varepsilon\right](m^\varepsilon)^{-1/2}\,\mathrm{d}\boldsymbol{u}+\frac{1}{2}\int_{\mathbb{R}^2}\left(\kappa w\,A_0\left(\theta^\varepsilon v,\,w\right)-\partial_w A_0\right)|\nu|^2\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,.$$

According to the definition of A_0 and applying Young's inequality, we obtain

$$\mathcal{A}_1 \le \frac{1}{2\,\eta_2\,(\theta^\varepsilon)^2}\,\mathcal{D}_{\rho^\varepsilon_0}[\,\nu\,]+\int_{\mathbb{R}^2}\left(\frac{C}{\eta_1}\,|\theta^\varepsilon v|^2+\left(C\eta_1+\frac{\eta_2-b\kappa}{2}\right)w^2\right)|\nu|^2\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}+C\|\nu\|^2_{L^2(m^\varepsilon)}\,,$$

for all positive η_1 and η_2 and for some positive constant C. We set $\alpha_- = (\alpha + 1/(2b\kappa))/2$, $\eta_2 = 1/(2\alpha_-)$ and $\eta_1 = (b\kappa-\eta_2)/(2C)$ which is positive according to the condition on α in Lemma 4.2. With this choice, we have $C\eta_1+(\eta_2-b\kappa)/2 = 0$ and consequently, we obtain

$$\mathcal{A}_1 \le \frac{\alpha_-}{(\theta^\varepsilon)^2}\,\mathcal{D}_{\rho^\varepsilon_0}[\,\nu\,]+C\int_{\mathbb{R}^2}|\theta^\varepsilon v|^2\,|\nu|^2\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}+C\|\nu\|^2_{L^2(m^\varepsilon)}\,, \tag{4.1}$$

for some positive constant C only depending on A_0, κ and α.

To estimate $\mathcal{A}_2$, we take advantage of the super-linear decay of N (see assumption (1.1a)) in order to control the terms growing at most linearly. We emphasize that the decaying property of N is prescribed at infinity. Consequently, it may not have confining property on bounded sets. Hence, the main point here consists in isolating the domain where N decays super linearly.
After some exact computations and an integration by part, $\mathcal{A}_2$ rewrites

$$\mathcal{A}_2 = \frac{1}{2}\int_{\mathbb{R}^2}\left[\frac{\rho^\varepsilon_0}{\theta^\varepsilon}\,v\,\left(N(\mathcal{V}^\varepsilon+\theta^\varepsilon v)-N(\mathcal{V}^\varepsilon)\right)-N'\left(\mathcal{V}^\varepsilon+\theta^\varepsilon v\right)\right]|\nu|^2\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,.$$

We consider some $R>0$ and split the former expression in three different parts

$$\mathcal{A}_2 = \mathcal{A}_{21}+\mathcal{A}_{22}+\mathcal{A}_{23}\,,$$

where

$$\begin{cases} \mathcal{A}_{21} = \dfrac{\rho^\varepsilon_0}{2\theta^\varepsilon}\displaystyle\int_{\mathbb{R}^2}\mathbb{1}_{|\theta^\varepsilon v|>R}\,v\,\left(N(\mathcal{V}^\varepsilon+\theta^\varepsilon v)-N(\mathcal{V}^\varepsilon)\right)|\nu|^2\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,,\\ \mathcal{A}_{22} = \dfrac{1}{2}\displaystyle\int_{\mathbb{R}^2}\mathbb{1}_{|\theta^\varepsilon v|\le R}\left[\frac{\rho^\varepsilon_0}{\theta^\varepsilon}\,v\,(N(\mathcal{V}^\varepsilon+\theta^\varepsilon v)-N(\mathcal{V}^\varepsilon))-N'\left(\mathcal{V}^\varepsilon+\theta^\varepsilon v\right)\right]|\nu|^2\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,,\\ \mathcal{A}_{23} = -\dfrac{1}{2}\displaystyle\int_{\mathbb{R}^2}\mathbb{1}_{|\theta^\varepsilon v|>R}\,N'\left(\mathcal{V}^\varepsilon+\theta^\varepsilon v\right)|\nu|^2\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,. \end{cases}$$

The first term $\mathcal{A}_{21}$ corresponds to the the contribution of N on the domain where it decays super-linearly. Consequently, $\mathcal{A}_{21}$ is non positive for R great enough. We take advantage of the helping term $\mathcal{A}_{21}$ to control $\mathcal{A}_{22}$, which corresponds to the contribution of N on bounded sets. We estimate $\mathcal{A}_3$ taking advantage of the confining term $\mathcal{A}_{21}$ coupled with the confining assumption (2.17) on N.

Let us estimate $\mathcal{A}_{21}$. According to item (2) in Proposition 2.5, $\mathcal{V}^\varepsilon$ is uniformly bounded with respect to both $(t,\boldsymbol{x}) \in \mathbb{R}^+ \times K$ and ε. Furthermore, since N is continuous, we obtain

$$\mathbb{1}_{|\theta^\varepsilon v|>R}\,\theta^\varepsilon v\,\left(N(\mathcal{V}^\varepsilon+\theta^\varepsilon v)\,-\,N(\mathcal{V}^\varepsilon)\right) \,\leq\, \mathbb{1}_{|\theta^\varepsilon v|>R}\left(|\mathcal{V}^\varepsilon+\theta^\varepsilon v|^2\,\omega(\mathcal{V}^\varepsilon+\theta^\varepsilon v)\,\frac{\theta^\varepsilon v}{\mathcal{V}^\varepsilon+\theta^\varepsilon v}\,+\,C\,|\theta^\varepsilon v|\right),$$

where ω is given by assumption (1.1a) and where the constant C depends on both N and the uniform upper bound on $\mathcal{V}^\varepsilon$. Hence, since $\theta^\varepsilon \leq 1$, we obtain a radius R only depending on N and the uniform bound on $|\mathcal{V}^\varepsilon|$ such that

$$\mathcal{A}_{21} \,\leq\, \int_{\mathbb{R}^2}\left(\frac{m_*}{4}\,\mathbb{1}_{|\theta^\varepsilon v|>R}\,|\mathcal{V}^\varepsilon+\theta^\varepsilon v|^2\,\omega(\mathcal{V}^\varepsilon+\theta^\varepsilon v)\,+\,C\,|\theta^\varepsilon v|\right)|\nu|^2\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,,$$

where m_* is the lower bound of ρ_0^ε given by assumption (2.7). From now on, **we fix** R such that the latter estimate holds. Furthermore, we introduce the following notation

$$\mathcal{N} \,=\, \int_{\mathbb{R}^2}\mathbb{1}_{|\theta^\varepsilon v|>R}\,|\mathcal{V}^\varepsilon+\theta^\varepsilon v|^2\,\omega(\mathcal{V}^\varepsilon+\theta^\varepsilon v)\,|\nu|^2\,m^\varepsilon\,\mathrm{d}\boldsymbol{u} \,\leq\, 0 \quad \text{when} \quad R \gg 1\,.$$

The term $\mathcal{N}$ corresponds to the contribution of N on the domain where it has super-linear decaying properties and according to assumption (1.1a), it is non positive for R sufficiently large. Therefore, we use $\mathcal{N}$ to control the contribution of the other terms. With this notation, the estimate on $\mathcal{A}_{21}$ rewrites

$$\mathcal{A}_{21} \,\leq\, C\int_{\mathbb{R}^2}|\theta^\varepsilon v|\,|\nu|^2\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,+\,\frac{m_*}{4}\mathcal{N}\,,$$

where C and R only depend on N and the uniform bound on $|\mathcal{V}^\varepsilon|$.

We turn to $\mathcal{A}_{22}$. Since N has $\mathscr{C}^1$ regularity and relying item (2) in Proposition 2.5 , which ensures that $\mathcal{V}^\varepsilon$ stays uniformly bounded, we obtain

$$\mathcal{A}_{22} \,\leq\, C\,\frac{\rho_0^\varepsilon}{2}\int_{\mathbb{R}^2}|v|^2\,|\nu|^2\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,+\,C\|\nu\|^2_{L^2(m^\varepsilon)}\,,$$

where C is a positive constant which may depend on m_*, N and the uniform bound on $|\mathcal{V}^\varepsilon|$. We estimate the quadratic term in the latter inequality thanks to the following relation

$$\frac{1}{2}\int_{\mathbb{R}^2}\left(\rho_0^\varepsilon\,|v|^2\,-\,1\right)|\nu|^2\,m^\varepsilon\,\mathrm{d}\boldsymbol{u} \,=\, \int_{\mathbb{R}^2} v\,\nu\,\partial_v\left(\nu\,m^\varepsilon\right)\,\mathrm{d}\boldsymbol{u}\,.$$

It is obtained after exact computations and an integration by part in the right hand side of the latter equality. We apply Young's inequality to the former relation and in the end it yields

$$\mathcal{A}_{22} \,\leq\, \frac{\eta}{(\theta^\varepsilon)^2}\,\mathcal{D}_{\rho_0^\varepsilon}[\,\nu\,]\,+\,C\left(\frac{1}{\eta}\int_{\mathbb{R}^2}|\theta^\varepsilon v|^2\,|\nu|^2\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,+\,\|\nu\|^2_{L^2(m^\varepsilon)}\right),$$

for all positive η and for some positive constant C depending on m_*, N and the uniform bound on $|\mathcal{V}^\varepsilon|$.

We estimate the last term $\mathcal{A}_{23}$ taking advantage of the confining properties corresponding to $\mathcal{N}$. Indeed we have

$$\mathcal{A}_{23}\,+\,\frac{m_*}{8}\mathcal{N} \,=\, \int_{\mathbb{R}^2}\mathbb{1}_{|\theta^\varepsilon v|>R}\left(\frac{m_*}{8}\,\left|v'\right|^2\omega(v')-\frac{1}{2}N'\left(v'\right)\right)|\nu|^2\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,,$$

where we used the shorthand notation $v' \,=\, \mathcal{V}^\varepsilon+\theta^\varepsilon v$. Hence, according to assumption (2.17), we deduce

$$\mathcal{A}_{23}\,+\,\frac{m_*}{8}\mathcal{N} \,\leq\, C\|\nu\|^2_{L^2(m^\varepsilon)}\,,$$

for some positive constant C depending on m_* and N. Gathering these computations, we obtain

$$\mathcal{A}_2 \,\leq\, \frac{\eta}{(\theta^\varepsilon)^2}\,\mathcal{D}_{\rho_0^\varepsilon}[\,\nu\,]\,+\,C\left(\frac{1}{\eta}+1\right)\int_{\mathbb{R}^2}|\theta^\varepsilon v|^2\,|\nu|^2\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,+\,C\|\nu\|^2_{L^2(m^\varepsilon)}\,+\,\frac{m_*}{8}\mathcal{N}\,, \tag{4.2}$$

for all $\eta > 0$ and where C is a positive constant which may depend on m_*, R, N and the uniform bound on $|\mathcal{V}^\varepsilon|$.

We turn to $\mathcal{A}_3$, which gathers terms of lower-order. We integrate by part and apply Young's inequality. It yields

$$\mathcal{A}_3 \,\leq\, \frac{\eta}{(\theta^\varepsilon)^2}\,\mathcal{D}_{\rho_0^\varepsilon}[\,\nu\,] \,+\, \frac{1}{2\eta}\left(|\Psi *_r \rho_0^\varepsilon|^2 \int_{\mathbb{R}^2} |\theta^\varepsilon v|^2\,|\nu|^2\; m^\varepsilon\,\mathrm{d}\boldsymbol{u} \,+\, |\mathcal{E}(\mu^\varepsilon)|^2\;\|\nu\|^2_{L^2(m^\varepsilon)}\right),$$

for all positive η. According to item (4) in Proposition 2.5 , $\mathcal{E}(\mu^\varepsilon)$ is uniformly bounded with respect to both $(t,\boldsymbol{x}) \in \mathbb{R}^+ \times K$ and $\varepsilon > 0$. On top of that, according to assumptions (1.2) and (2.7) on Ψ and ρ_0^ε, $\Psi *_r \rho_0^\varepsilon$ is uniformly bounded with respect to both $\boldsymbol{x} \in K$ and $\varepsilon > 0$. Consequently, we obtain

$$\mathcal{A}_3 \,\leq\, \frac{\eta}{(\theta^\varepsilon)^2}\,\mathcal{D}_{\rho_0^\varepsilon}[\,\nu\,] \,+\, \frac{C}{\eta}\left(\int_{\mathbb{R}^2} |\theta^\varepsilon v|^2\,|\nu|^2\; m^\varepsilon\,\mathrm{d}\boldsymbol{u} \,+\, \|\nu\|^2_{L^2(m^\varepsilon)}\right), \tag{4.3}$$

for some positive constant C which may depend on m_* (see assumption (2.7)), m_p and $\overline{m}_p$ (see assumptions (2.8a) and (2.8b)) and the data of our problem N, Ψ and A_0.

Gathering estimates (4.1), (4.2) and (4.3), it yields

$$-\,\langle\,\mathrm{div}_{\boldsymbol{u}}[\,\mathbf{b}_0^\varepsilon\,\nu\,]\;,\;\nu\,\rangle \,\leq\, \frac{\alpha_- + 2\eta}{(\theta^\varepsilon)^2}\,\mathcal{D}_{\rho_0^\varepsilon}[\,\nu\,] \,+\, C\left(1+\frac{1}{\eta}\right)\int_{\mathbb{R}^2}\left(|\theta^\varepsilon v|^2 \,+\, 1\right)|\nu|^2\; m^\varepsilon\,\mathrm{d}\boldsymbol{u} \,+\, \frac{m_*}{8}\,\mathcal{N}\,,$$

for all positive η. Hence, we choose $2\eta = \alpha - \alpha_-$. Therefore, replacing $\mathcal{N}$ by its definition, the former estimate rewrites

$$-\,\langle\,\mathrm{div}_{\boldsymbol{u}}[\,\mathbf{b}_0^\varepsilon\,\nu\,]\;,\;\nu\,\rangle \,\leq\, \frac{\alpha}{(\theta^\varepsilon)^2}\,\mathcal{D}_{\rho_0^\varepsilon}[\,\nu\,] + \int_{\mathbb{R}^2}\left(C\left(|\theta^\varepsilon v|^2+1\right) + \mathbb{1}_{|\theta^\varepsilon v|>R}\,\frac{m_*}{8}\;\left|v'\right|^2\omega(v')\right)|\nu|^2\, m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,,$$

where we used the shorthand notation $v' = \mathcal{V}^\varepsilon + \theta^\varepsilon v$. To conclude, we estimate the right-hand side in the latter inequality applying assumption (1.1a) on N. Since $\mathcal{V}^\varepsilon$ is uniformly bounded, we obtain

$$-\,\langle\,\mathrm{div}_{\boldsymbol{u}}[\,\mathbf{b}_0^\varepsilon\,\nu\,]\;,\;\nu\,\rangle_{L^2(m^\varepsilon)} \,\leq\, \frac{\alpha}{(\theta^\varepsilon)^2}\,\mathcal{D}_{\rho_0^\varepsilon}[\,\nu\,] \,+\, C\|\nu\|^2_{L^2(m^\varepsilon)}\,,$$

for some constant C only depending on α, κ, m_*, m_p, $\overline{m}_p$ and the data of the problem: N, A_0 and Ψ. □

We also mention the following general result, which may be interpreted as a Poincaré inequality in the functional space $L^2(m^\varepsilon)$

Lemma 4.4. *For all $\boldsymbol{x} \in K$ and all function ν in $H^1_w(m^\varepsilon_{\boldsymbol{x}})$, holds the following estimates*

$$\|\,\nu\,\|_{L^2(m^\varepsilon)} \,\leq\, \frac{1}{\sqrt{\kappa}}\,\|\,\partial_w\,\nu\,\|_{L^2(m^\varepsilon)} \quad\text{and}\quad \|\,w\,\nu\,\|_{L^2(m^\varepsilon)} \,\leq\, \frac{2}{\kappa}\,\|\,\partial_w\,\nu\,\|_{L^2(m^\varepsilon)}\,.$$

Proof. The proof relies on the following relation

$$\frac{1}{2}\int_{\mathbb{R}^2}\left(1 \,+\, \kappa\,w^2\right)|\,\nu\,|^2\; m^\varepsilon_{\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} \,=\, -\int_{\mathbb{R}^2} w\,\nu\;(\partial_w\,\nu)\; m^\varepsilon_{\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,,$$

which is obtained by an integration by part in the left-hand side of the equality. From the latter relation we obtain the result applying Young's inequality in the right-hand side for the first estimate and Cauchy-Schwarz inequality for the second one.

□

From Lemma 4.2, we deduce regularity estimates for the solution ν^ε to equation (2.6). The main challenge consists in propagating the $\mathscr{H}^0$-norm. Then we easily adapt our analysis to the case of the $\mathscr{H}^k$-norms, when k is greater than 0. Indeed, the w-derivatives of ν^ε solve equation (2.6) with additional source terms which we are able to control with the dissipation brought by the Fokker-Planck operator. More precisely, equation (2.6) on ν^ε reads as follows

$$\partial_t\,\nu^\varepsilon \,+\, \mathscr{A}^\varepsilon[\,\nu^\varepsilon\,] \,=\, 0\,,$$

where the operator $\mathscr{A}^\varepsilon$ is given by

$$\mathscr{A}^\varepsilon[\,\nu^\varepsilon\,] \,=\, \mathrm{div}_{\boldsymbol{u}}[\,\mathbf{b}_0^\varepsilon\,\nu^\varepsilon\,] \,-\, \frac{1}{(\theta^\varepsilon)^2}\,\mathcal{F}_{\rho_0^\varepsilon}[\,\nu^\varepsilon\,]\,. \tag{4.4}$$

With this notation, the equations on the w-derivatives read as follows

$$\partial_t\,h^\varepsilon \,+\, \mathscr{A}^\varepsilon[\,h^\varepsilon\,] \,=\, \frac{1}{\theta^\varepsilon}\,\partial_v\,\nu^\varepsilon \,+\, b\,h^\varepsilon\,, \tag{4.5}$$

where $h^\varepsilon = \partial_w \nu^\varepsilon$, and

$$\partial_t g^\varepsilon + \mathscr{A}^\varepsilon [g^\varepsilon] = \frac{2}{\theta^\varepsilon} \partial_v h^\varepsilon + 2 b g^\varepsilon , \tag{4.6}$$

where g^ε is given by $\partial_w^2 \nu^\varepsilon$.

Proposition 4.5. *Under assumptions (1.1a)-(1.1b) and (2.17) on the drift N, (1.2) on the interaction kernel Ψ, consider a sequence of smooth solutions $(\mu^\varepsilon)_{\varepsilon>0}$ to (1.3) with initial conditions satisfying assumptions (2.7)-(2.8b) and (2.19) with an exponent κ greater than $1/(2b)$. Then, there exists a constant $C > 0$, such that, for all $\varepsilon > 0$, we have*

$$\left\| \partial_w^k \nu^\varepsilon(t\,,\,\boldsymbol{x}) \right\|_{L^2(m^\varepsilon_{\boldsymbol{x}})} \leq e^{Ct} \left\| \partial_w^k \nu_0^\varepsilon(\boldsymbol{x}) \right\|_{L^2(m^\varepsilon_{\boldsymbol{x}})} , \quad \forall\, (t, \boldsymbol{x}) \in \mathbb{R}^+ \times K ,$$

for all integer k less than or equal to 2

Proof. We start with $k = 0$. We compute the time derivative of $\| \nu^\varepsilon \|^2_{L^2(m^\varepsilon)}$ multiplying equation (2.6) by $\nu^\varepsilon\, m$ and integrating with respect to $\boldsymbol{u}$. After integrating by part the stiffer term, we obtain

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \| \nu^\varepsilon \|^2_{L^2(m^\varepsilon)} + \frac{1}{(\theta^\varepsilon)^2} \mathcal{D}_{\rho_0^\varepsilon} [\nu^\varepsilon] = - \langle \operatorname{div}_{\boldsymbol{u}} [\mathbf{b}_0^\varepsilon \nu^\varepsilon] , \nu^\varepsilon \rangle_{L^2(m^\varepsilon)} ,$$

for all $\varepsilon > 0$ and all $(t, \boldsymbol{x}) \in \mathbb{R}_+ \times K$. Since κ is greater than $1/(2b)$, we apply Lemma 4.2 with $\alpha = 1$. This leads to the following inequality

$$\frac{\mathrm{d}}{\mathrm{d}t} \| \nu^\varepsilon \|^2_{L^2(m^\varepsilon)} \leq C \| \nu^\varepsilon \|^2_{L^2(m^\varepsilon)} ,$$

for some constant C only depending on κ, m_*, m_p, $\overline{m}_p$ and on the data of the problem: N, A_0 and Ψ. According to Gronwall's lemma, it yields

$$\| \nu^\varepsilon(t\,,\,\boldsymbol{x}) \|_{L^2(m^\varepsilon_{\boldsymbol{x}})} \leq e^{Ct} \| \nu^\varepsilon(0\,,\,\boldsymbol{x}) \|_{L^2(m^\varepsilon_{\boldsymbol{x}})} , \quad \forall\, (t, \boldsymbol{x}) \in \mathbb{R}^+ \times K .$$

Let us now treat the case $k = 1$. We write $h^\varepsilon = \partial_w \nu^\varepsilon$. We compute the derivative of $\| h^\varepsilon \|^2_{L^2(m^\varepsilon)}$ multiplying equation (4.5) by $h^\varepsilon\, m^\varepsilon$ and integrating with respect to $\boldsymbol{u}$. After integrating by part the stiffer term, we obtain

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \| h^\varepsilon \|^2_{L^2(m^\varepsilon)} + \frac{1}{(\theta^\varepsilon)^2} \mathcal{D}_{\rho_0^\varepsilon} [h^\varepsilon] = - \langle \operatorname{div}_{\boldsymbol{u}} [\mathbf{b}_0^\varepsilon h^\varepsilon] , h^\varepsilon \rangle_{L^2(m^\varepsilon)} + b \| h^\varepsilon \|^2_{L^2(m^\varepsilon)} + \mathcal{B} ,$$

for all $\varepsilon > 0$ and all $(t, \boldsymbol{x}) \in \mathbb{R}_+ \times K$, where $\mathcal{B}$ is given by

$$\mathcal{B} = \frac{1}{\theta^\varepsilon} \int_{\mathbb{R}^2} \partial_v \nu^\varepsilon\, h^\varepsilon\, m^\varepsilon \, \mathrm{d}\boldsymbol{u} .$$

We estimate $\mathcal{B}$ integrating by part and applying Young's inequality. It yields

$$\mathcal{B} \leq \frac{C}{\eta} \| \nu^\varepsilon \|^2_{L^2(m^\varepsilon)} + \frac{\eta}{(\theta^\varepsilon)^2} \mathcal{D}_{\rho_0^\varepsilon} [h^\varepsilon] .$$

for some positive constant C and for all positive η. Then we apply Lemma 4.4, which yields

$$\mathcal{B} \leq \frac{C}{\eta} \| h^\varepsilon \|^2_{L^2(m^\varepsilon)} + \frac{\eta}{(\theta^\varepsilon)^2} \mathcal{D}_{\rho_0^\varepsilon} [h^\varepsilon] ,$$

and conclude this step following the same method as in the former step of the proof.

The last case $k = 2$ relies on the same arguments as the former step. Indeed, equation (4.6) on $\partial_w^2 \nu^\varepsilon$ is the same as equation (4.5) on $\partial_w \nu^\varepsilon$ up to a constant. Consequently, we skip the details and conclude this proof.

□

Due to the cross terms between the v and w variables in equation (1.3), we are led to estimate mixed quantities of the form $w^{k_1}\, \partial_w^{k_2} \nu^\varepsilon$. These estimates are easily obtained from Proposition 4.5 and Lemma 4.4

Corollary 4.6. *Under the assumptions of Proposition 4.5, we consider* $(k\,,\,k_1\,,\,k_2)$ *in* $\mathbb{N}^3$ *such that* $k_1 + k_2 = k$ *and* $k \leq 2$. *there exists a constant* $C > 0$, *such that, for all* $\varepsilon > 0$, *we have*

$$\left\| \left(w^{k_1}\, \partial_w^{k_2} \right) \nu^\varepsilon(t\,,\,\boldsymbol{x}) \right\|_{L^2(m^\varepsilon_{\boldsymbol{x}})} \leq \left(\frac{2}{\kappa} \right)^{k_1} e^{Ct} \left\| \partial_w^k\, \nu_0^\varepsilon(\boldsymbol{x}) \right\|_{L^2(m^\varepsilon_{\boldsymbol{x}})}, \quad \forall\, (t, \boldsymbol{x}) \in \mathbb{R}^+ \times K\,,$$

Proof. We consider $(k\,,\,k_1\,,\,k_2)$ in $\mathbb{N}^3$ such that $k_1 + k_2 = k$ and $k \leq 2$ and point out that according to Lemma 4.4, we have

$$\left\| \left(w^{k_1}\, \partial_w^{k_2} \right) \nu^\varepsilon(t\,,\,\boldsymbol{x}) \right\|_{L^2(m^\varepsilon_{\boldsymbol{x}})} \leq \left(\frac{2}{\kappa} \right)^{k_1} \left\| \partial_w^k\, \nu^\varepsilon(t\,,\,\boldsymbol{x}) \right\|_{L^2(m^\varepsilon_{\boldsymbol{x}})}.$$

Consequently, we obtain the result applying Proposition 4.5. □

We conclude this section with providing regularity estimates for the limiting distribution $\bar{\nu}$ with respect to the adaptation variable, which solves (1.7). The proof for this result is mainly computational since we have an explicit formula for the solutions to equation (1.7).

Lemma 4.7. *Consider some index* k *lying in* $\{0\,,\,1\}$ *and some* $\bar{\nu}_0$ *lying in* $\mathscr{H}^k(\overline{m})$. *The solution* $\bar{\nu}$ *to equation (1.7) with initial condition* $\bar{\nu}_0$ *verifies*

$$\| \bar{\nu}(t) \|_{\mathscr{H}^k(\overline{m})} \leq \exp\left(\left(k + \frac{1}{2} \right) b\, t \right) \| \bar{\nu}_0 \|_{\mathscr{H}^k(\overline{m})}, \quad \forall\, t \in \mathbb{R}^+.$$

Proof. Since $\bar{\nu}$ solves (1.7), it is given by the following formula

$$\bar{\nu}_{t,\,\boldsymbol{x}}(w) = e^{bt}\, \bar{\nu}_{0,\,\boldsymbol{x}}\left(e^{bt}\, w \right), \quad \forall\, (t, \boldsymbol{x}) \in \mathbb{R}^+ \times K\,.$$

Consequently, we easily obtain the expected result. □

4.2. **Proof of Theorem 2.3.** In the forthcoming analysis we quantify the convergence of ν^ε towards the asymptotic profile ν given by

$$\nu = \mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}\,,$$

in the functional spaces $\mathscr{H}^k(m^\varepsilon)$. We introduce the orthogonal projection of ν^ε onto the space of function with marginal $\mathcal{M}_{\rho_0^\varepsilon}$ with respect to the voltage variable

$$\Pi\, \nu^\varepsilon = \mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}^\varepsilon\,.$$

Furthermore, we consider the orthogonal component $\nu_\perp^\varepsilon$ of ν^ε with respect to the latter projection

$$\nu_\perp^\varepsilon = \nu^\varepsilon - \Pi\, \nu^\varepsilon\,.$$

With these notations we have

$$\| \nu^\varepsilon - \nu \|^2_{\mathscr{H}^k(m^\varepsilon)} = \| \nu_\perp^\varepsilon \|^2_{\mathscr{H}^k(m^\varepsilon)} + \| \bar{\nu}^\varepsilon - \bar{\nu} \|^2_{\mathscr{H}^k(\overline{m})}\,,$$

for k in $\{0\,,\,1\}$. Therefore, we prove that $\nu_\perp^\varepsilon$ and $\bar{\nu}^\varepsilon - \bar{\nu}$ vanish as ε goes to zero in both $\mathscr{H}^0$ and $\mathscr{H}^1$.

4.2.1. *Estimates for* $\nu_\perp^\varepsilon$. Our strategy relies on the same arguments as the ones we developed in the former section to prove Proposition 4.5. Indeed, the equation satisfied by $\nu_\perp^\varepsilon$ is the same equation as equation (2.6) solved by ν^ε with additional source terms. It reads as follows

$$\partial_t\, \nu_\perp^\varepsilon + \mathscr{A}^\varepsilon\left[\nu_\perp^\varepsilon \right] = \mathscr{S}\left[\nu^\varepsilon\,,\, \Pi\, \nu^\varepsilon \right], \tag{4.7}$$

where the operator $\mathscr{A}^\varepsilon$ is given by (4.4) and the source terms are given by

$$\mathscr{S}^\varepsilon\left[\nu^\varepsilon\,,\, \Pi\, \nu^\varepsilon \right] = \partial_w \left[a\, \theta^\varepsilon \int v\, \nu^\varepsilon\, \mathrm{d}v\, \mathcal{M}_{\rho_0^\varepsilon} - b\, w\, \Pi\, \nu^\varepsilon \right] - \mathrm{div}_{\boldsymbol{u}}\left[\mathbf{b}_0^\varepsilon\, \Pi\, \nu^\varepsilon \right].$$

Consequently, our strategy consists in estimating the source terms using the regularity estimates provided by Proposition 4.5. Then, we adapt our analysis to the case of $\partial_w\, \nu_\perp^\varepsilon$, which we write $h_\perp^\varepsilon = \partial_w\, \nu_\perp^\varepsilon$ and which solves again the same equation up to extra terms that add no difficulty

$$\partial_t\, h_\perp^\varepsilon + \mathscr{A}^\varepsilon\left[h_\perp^\varepsilon \right] = \mathscr{S}\left[h^\varepsilon\,,\, \Pi\, h^\varepsilon \right] + b\, h_\perp^\varepsilon + \frac{1}{\theta^\varepsilon}\, \partial_v\, \nu^\varepsilon\,, \tag{4.8}$$

where we used the notation $h^\varepsilon = \partial_w\, \nu^\varepsilon$.

Proposition 4.8. *Under assumptions (1.1a)-(1.1b) and (2.17) on the drift N and (1.2) on the interaction kernel Ψ, consider a sequence of solutions $(\mu^\varepsilon)_{\varepsilon>0}$ to (1.3) with initial conditions satisfying assumptions (2.7)-(2.8b) and (2.19) with an index k in $\{0\,,\,1\}$ and an exponent κ greater than $1/(2b)$. Then, for all ε between 0 and 1 and any α_* lying in $\left(0\,,\,1-(2b\kappa)^{-1}\right)$, there exists a constant $C>0$, independent of ε, such that*

$$\|\nu_\perp^\varepsilon(t)\|_{\mathscr{H}^k(m^\varepsilon)} \,\leq\, e^{Ct}\,\|\,\nu_0^\varepsilon\,\|_{\mathscr{H}^{k+1}(m^\varepsilon)}\left(C\,\sqrt{\varepsilon}\,+\,\min\left\{\,1\,,\,e^{-\,\alpha_*\,t\,/\,\varepsilon}\,\varepsilon^{-\,\alpha_*\,/\,(2\,m_*)}\right\}\right)\,,\quad\forall\,t\in\mathbb{R}^+\,.$$

Proof. We first treat the case $k=0$. All along this step of the proof, we consider some $\varepsilon\,>\,0$ and some $(t,\,\boldsymbol{x})$ in $\mathbb{R}_+\times K$; we omit the dependence with respect to $(t,\boldsymbol{x})$ when the context is clear. We compute the time derivative of $\|\,\nu_\perp^\varepsilon\,\|^2_{L^2(m^\varepsilon)}$ multiplying equation (4.7) by $\nu_\perp^\varepsilon\,m^\varepsilon$ and integrating with respect to $\boldsymbol{u}$

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\,\|\nu_\perp^\varepsilon\|^2_{L^2(m^\varepsilon)}\,=\,\langle\,\mathscr{S}^\varepsilon\,[\,\nu^\varepsilon\,,\,\Pi\,\nu^\varepsilon\,]\,,\,\nu_\perp^\varepsilon\,\rangle_{L^2(m^\varepsilon)}\,-\,\langle\,\mathscr{A}^\varepsilon\,[\,\nu_\perp^\varepsilon\,]\,,\,\nu_\perp^\varepsilon\,\rangle_{L^2(m^\varepsilon)}\,,$$

and we split the contribution of the source terms as follows

$$\langle\,\mathscr{S}^\varepsilon\,[\,\nu^\varepsilon\,,\,\Pi\,\nu^\varepsilon\,]\,,\,\nu_\perp^\varepsilon\,\rangle_{L^2(m^\varepsilon)}\,=\,\mathcal{A}_1\,+\,\mathcal{A}_2\,,$$

where the terms $\mathcal{A}_1$ and $\mathcal{A}_2$ are given by

$$\left\{\begin{array}{l}\displaystyle\mathcal{A}_1\,=\,-\,\frac{1}{\theta^\varepsilon}\int_{\mathbb{R}^2}\partial_v\,[\,B_0^\varepsilon\,(t\,,\,\boldsymbol{x}\,,\,\theta^\varepsilon\,v\,,\,w)\,\Pi\,\nu^\varepsilon\,]\,\nu_\perp^\varepsilon\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,,\\[2ex]\displaystyle\mathcal{A}_2\,=\,-a\,\theta^\varepsilon\int_{\mathbb{R}^2}v\,\partial_w\,[\,\Pi\,\nu^\varepsilon\,]\,\nu_\perp^\varepsilon\,m^\varepsilon\,\mathrm{d}\boldsymbol{u}\,,\end{array}\right.$$

where B_0^ε is given by (2.3).

Let us estimate $\mathcal{A}_1$. After an integration by part, this term rewrites as follows

$$\mathcal{A}_1\,=\,\int_{\mathbb{R}^2}B_0^\varepsilon\,(t\,,\,\boldsymbol{x}\,,\,\theta^\varepsilon\,v\,,\,w)\;\Pi\,\nu^\varepsilon\;(m^\varepsilon)^{\frac{1}{2}}\;\frac{1}{\theta^\varepsilon}\,\partial_v\,[\,\nu_\perp^\varepsilon\,m^\varepsilon\,]\;(m^\varepsilon)^{-\frac{1}{2}}\;\mathrm{d}\boldsymbol{u}\,.$$

According to items (2) and (4) in Proposition 2.5 , $\mathcal{V}^\varepsilon$ and $\mathcal{E}\,(\mu^\varepsilon)$ are uniformly bounded with respect to both $(t,\,\boldsymbol{x})\in\mathbb{R}^+\times K$ and ε. On top of that, according to assumptions (1.2) and (2.7), $\Psi\,*_r\,\rho_0^\varepsilon$ stays uniformly bounded with respect to both $\boldsymbol{x}\in K$ and ε as well. Consequently, applying Young's inequality to the former relation and using assumption (1.1b), which ensures that N has polynomial growth, we obtain

$$\mathcal{A}_1\,\leq\,\frac{\eta}{(\theta^\varepsilon)^2}\,\mathcal{D}_{\rho_0^\varepsilon}\,[\,\nu_\perp^\varepsilon\,]\,+\,\frac{C}{\eta}\int_{\mathbb{R}^2}\Big(\,|\theta^\varepsilon v|^{2p}\,+\,|w|^2\,+\,1\,\Big)\;|\bar\nu^\varepsilon|^2\;\mathcal{M}_{\rho_0^\varepsilon}\,\overline{m}\,\mathrm{d}\boldsymbol{u}\,,$$

for all positive η and for some positive constant C which only depends on m_*, m_p and $\overline{m}_p$ and the data of the problem N, Ψ and A_0. Taking advantage of the properties of the Maxwellian $\mathcal{M}_{\rho_0^\varepsilon}$ and since ρ_0^ε meets assumption (2.7), the latter estimate simplifies into

$$\mathcal{A}_1\,\leq\,\frac{\eta}{(\theta^\varepsilon)^2}\,\mathcal{D}_{\rho_0^\varepsilon}\,[\,\nu_\perp^\varepsilon\,]\,+\,\frac{C}{\eta}\left(\,\|\,\bar\nu^\varepsilon\,\|^2_{L^2(\overline{m})}\,+\,\|\,w\,\bar\nu^\varepsilon\,\|^2_{L^2(\overline{m})}\,\right).$$

Furthermore, according to Jensen's inequality, it holds

$$\|\,\bar\nu^\varepsilon\,\|^2_{L^2(\overline{m})}\,+\,\|\,w\,\bar\nu^\varepsilon\,\|^2_{L^2(\overline{m})}\,\leq\,\|\,\nu^\varepsilon\,\|^2_{L^2(m^\varepsilon)}\,+\,\|\,w\,\nu^\varepsilon\,\|^2_{L^2(m^\varepsilon)}\,.$$

Therefore, we apply Corollary 4.6 and obtain the following estimate for $\mathcal{A}_1$

$$\mathcal{A}_1\,\leq\,\frac{\eta}{(\theta^\varepsilon)^2}\,\mathcal{D}_{\rho_0^\varepsilon}\,[\,\nu_\perp^\varepsilon\,]\,+\,\frac{C}{\eta}\,e^{Ct}\,\|\,\nu_0^\varepsilon\,\|^2_{H^1_w(m^\varepsilon)}\,.$$

To estimate $\mathcal{A}_2$, we apply the Cauchy-Schwarz inequality, use the properties of the Maxwellian $\mathcal{M}_{\rho_0^\varepsilon}$ and assumption (2.7). It yields

$$\mathcal{A}_2\,\leq\,C\,\left(\,\|\,\partial_w\,\bar\nu^\varepsilon\,\|^2_{L^2(m^\varepsilon)}\,+\,\|\,\nu_\perp^\varepsilon\,\|^2_{L^2(m^\varepsilon)}\right).$$

According to the same remark as in the former step, it holds

$$\|\,\partial_w\,\bar\nu^\varepsilon\,\|^2_{L^2(\overline{m})}\,+\,\|\,\nu_\perp^\varepsilon\,\|^2_{L^2(m^\varepsilon)}\,\leq\,\|\,\nu^\varepsilon\,\|^2_{L^2(m^\varepsilon)}\,+\,\|\,\partial_w\,\nu^\varepsilon\,\|^2_{L^2(m^\varepsilon)}\,,$$

hence, applying Proposition 4.5, we obtain

$$\mathcal{A}_2 \,\leq\, C\, e^{Ct}\, \|\, \nu_0^\varepsilon \,\|^2_{H^1_w(m^\varepsilon)}\,.$$

To evaluate the contribution of $\mathscr{A}^\varepsilon$ we replace it by its definition (4.4) and integrate by part the stiffer term. It yields

$$-\,\langle\, \mathscr{A}^\varepsilon\,[\,\nu_\perp^\varepsilon\,]\,,\,\nu_\perp^\varepsilon\,\rangle_{L^2(m^\varepsilon)} \,=\, -\,\frac{1}{(\theta^\varepsilon)^2}\mathcal{D}_{\rho_0^\varepsilon}[\,\nu_\perp^\varepsilon\,] \,-\, \langle\, \mathrm{div}_{\boldsymbol{u}}\,[\,\mathbf{b}_0^\varepsilon\,\nu_\perp^\varepsilon\,]\,,\,\nu_\perp^\varepsilon\,\rangle_{L^2(m^\varepsilon)}\,.$$

In order to close the estimate, we apply Lemma 4.2 and Proposition 4.5 to control the term associated to linear transport. It yields

$$-\,\langle\, \mathrm{div}_{\boldsymbol{u}}\,[\,\mathbf{b}_0^\varepsilon\,\nu_\perp^\varepsilon\,]\,,\,\nu_\perp^\varepsilon\,\rangle_{L^2(m^\varepsilon)} \,\leq\, \frac{\alpha}{(\theta^\varepsilon)^2}\,\mathcal{D}_{\rho_0^\varepsilon}[\,\nu_\perp^\varepsilon\,] \,+\, C\, e^{Ct}\, \|\, \nu_0^\varepsilon \,\|^2_{L^2(m^\varepsilon)}\,,$$

for all positive constant α greater than $1/(2b\kappa)$. Consequently, the former computations lead to the following differential inequality

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\,\|\nu_\perp^\varepsilon\|^2_{L^2(m^\varepsilon)} \,+\, \frac{1-\alpha-\eta}{(\theta^\varepsilon)^2}\,\mathcal{D}_{\rho_0^\varepsilon}[\,\nu_\perp^\varepsilon\,] \,\leq\, \frac{C}{\eta}\,e^{Ct}\,\|\,\nu_0^\varepsilon\,\|^2_{H^1_w(m^\varepsilon)}\,.$$

Based on our assumptions, it holds $1/(2b\kappa) < 1$ therefore we may choose α and η such that $\alpha_* = 1-\alpha-\eta$ lies in $\left]\,0\,,\,1-(2b\kappa)^{-1}\,\right[$. Furthermore, in our context, the Gaussian-Poincaré inequality [13] rewrites

$$\|\nu_\perp^\varepsilon\|^2_{L^2(m^\varepsilon)} \,\leq\, \mathcal{D}_{\rho_0^\varepsilon}[\,\nu_\perp^\varepsilon\,]\,.$$

According to the latter remarks, the former inequality rewrites

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\|\nu_\perp^\varepsilon\|^2_{L^2(m^\varepsilon)} \,+\, \frac{2\,\alpha_*}{(\theta^\varepsilon)^2}\,\|\nu_\perp^\varepsilon\|^2_{L^2(m^\varepsilon)} \,\leq\, C\,e^{Ct}\,\|\,\nu_0^\varepsilon\,\|^2_{H^1_w(m^\varepsilon)}\,.$$

We multiply this estimate by

$$\exp\left(2\,\alpha_*\int_0^t \frac{1}{(\theta^\varepsilon)^2}\,\mathrm{d}s\right),$$

and integrate between 0 and t. In the end, we deduce the following inequality

$$\|\,\nu_\perp^\varepsilon\,\|^2_{L^2(m^\varepsilon)} \,\leq\, \|\,\nu_0^\varepsilon\,\|^2_{H^1_w(m^\varepsilon)}\, e^{-2\,\alpha_*\,I(t)}\left(1\,+\,C\int_0^t e^{Cs}\,e^{2\,\alpha_*\,I(s)}\,\mathrm{d}s\right),$$

where I is given by

$$I(t) \,=\, \int_0^t \frac{1}{(\theta^\varepsilon)^2}\,\mathrm{d}s\,.$$

Taking advantage of the ODE solved by θ^ε (see (2.5)), we compute explicitly I

$$I(t) \,=\, \frac{t}{\varepsilon} \,+\, \frac{1}{2\,\rho_0^\varepsilon}\ln\left(\,\theta^\varepsilon(t)^2\,\right).$$

Consequently, the latter estimate rewrites

$$\|\,\nu_\perp^\varepsilon\,\|^2_{L^2(m^\varepsilon)} \,\leq\, \|\,\nu_0^\varepsilon\,\|^2_{H^1_w(m^\varepsilon)}\, e^{-2\,\alpha_*\,\frac{t}{\varepsilon}}\left(\left(\theta^\varepsilon(t)\right)^{-2\,\frac{\alpha_*}{\rho_0^\varepsilon}} + C\int_0^t e^{Cs}\,e^{2\,\alpha_*\,\frac{s}{\varepsilon}}\,\left(\frac{\theta^\varepsilon(s)}{\theta^\varepsilon(t)}\right)^{2\,\frac{\alpha_*}{\rho_0^\varepsilon}}\,\mathrm{d}s\right).$$

Then we notice that according to the explicit formula (2.5) for θ^ε, given that ε lies in $(0\,,\,1)$, we have on the one hand

$$e^{-2\,\alpha_*\,\frac{t}{\varepsilon}}\,\left(\theta^\varepsilon(t)\right)^{-2\,\frac{\alpha_*}{\rho_0^\varepsilon}} \,\leq\, \min\left(1\,,\,e^{-2\,\alpha_*\,\frac{t}{\varepsilon}}\,\varepsilon^{-\,\frac{\alpha_*}{\rho_0^\varepsilon}}\right),$$

and on the other hand

$$\left(\frac{\theta^\varepsilon(s)}{\theta^\varepsilon(t)}\right)^{2\,\frac{\alpha_*}{\rho_0^\varepsilon}} \,\leq\, \min\left(2\,e^{2\,\alpha_*\,\frac{t-s}{\varepsilon}}\,,\,C\left(1\,+\,e^{-2\,\alpha_*\,\frac{s}{\varepsilon}}\,\varepsilon^{-\,\frac{\alpha_*}{\rho_0^\varepsilon}}\right)\right).$$

We inject these bounds and take the supremum over all $\boldsymbol{x}$ in K in the latter estimate. In the end, we obtain the estimate for the case where $k=0$ in Proposition 4.8.

We turn to the case $k = 1$ in Proposition 4.8. We make use of the shorthand notation $h^\varepsilon_\perp = \partial_w \nu^\varepsilon_\perp$. We compute the time derivative of $\| h^\varepsilon_\perp \|^2_{L^2(m^\varepsilon)}$ multiplying equation (4.8) by $h^\varepsilon_\perp m^\varepsilon$ and integrating with respect to $\boldsymbol{u}$

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\| h^\varepsilon_\perp \|^2_{L^2(m^\varepsilon)} = \langle \mathscr{S}^\varepsilon [h^\varepsilon , \Pi h^\varepsilon] , h^\varepsilon_\perp \rangle_{L^2(m^\varepsilon)} - \langle \mathscr{A}^\varepsilon [h^\varepsilon_\perp] , h^\varepsilon_\perp \rangle_{L^2(m^\varepsilon)} + b \| h^\varepsilon_\perp \|^2_{L^2(m^\varepsilon)} + \mathcal{A} ,$$

where $\mathcal{A}$ is given by

$$\mathcal{A} = \frac{1}{\theta^\varepsilon} \int_{\mathbb{R}^2} \partial_v \nu^\varepsilon h^\varepsilon_\perp m^\varepsilon \, \mathrm{d}\boldsymbol{u} .$$

We estimate $\mathcal{A}$ integrating by part with respect to v and applying Young's inequality. After applying Proposition 4.5, it yields

$$\mathcal{A} \leq \frac{1}{4\eta} e^{Ct} \| \nu^\varepsilon_0 \|^2_{L^2(m^\varepsilon)} + \frac{\eta}{(\theta^\varepsilon)^2} \mathcal{D}_{\rho^\varepsilon_0} [h^\varepsilon_\perp] ,$$

for some positive constant C and all positive η. Then we follow the same argument as in the last step and obtain the expected result.

□

4.2.2. *Estimate for* $(\bar\nu^\varepsilon - \bar\nu)$. It solves the following equation

$$\partial_t (\bar\nu^\varepsilon - \bar\nu) - b \partial_w [w (\bar\nu^\varepsilon - \bar\nu)] = -a \theta^\varepsilon \partial_w \left[\int_{\mathbb{R}} v \nu^\varepsilon \, \mathrm{d}v \right] , \tag{4.9}$$

obtained taking the difference between equations (1.7) and (2.14). It is the same equation as (1.7) solved by $\bar\nu$ with the additional source term on the right-hand side of the latter equation. Consequently, our strategy consists in estimating the source term. We point out that since the source term is weighted by θ^ε, it sufficient to prove that it is bounded in order to obtain convergence. However, it is not hard to check that the source term cancels if we replace ν^ε by its projection $\Pi \nu^\varepsilon$

$$\partial_w \left[\int_{\mathbb{R}} v \, \Pi \nu^\varepsilon \, \mathrm{d}v \right] = 0 .$$

Consequently, based on the estimates obtained in the first step on $\nu^\varepsilon_\perp$ (see Proposition 4.8), we expect

$$\theta^\varepsilon \partial_w \left[\int_{\mathbb{R}} v \nu^\varepsilon \, \mathrm{d}v \right] \underset{\varepsilon \to 0}{=} O(\varepsilon) .$$

This formal approach was already rigorously justified in a weak convergence setting in [2]. In our setting and due to the structure of the source term, we use regularity estimates to achieve the latter convergence rate.

Lemma 4.9. *Consider a sequence of solutions $(\mu^\varepsilon)_{\varepsilon > 0}$ to (1.3) with initial conditions satisfying assumption (2.19) with an index k in $\{0, 1\}$ as well as the solution $\bar\nu$ to equation (1.7) with some initial condition $\bar\nu_0$ lying in $\mathscr{H}^k(\overline{m})$. The following estimate holds for all positive ε*

$$\| \bar\nu^\varepsilon - \bar\nu \|_{H^k(\overline{m})} \leq e^{Ct} \left(\| \bar\nu^\varepsilon_0 - \bar\nu_0 \|_{H^k(\overline{m})} + C \int_0^t e^{-Cs} \theta^\varepsilon \| \nu^\varepsilon_\perp \|_{H^{k+1}_w(m^\varepsilon)} \, \mathrm{d}s \right) , \quad \forall (t, \boldsymbol{x}) \in \mathbb{R}^+ \times K ,$$

where k lies in $\{0, 1\}$ and for some positive constant C only depending on κ, m_* and A.

Proof. We start with the case $k = 0$. We consider some $\varepsilon > 0$ and some $(t, \boldsymbol{x})$ in $\mathbb{R}_+ \times K$; we omit the dependence with respect to $(t, \boldsymbol{x})$ when the context is clear. We compute the time derivative of $\| \bar\nu^\varepsilon - \bar\nu \|^2_{L^2(\overline{m})}$ multiplying equation (4.9) by $(\bar\nu^\varepsilon - \bar\nu) \overline{m}$ and integrating with respect to w. We integrate by part the term associated to linear transport and end up with the following relation

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\| \bar\nu^\varepsilon - \bar\nu \|^2_{L^2(\overline{m})} = \frac{b}{2} \int_{\mathbb{R}} (1 - \kappa w^2) | \bar\nu^\varepsilon - \bar\nu |^2 \, \overline{m} \, \mathrm{d}w - a \theta^\varepsilon \int_{\mathbb{R}^2} v \partial_w \nu^\varepsilon_\perp (\bar\nu^\varepsilon - \bar\nu) \, \overline{m} \, \mathrm{d}\boldsymbol{u} .$$

According to Cauchy-Schwarz inequality and applying assumption (2.7), the source term admits the bound

$$- a \theta^\varepsilon \int_{\mathbb{R}^2} v \partial_w \nu^\varepsilon_\perp (\bar\nu^\varepsilon - \bar\nu) \, \overline{m} \, \mathrm{d}\boldsymbol{u} \leq C \theta^\varepsilon \| h^\varepsilon_\perp \|_{L^2(m^\varepsilon)} \| \bar\nu^\varepsilon - \bar\nu \|_{L^2(\overline{m})} ,$$

for some positive constant C only depending on A and m_*. On top of that we bound the term associated to linear transport using that the polynomial $1-\kappa\, w^2$ is upper-bounded over $\mathbb{R}$. Gathering the former considerations we end up with the following differential inequality

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\,\|\,\bar\nu^\varepsilon\,-\,\bar\nu\,\|^2_{L^2(\overline m)}\,\le\,C\,\Big(\|\,\bar\nu^\varepsilon\,-\,\bar\nu\,\|^2_{L^2(\overline m)}\,+\,\theta^\varepsilon\,\|\,\partial_w\,\nu^\varepsilon_\perp\,\|_{L^2(m^\varepsilon)}\,\|\,\bar\nu^\varepsilon\,-\,\bar\nu\,\|_{L^2(\overline m)}\Big)\,,$$

for some positive constant C only depending on κ, m_* and A. we divide the latter inequality by $\|\,\bar\nu^\varepsilon\,-\,\bar\nu\,\|_{L^2(\overline m)}$ and obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\|\,\bar\nu^\varepsilon\,-\,\bar\nu\,\|_{L^2(\overline m)}\,\le\,C\,\Big(\|\,\bar\nu^\varepsilon\,-\,\bar\nu\,\|_{L^2(\overline m)}\,+\,\theta^\varepsilon\,\|\,\partial_w\,\nu^\varepsilon_\perp\,\|_{L^2(m^\varepsilon)}\Big)\,,$$

We conclude this step applying Gronwall's Lemma.

We treat the case $k\,=\,1$ applying the same method. Indeed, $\bar\nu^\varepsilon\,-\,\bar\nu$ and $\partial_w\,(\bar\nu^\varepsilon\,-\,\bar\nu)$ solve the same equation up to an additional source term which adds no difficulty.

□

Proposition 4.10. *Under assumptions (1.1a)-(1.1b) and (2.17) on the drift N and (1.2) on the interaction kernel Ψ, consider a sequence of solutions $(\mu^\varepsilon)_{\varepsilon>0}$ to (1.3) with initial conditions satisfying assumptions (2.7)-(2.8b) as well as the solution $\bar\nu$ to equation (1.7) with some initial condition $\bar\nu_0$ and an exponent κ greater than $1/(2b)$. Then there exists a constant $C>0$, such that for all $\varepsilon\in(0,1)$, the following statements hold*

(1) suppose that assumptions (2.19)-(2.20) are fulfilled with an index k in $\{0\,,\,1\}$, then for all $t\ge 0$,

$$\|\,\bar\nu^\varepsilon(t)\,-\,\bar\nu(t)\,\|_{\mathscr{H}^k(\overline m)}\,\le\,e^{Ct}\,\Big(\|\bar\nu^\varepsilon_0\,-\,\bar\nu_0\|_{\mathscr{H}^k(\overline m)}\,+\,C\,\|\,\nu^\varepsilon_0\,\|_{\mathscr{H}^{k+1}(m^\varepsilon)}\,\sqrt{\varepsilon}\Big)\;;$$

(2) supposing assumption (2.19) with index $k\,=\,1$ and assumption (2.20) with index $k\,=\,0$, it holds for all $t\ge 0$,

$$\|\,\bar\nu^\varepsilon(t)\,-\,\bar\nu(t)\,\|_{\mathscr{H}^0(\overline m)}\,\le\,e^{Ct}\,\Big(\|\bar\nu^\varepsilon_0\,-\,\bar\nu_0\|_{\mathscr{H}^0(\overline m)}\,+\,C\,\|\,\nu^\varepsilon_0\,\|_{\mathscr{H}^2(m^\varepsilon)}\,\varepsilon\,\sqrt{|\ln\varepsilon\,|\,+\,1}\Big)\,.$$

Proof. We prove item (2) in the latter proposition. According to Lemma 4.2, we have

$$\|\,\bar\nu^\varepsilon\,-\,\bar\nu\,\|_{L^2(\overline m)}\,\le\,e^{C\,t}\,\Big(\|\,\bar\nu^\varepsilon_0\,-\,\bar\nu_0\,\|_{L^2(\overline m)}\,+\,C\int_0^t e^{-C\,s}\,\theta^\varepsilon\,\|\,\nu^\varepsilon_\perp\|_{H^1_w(m^\varepsilon)}\,\mathrm{d}s\Big)\,.$$

Therefore, the proof comes down to estimating the integral in the right-hand side of the latter inequality

$$\mathcal{A}\,:=\,\int_0^t e^{-C\,s}\,\theta^\varepsilon\,\|\,\nu^\varepsilon_\perp\|_{H^1_w(m^\varepsilon)}\,\mathrm{d}s\,.$$

We apply the second estimate in Proposition 4.8 and Cauchy-Schwarz inequality. This yields

$$\mathcal{A}\,\le\,C\,\|\,\nu^\varepsilon_0\,\|_{H^2_w(m^\varepsilon)}\,\Big(\int_0^t|\theta^\varepsilon|^2\,\mathrm{d}s\Big)^{1/2}\,\Big(\int_0^t\varepsilon\,+\,\min\Big\{\,1\,,\,e^{-\,2\,\alpha_*\frac{s}{\varepsilon}}\,\varepsilon^{-\,\frac{\alpha_*}{m_*}}\Big\}\,\mathrm{d}s\Big)^{1/2}\,.$$

Then we inject the following estimate in the latter inequality

$$\min\Big\{\,1\,,\,e^{-\,2\,\alpha_*s\,/\,\varepsilon}\,\varepsilon^{-\,\alpha_*\,/\,m_*}\Big\}\,\le\,\mathbb{1}_{\left\{s\,\le\,-\frac{1}{2m_*}\varepsilon\ln\varepsilon\right\}}\,+\,\mathbb{1}_{\left\{s\,>\,-\frac{1}{2m_*}\varepsilon\ln\varepsilon\right\}}\,e^{-\,2\,\alpha_*\frac{s}{\varepsilon}}\,\varepsilon^{-\,\frac{\alpha_*}{m_*}}\,.$$

On top of that, we use (2.5) to compute the time integral of $|\theta^\varepsilon|^2$. In the end, we obtain

$$\mathcal{A}\,\le\,C\,\|\,\nu^\varepsilon_0\,\|_{H^2_w(m^\varepsilon)}\,\varepsilon\,\sqrt{t\,+\,1}\,\sqrt{|\ln\varepsilon\,|\,+\,1}\,.$$

Hence, we obtain the expected result taking the supremum over all $\boldsymbol{x}$ in K in the latter estimate.

Item (1) in Proposition 4.10 is obtained following the same method excepted that we estimate $\mathcal{A}$ using Proposition 4.5 with index k instead of Proposition 4.8.

□

Let us now conclude the proof of Theorem 2.3. On the one hand, we observe that item (2) corresponds to item (2) of Proposition 4.10. On the other hand, item (1) is obtained by gathering the estimate in Proposition 4.8 and item (1) in Proposition 4.10.

4.3. **Proof of Theorem 2.4.** All along this proof, we consider some ε_0 small enough so that the following condition is fulfilled

$$\| \rho_0^\varepsilon - \rho_0 \|_{L^\infty(K)} < m_* / 2 ,$$

for all ε less than ε_0. We omit the dependence with respect to $(t, \boldsymbol{x}, \boldsymbol{u}) \in \mathbb{R}^+ \times K \times \mathbb{R}^2$ when the context is clear. We start by proving item (1) in Theorem 2.4. Since the cases $k = 0$ and $k = 1$ are treated the same way, we only detail the case $k = 0$. We consider some integer i and take some ε less than ε_0. Then we decompose the error as follows

$$\| (v - \mathcal{V})^i \, (\mu^\varepsilon - \mu) \, (t) \|_{\mathcal{H}^0(m^-)} \leq \mathcal{A} + \mathcal{B} ,$$

where $\mathcal{A}$ and $\mathcal{B}$ are given by

$$\begin{cases} \mathcal{A} = \| (v - \mathcal{V})^i \, (\mu^\varepsilon - \tau_{-\mathcal{U}^\varepsilon} \circ D_{\theta^\varepsilon} (\nu)) \|_{\mathcal{H}^0(m^-)} , \\ \mathcal{B} = \| (v - \mathcal{V})^i \, (\tau_{-\mathcal{U}^\varepsilon} \circ D_{\theta^\varepsilon} (\nu) - \mu) \|_{\mathcal{H}^0(m^-)} , \end{cases}$$

where ν is the limit of ν^ε in Theorem 2.3 and is defined by (1.6) and where the operators $\tau_{-\mathcal{U}^\varepsilon}$ and D_{θ^ε} respectively stand for the translation of vector $-\mathcal{U}^\varepsilon$ with respect to the $\boldsymbol{u}$-variable and the dilatation with parameter $(\theta^\varepsilon)^{-1}$ with respect to the v-variable, that is

$$\tau_{-\mathcal{U}^\varepsilon} \circ D_{\theta^\varepsilon} (\nu) (t , \boldsymbol{x} , \boldsymbol{u}) = \frac{1}{\theta^\varepsilon} \nu \left(t , \boldsymbol{x} , \frac{v - \mathcal{V}^\varepsilon}{\theta^\varepsilon} , w - \mathcal{W}^\varepsilon \right) .$$

The first term $\mathcal{A}$ corresponds to the convergence of the re-scaled version ν^ε of μ^ε towards ν whereas $\mathcal{B}$ corresponds to the convergence of the macroscopic quantities.

We estimate $\mathcal{A}$ as follows

$$\mathcal{A} \leq C (\mathcal{A}_1 + \mathcal{A}_2) ,$$

where C is a positive constant which only depends on i and where $\mathcal{A}_1$ and $\mathcal{A}_2$ are given by

$$\begin{cases} \mathcal{A}_1 = \left\| (v - \mathcal{V}^\varepsilon)^i \, (\mu^\varepsilon - \tau_{-\mathcal{U}^\varepsilon} \circ D_{\theta^\varepsilon} (\nu)) \right\|_{\mathcal{H}^0(m^-)} , \\ \mathcal{A}_2 = \| \mathcal{V} - \mathcal{V}^\varepsilon \|^i_{L^\infty(K)} \, \| \mu^\varepsilon - \tau_{-\mathcal{U}^\varepsilon} \circ D_{\theta^\varepsilon} (\nu) \|_{\mathcal{H}^0(m^-)} . \end{cases}$$

According to item (2) in Proposition 2.5 , $\mathcal{V}^\varepsilon$ and $\mathcal{W}^\varepsilon$ are uniformly bounded with respect to both $(t, \boldsymbol{x}) \in \mathbb{R}^+ \times K$ and $\varepsilon > 0$, and $0 \leq \theta^\varepsilon \leq 1$, hence

$$m^-(\boldsymbol{x}, v, w) \leq C \, m^\varepsilon \left(\boldsymbol{x} , \frac{v - \mathcal{V}^\varepsilon}{\sqrt{2} \, \theta^\varepsilon} , w - \mathcal{W}^\varepsilon \right) ,$$

which yields

$$\mathcal{A}_1 \leq C \sup_{\boldsymbol{x} \in K} \left(\int_{\mathbb{R}^2} (v - \mathcal{V}^\varepsilon)^{2\,i} \, | \mu^\varepsilon - \tau_{-\mathcal{U}^\varepsilon} \circ D_{\theta^\varepsilon} (\nu) |^2 \, m^\varepsilon \left(\boldsymbol{x} , \frac{v - \mathcal{V}^\varepsilon}{\sqrt{2} \, \theta^\varepsilon} , w - \mathcal{W}^\varepsilon \right) \mathrm{d}\boldsymbol{u} \right)^{\frac{1}{2}} ,$$

for another constant $C > 0$ depending only on κ, m_*, m_p and $\overline{m}_p$ (see assumptions (2.7)-(2.8b)) and on the data of the problem N, Ψ and A_0. Then we invert the change of variable (2.1) and notice that

$$v^{2\,i} \, m^\varepsilon \left(\boldsymbol{x} , \frac{v}{\sqrt{2}} , w \right) \leq C \, m^\varepsilon (\boldsymbol{x} , \boldsymbol{u}) ,$$

for some constant $C > 0$ only depending on i and m_*. Consequently, we deduce

$$\mathcal{A}_1 \leq C \left\| (\theta^\varepsilon)^{i - \frac{1}{2}} \right\|_{L^\infty(K)} \| \nu^\varepsilon - \nu \|_{\mathcal{H}^0(m^\varepsilon)} .$$

Therefore, applying Theorem 2.3, using the compatibility assumption (2.22), and thanks to the constraint $\theta^\varepsilon (t = 0) = 1$, which ensures

$$\| \nu_0^\varepsilon \|_{\mathscr{H}^1(m^\varepsilon)} \leq C \, \| \mu_0^\varepsilon \|_{\mathscr{H}^1(m^+)} ,$$

for some constant C depending only on m_p, κ and m_*, we finally get

$$\mathcal{A}_1 \leq C \, e^{Ct} \, \varepsilon^{-\frac{1}{4}} \left(\varepsilon^{\frac{i}{2}} + e^{-i \, m_* \frac{t}{\varepsilon}} \right) \left(\varepsilon^{\frac{1}{2}} + e^{-\alpha_* \frac{t}{\varepsilon}} \, \varepsilon^{-\frac{\alpha_*}{2 \, m_*}} \right) .$$

Moreover, since $\alpha_* < m_*/2$, we deduce

$$\mathcal{A}_1 \leq C\, e^{Ct} \left(\varepsilon^{\frac{i}{2}+\frac{1}{4}} + e^{-\alpha_* \frac{t}{\varepsilon}}\, \varepsilon^{-\frac{1}{2}} \right).$$

To estimate $\mathcal{A}_2$, we apply item (1) in Proposition 2.5 and the compatibility assumption (2.22), which ensure

$$\| \mathcal{V} - \mathcal{V}^\varepsilon \|^i_{L^\infty(K)} \leq C\, e^{Ct}\, \varepsilon^i .$$

Then we follow the same method as before. In the end, we end up with the following bound for $\mathcal{A}_2$

$$\mathcal{A}_2 \leq C\, e^{Ct} \left(\varepsilon^{i+\frac{1}{4}} + e^{-\alpha_* \frac{t}{\varepsilon}}\, \varepsilon^{i-\frac{1}{2}} \right).$$

Gathering these results, we obtain the following estimate for $\mathcal{A}$

$$\mathcal{A} \leq C\, e^{Ct} \left(\varepsilon^{\frac{i}{2}+\frac{1}{4}} + e^{-\alpha_* \frac{t}{\varepsilon}}\, \varepsilon^{-\frac{1}{2}} \right).$$

We turn to $\mathcal{B}$. Similarly as before, we apply the triangular inequality and invert the change of variable (2.1). Then we apply Proposition 2.5, which yields

$$\mathcal{B} \leq C\, e^{Ct} \varepsilon^{-\frac{1}{4}} \left(\varepsilon^{\frac{i}{2}} + e^{-i\, m_* \frac{t}{\varepsilon}} \right) \left\| \nu - \tau_{\left(\frac{\mathcal{V}^\varepsilon - \mathcal{V}}{\theta^\varepsilon}, \mathcal{W}^\varepsilon - \mathcal{W}\right)} \left(\mathcal{M}_{\rho_0} \otimes \bar{\nu} \right) \right\|_{\mathcal{H}^0\left(D_{\sqrt{2}}\,(m^\varepsilon)\right)},$$

where $D_{\sqrt{2}}\,(m^\varepsilon)$ is a short-hand notation for

$$D_{\sqrt{2}}\,(m^\varepsilon)\,(\boldsymbol{x}\,,\boldsymbol{u}) = m^\varepsilon \left(\boldsymbol{x}\,, \frac{v}{\sqrt{2}}\,, w \right).$$

Then we decompose the right-hand side of the latter inequality as follows

$$\left\| \nu - \tau_{\left(\frac{\mathcal{V}^\varepsilon - \mathcal{V}}{\theta^\varepsilon}, \mathcal{W}^\varepsilon - \mathcal{W}\right)} \left(\mathcal{M}_{\rho_0} \otimes \bar{\nu} \right) \right\|_{\mathcal{H}^0\left(D_{\sqrt{2}}\,(m^\varepsilon)\right)} \leq \mathcal{B}_1 + \mathcal{B}_2 + \mathcal{B}_3\,,$$

where $\mathcal{B}_1$, $\mathcal{B}_2$ and $\mathcal{B}_3$ are given by

$$\begin{cases} \mathcal{B}_1 = \left\| \bar{\nu} - \tau_{(\mathcal{W}^\varepsilon - \mathcal{W})}\, \bar{\nu} \right\|_{\mathcal{H}^0(\overline{m})}, \\ \mathcal{B}_2 = \left\| \tau_{(\mathcal{W}^\varepsilon - \mathcal{W})}\, \bar{\nu} \right\|_{\mathcal{H}^0(\overline{m})} \left\| \mathcal{M}_{\rho_0^\varepsilon} - \tau_{\left(\frac{\mathcal{V}^\varepsilon - \mathcal{V}}{\theta^\varepsilon}\right)} \mathcal{M}_{\rho_0^\varepsilon} \right\|_{\mathcal{H}^0\left(\mathcal{M}^{-1}_{\rho_0^\varepsilon}\right)}, \\ \mathcal{B}_3 = \left\| \tau_{(\mathcal{W}^\varepsilon - \mathcal{W})}\, \bar{\nu} \right\|_{\mathcal{H}^0(\overline{m})} \left\| \mathcal{M}_{\rho_0^\varepsilon} - \mathcal{M}_{\rho_0} \right\|_{\mathcal{H}^0\left(\mathcal{M}^{-1}_{\rho_0^\varepsilon}\right)} e^{\|\mathcal{V}^\varepsilon - \mathcal{V}\|^2_{L^\infty(K)}/(2\, m_*\, \varepsilon)}, \end{cases}$$

where we used that

$$D_{\sqrt{2}}\,(m^\varepsilon) \leq m^\varepsilon\,, \quad \text{and} \quad m^\varepsilon = \mathcal{M}^{-1}_{\rho_0^\varepsilon}\, \overline{m}\,.$$

Since equation (1.7) is linear, $\bar{\nu} - \tau_{w_0}\, \bar{\nu}$ also solves the equation, therefore, applying Lemma 4.7 with $w_0 = \mathcal{W}^\varepsilon - \mathcal{W}$, it yields

$$\mathcal{B}_1 \leq C\, e^{\frac{b}{2}t} \left\| \bar{\nu}_0 - \tau_{(\mathcal{W}^\varepsilon - \mathcal{W})}\, \bar{\nu}_0 \right\|_{\mathcal{H}^0(\overline{m})}.$$

Furthermore, since $\overline{m} \leq \overline{m}^+$ and relying on assumption (2.23), we deduce

$$\mathcal{B}_1 \leq C\, e^{\frac{b}{2}t} \| \mathcal{W}^\varepsilon - \mathcal{W} \|_{L^\infty(K)}.$$

Therefore, according to item (1) in Proposition 2.5 we conclude

$$\mathcal{B}_1 \leq C\, e^{Ct}\, \varepsilon\,.$$

To estimate $\mathcal{B}_2$ and $\mathcal{B}_3$, we follow the same method as for $\mathcal{B}_1$: we first apply the following relation

$$\left\| \mathcal{M}_{\rho_0^\varepsilon} - \tau_{\left(\frac{\mathcal{V}^\varepsilon - \mathcal{V}}{\theta^\varepsilon}\right)} \mathcal{M}_{\rho_0^\varepsilon} \right\|_{\mathcal{H}^0\left(\mathcal{M}^{-1}_{\rho_0^\varepsilon}\right)} = \left\| e^{\rho_0^\varepsilon \left| \frac{\mathcal{V}^\varepsilon - \mathcal{V}}{\theta^\varepsilon} \right|^2} - 1 \right\|^{\frac{1}{2}}_{L^\infty(K)},$$

which ensures

$$\left\| \mathcal{M}_{\rho_0^\varepsilon} - \tau_{\left(\frac{\mathcal{V}^\varepsilon - \mathcal{V}}{\theta^\varepsilon}\right)} \mathcal{M}_{\rho_0^\varepsilon} \right\|_{\mathcal{H}^0\left(\mathcal{M}^{-1}_{\rho_0^\varepsilon}\right)} \leq \left\| e^{\rho_0^\varepsilon \left| \frac{\mathcal{V}^\varepsilon - \mathcal{V}}{\theta^\varepsilon} \right|^2} \rho_0^\varepsilon \left| \frac{\mathcal{V}^\varepsilon - \mathcal{V}}{\theta^\varepsilon} \right|^2 \right\|^{\frac{1}{2}}_{L^\infty(K)}.$$

Furthermore, we apply Lemma 4.7, which ensures that

$$\left\| \tau_{(\mathcal{W}^\varepsilon - \mathcal{W})}\, \bar{\nu} \right\|_{\mathcal{H}^0(\overline{m})} \leq C\, e^{Ct}\,.$$

Therefore, applying item (1) in Proposition 2.5, we obtain

$$\mathcal{B}_2 \leq C\, e^{C\left(t + \varepsilon\, e^{Ct}\right)}\, \varepsilon^{\frac{1}{2}}\,.$$

Then to estimate $\mathcal{B}_3$, an exact computation yields

$$\left\| \mathcal{M}_{\rho_0^\varepsilon} - \mathcal{M}_{\rho_0} \right\|_{\mathcal{H}^0\left(\mathcal{M}^{-1}_{\rho_0^\varepsilon}\right)} = \sup_{\boldsymbol{x}\in K} \left(\frac{|\,\rho_0 - \rho_0^\varepsilon\,|^2}{\sqrt{\rho_0^2 - (\rho_0 - \rho_0^\varepsilon)^2}\left(\rho_0 + \sqrt{\rho_0^2 - (\rho_0 - \rho_0^\varepsilon)^2}\right)} \right)^{\frac{1}{2}}\,.$$

Therefore, according to assumption (2.22) and item (1) in Proposition 2.5, which ensures that

$$e^{\|\mathcal{V}^\varepsilon - \mathcal{V}\|^2_{L^\infty(K)}\,/\,(2\,m_*\,\varepsilon)} \leq e^{C\,e^{Ct}\,\varepsilon}\,,$$

this yields

$$\mathcal{B}_3 \leq C\, e^{C\left(t + e^{Ct}\,\varepsilon\right)}\,\varepsilon\,.$$

In the end, we deduce the following estimate for $\mathcal{B}$

$$\mathcal{B} \leq C\, e^{C\left(t + \varepsilon\, e^{Ct}\right)} \left(\varepsilon^{\frac{i}{2} + \frac{1}{4}} + e^{-\,i\,m_*\,\frac{t}{\varepsilon}}\, \varepsilon^{\frac{1}{4}} \right)\,.$$

The proof for the statement (2) in Theorem 2.4 follows the same lines as the former one excepted that we apply item (2) in Theorem 2.3 instead of item (1) to quantify the convergence of $\bar{\nu}^\varepsilon$ towards $\bar{\nu}$. Therefore, we do not detail the proof.

5. Conclusion

In this paper, we have characterized the blow-up profile of the voltage distribution in the regime of strong and short-range coupling between neurons and have computed the limiting distribution for the adaptation variable. Our result should be interpreted as the expansion up to the first term of the network's distribution as ε vanishes. A significant aspect of our analysis lies in the fact that we obtained quantitative convergence estimates in strong topology. More specifically, we provide two results. On the one hand, we prove convergence in a L^1 functional framework. Our analysis relies on a modified Boltzmann entropy (see Appendix A) which is original to our knowledge. On the other hand, we prove convergence in a weighted L^2 setting, in which we take advantage of the variational structure in order to obtain regularity estimates. This enables to recover the optimal convergence rates obtained in a weak convergence setting in probability spaces [2]. In many ways, the problem at hands in this article shows similarities with other problems coming from kinetic theory. Therefore, the methods developed in this article might have applications in the asymptotic analysis of models sharing similar structure such as flocking models [18, 23, 24, 32], Vlasov-Navier-Stokes models with Brinkman force [15] and Vlasov-Poisson-Fokker-Planck models [29, 19].

Acknowledgment

The author gratefully acknowledges the support of ANITI (Artificial and Natural Intelligence Toulouse Institute). This project has received support from ANR ChaMaNe No: ANR-19-CE40-0024.

Appendix A. Proof of Lemma 3.1

Instead of estimating directly the L^1 norm of $f - g$, we estimate the following intermediate quantity, introduced in [2]

$$H_{1/2}\,[\,f\,|\,g\,] = \int_{\mathbb{R}^{d_1+d_2}} f\, \ln\left(\frac{2\,f}{f+g}\right) \mathrm{d}\boldsymbol{y}\,\mathrm{d}\xi\,,$$

which is easily comparable to the classical L^1 norm thanks to the following Lemma (see [2], Lemma A.1. for detailed proof)

Lemma A.1. *For any two non-negative functions $f,\, g \in L^1\left(\mathbb{R}^{d_1+d_2}\right)$ with integral equal to one, the following estimate holds*

$$\frac{1}{8}\,\|\,f\,-\,g\,\|^2_{L^1\left(\mathbb{R}^{d_1+d_2}\right)} \,\leq\, H_{1/2}\,[\,f\,|\,g\,] \,\leq\, \|\,f\,-\,g\,\|_{L^1(\mathbb{R}^{d_1+d_2})}\,.$$

Unlike the L^1 norm, $H_{1/2}$ has an explicit dissipation with respect to the Laplace operator which is given by

$$I_{1/2}\,[\,f\,|\,g\,] \,:=\, \int_{\mathbb{R}^{d_1+d_2}} \left|\nabla_\xi \ln\left(\frac{2\,f}{f+g}\right)\right|^2 f\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\xi\,.$$

We consider the quantity $h \,:=\, (f+g)\,/\,2$, whose equation is obtained multiplying by $1/2$ the sum of the equations solved by f and g, that is

$$\partial_t\,h \,+\, \mathrm{div}_{\boldsymbol{y}}\,[\,\mathbf{a}\,h\,] \,+\, \frac{\lambda(t)}{2}\,\mathrm{div}_\xi\,[\,(\mathbf{b}_1\,+\,\mathbf{b}_3)\;f\,+\,\mathbf{b}_2\,g\,] \,-\, \lambda(t)^2\,\Delta_\xi\,h \,=\, (\delta-1)\,\frac{\lambda(t)^2}{2}\,\Delta_\xi\,g\,.$$

Then, we compute the time derivative of the quantity $H_{1/2}\,[\,f\,|\,g\,]$ integrating with respect to both ξ and $\boldsymbol{y}$ the difference between the equation solved by f multiplied by $\ln\,(f\,/\,h)$ and the equation solved by h multiplied by $f\,/\,h$. After an integration by part, it yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\,H_{1/2}\,[\,f\,|\,g\,] \,+\, \lambda(t)^2\,I_{1/2}\,[\,f\,|\,g\,] \,=\, \mathcal{A}\,+\,\mathcal{B}\,,$$

where $\mathcal{A}$ and $\mathcal{B}$ are given by

$$\begin{cases} \mathcal{A} \,=\, \dfrac{\lambda(t)}{2}\displaystyle\int_{\mathbb{R}^{d_1+d_2}} (\,\mathbf{b}_1\,-\,\mathbf{b}_2\,)\,\nabla_\xi\left(\ln\left(\frac{f}{h}\right)\right)\,\frac{g}{h}\,f\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\xi\,, \\[2ex] \mathcal{B} \,=\, -\,\dfrac{\lambda(t)}{2}\displaystyle\int_{\mathbb{R}^{d_1+d_2}} \mathrm{div}_\xi\,[\,\mathbf{b}_3\,g\,+\,(\delta\,-\,1)\;\lambda(t)\,\nabla_\xi\,g\,]\;\frac{f}{h}\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\xi\,. \end{cases}$$

To estimate $\mathcal{A}$, we notice that $|\,g\,/\,h\,| \,\leq\, 2$ and we apply Young's inequality. This yields

$$\mathcal{A} \,\leq\, \lambda(t)^2\,I_{1/2}\,[\,f\,|\,g\,] \,+\, \frac{1}{4}\int_{\mathbb{R}^{d_1+d_2}} |\,\mathbf{b}_1\,-\,\mathbf{b}_2\,|^2\,f\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\xi\,.$$

To estimate $\mathcal{B}$, we simply notice that $|\,f\,/\,h\,| \,\leq\, 2$ and take the absolute value inside the integral. In the end, we obtain

$$\mathcal{A}\,+\,\mathcal{B} \,\leq\, \mathcal{R}(t) \,+\, \lambda(t)^2\,I_{1/2}\,[\,f\,|\,g\,]\,,$$

where $\mathcal{R}$ is defined as in Lemma 3.1. Therefore, it holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\,H_{1/2}\,[\,f(t)\,|\,g(t)\,] \,\leq\, \mathcal{R}(t)\,,\quad \forall\,t\in\mathbb{R}^+\,.$$

Then, we deduce (3.1) by integrating the latter inequality between 0 and t and applying Lemma A.1 in order to substitute $H_{1/2}$ with the L^1-norm.

References

[1] J. Baladron, D. Fasoli, O. Faugeras and J. Touboul — *Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons,* The Journal of Mathematical Neuroscience, **10**, (2012).

[2] A. Blaustein, F. Filbet — *Concentration phenomena in Fitzhugh-Nagumo's equations : a mesoscopic approach,* submitted.

[3] M. Bossy, O. Faugeras and D. Talay — *Clarification and complement to mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons,* The Journal of Mathematical Neuroscience, **5**, (2015).

[4] M.J. Caceres, J.A. Carrillo, and B. Perthame — *Analysis of nonlinear noisy integrate and fire neuron models: blow-up and steady states,* Journal of Mathematical Neuroscience, **1** (2011) pp. 33.

[5] M.J. Caceres, P. Roux, D. Salort and R. Schneider — *Global-in-time solutions and qualitative properties for the NNLIF neuron model with synaptic delay* Comm. Partial Differential Equations **44** (2019), pp. 1358-1386.

[6] J.A. Carrillo, X. Dou and Z. Zhou — *A simplified voltage-conductance kinetic model for interacting neurons and its asymptotic limit,* arXiv:2203.02746 [math.AP].

[7] J.A. Carrillo, B. Perthame, D. Salort, and D. Smets — *Qualitative properties of solutions for the noisy integrate and fire model in computational neuroscience,* Nonlinearity, **28**, (2015) pp. 3365-3388.

[8] J. Chevallier — *Mean-field limit of generalized hawkes processes*, Stochastic Processes and their Applications, **127**, (2017) pp. 3870-3912.

[9] J. Chevallier, M.J. Caceres, M. Doumic, and P. Reynaud-Bouret — *Microscopic approach of a time elapsed neural model,* Mathematical Models and Methods in Applied Sciences, **25**, (2015) pp. 2669-2719.

[10] J. Chevallier, A. Duarte, E. Löcherbach and G. Ost — *Mean-field limits for nonlinear spatially extended Hawkes processes with exponential memory kernels,* Stochastic Process. Appl. **129** (2019), pp. 1-27.
[11] J. Crevat, G. Faye and F. Filbet — *Rigorous derivation of the nonlocal reaction-diffusion FitzHugh-Nagumo system,* SIAM J. Math. Anal., **51**, (2019) pp. 346-373.
[12] J. Crevat — *Mean-field limit of a spatially-extended FitzHugh-Nagumo neural network,* Kinetic and Related Models, **12**, (2019), pp. 1329–1358.
[13] J. Dolbeault, B. Volzone — *Improved Poincaré inequalities*, Nonlinear Analysis: Theory, Methods and Applications, Elsevier, **75**, (2012), pp.5985 - 6001.
[14] N. El Ghani, N. Masmoudi — *Diffusion limit of the Vlasov-Poisson-Fokker-Planck system*, Commun. Math. Sci, **8**, (2010), no. 2, pp. 463 - 479.
[15] L. Ertzbischoff, D. Han-Kwan and A. Moussa,— *Concentration versus absorption for the Vlasov-Navier-Stokes system on bounded domains* Nonlinearity **34** (2021), pp. 68436900.
[16] M. Fathi, E. Indrei, M. Ledoux — *Quantitative logarithmic Sobolev inequalities and stability estimates,* Discrete Contin. Dyn. Syst, **36**, (2016), no. 12, 68356853.
[17] R. FitzHugh — *Impulses and physiological sates in theoretical models of nerve membrane,* Biophysical Journal, **1**, (1961), pp. 445466.
[18] A. Figalli and M. J. Kang — *A rigorous derivation from the kinetic Cucker-Smale model to the pressureless Euler system with nonlocal alignment* Anal. PDE **12** (2019), pp. 843866.
[19] M. Herda — *On massless electron limit for a multispecies kinetic system with external magnetic field,* J. Differential Equations, **260** (2016), no. 11, pp. 78617891.
[20] M. Herda and L. Rodrigues — *Large-time behavior of solutions to Vlasov-Poisson-Fokker-Planck equations: from evanescent collisions to diffusive limit,* J. Stat. Phys., **170** (2018), no. 5, 895931.
[21] A. L. Hodgkin and A. F. Huxley — *A quantitative description of membrane current and its application to conduction and excitation in nerve,* The Journal of Physiology, **117** (1952).
[22] P.-E. Jabin, D. Poyato and J. Soler — *Mean-field limit of non-exchangeable systems*, arXiv:2112.15406
[23] M.J. Kang and A. F. Vasseur — *Asymptotic analysis of Vlasov-type equations under strong local alignment regime* Math. Models Methods Appl. Sci. **25** (2015), 21532173.
[24] M.J. Kang and J. Kim — *Propagation of the mono-kinetic solution in the Cucker-Smale-type kinetic equations* Commun. Math. Sci. **18** (2020), no. 5, 12211231.
[25] E. Luon and W. Stannat — *Mean-field limit for disordered diffusions with singular interactions,* The annals of applied probability, **24**, (2014) pp. 1946-1993.
[26] N. Masmoui and M.L. Tayeb — *On the diffusion limit of a semiconductor Boltzmann-Poisson system without micro-reversible process,* Comm. Partial Differential Equations, **35**, (2010) , no. 7, pp. 11631175.
[27] S. Mischler, C. Quiinao and J. Touboul — *On a kinetic FitzHugh-Nagumo model of neuronal network,* Comm. Math. Phys., **342**, (2015) pp. 1001-1042.
[28] J. Nagumo, S. Arimoto and S. Yoshizawa — *An active pulse transmission line simulating nerve axon,* Proceedings of the IRE, **50**, (1962) pp. 2061-2070.
[29] J. Nieto, F. Poupaud and J. Soler— *High-field limit for the Vlasov-Poisson-Fokker-Planck system,* Arch. Ration. Mech. Anal. **158** (2001), pp. 29-59.
[30] K. Pakdaman, B. Perthame and D. Salort *Dynamics of a structured neuron population,* Nonlinearity **23**, (2010), pp. 55–75.
[31] B. Perthame and D. Salort — *On a voltage-conductance kinetic system for integrate and fire neural networks,* Kinet. Relat. Models **6** (2013)no. 4, pp 841–864.
[32] D. Poyato and J. Soler — *Euler-type equations and commutators in singular and hyperbolic limits of kinetic Cucker-Smale models.* Math. Models Methods Appl. Sci. **27** (2017), pp. 10891152.
[33] C. Quiinao and J. Touboul — *Clamping and synchronization in the strongly coupled FitzHugh-Nagumo model,* SIAM J. Appl. Dyn. Syst. **19** (2020), pp 788-827.