CONCENTRATION PHENOMENA IN FITZHUGH-NAGUMO'S EQUATIONS: A MESOSCOPIC APPROACH

Alain Blaustein, Francis Filbet

CONCENTRATION PHENOMENA IN FITZHUGH-NAGUMO'S EQUATIONS: A MESOSCOPIC APPROACH

ALAIN BLAUSTEIN AND FRANCIS FILBET

Alain Blaustein
Institut de Mathématiques de Toulouse, Université Paul Sabatier
Toulouse, France, alain.blaustein@math.univ-toulouse.fr

Francis Filbet
Institut de Mathématiques de Toulouse, Université Paul Sabatier
Toulouse, France, francis.filbet@math.univ-toulouse.fr

Abstract. We consider a spatially extended mesoscopic FitzHugh-Nagumo model with strong local interactions and prove that its asymptotic limit converges towards the classical nonlocal reaction-diffusion FitzHugh-Nagumo system. As the local interactions strongly dominate, the weak solution to the mesoscopic equation under consideration converges to the local equilibrium, which has the form of Dirac distribution concentrated to an averaged membrane potential.

Our approach is based on techniques widely developed in kinetic theory (Wasserstein distance, relative entropy method), where macroscopic quantities of the mesoscopic model are compared with the solution to the nonlocal reaction-diffusion system. This approach allows to make the rigorous link between microscopic and reaction-diffusion models.

Keywords: FitzHugh-Nagumo system; Wasserstein distance; blow-up profile.

2010MSC: 35K57; 35Q92; 35D30.

Contents

Date: January 6, 2022.

1. Introduction

1.1. **Physical model and motivations.** Neuron models often focus on the dynamics of the electrical potential through the membrane of a nerve cell. These dynamics are driven by ionic exchanges between the neuron and its environment through its cellular membranes. A very precise modeling of these ion exchanges led to the well-known Hodgkin-Huxley model [17]. In this paper, we shall focus on a simplified version, called the FitzHugh-Nagumo model [15] [20], which keeps its most valuable aspects and remains relatively simple mathematically. More precisely, the FitzHugh-Nagumo model accounts for the variations of the membrane potential v of a neuron coupled to an auxiliary variable w called the adaptation variable. It is usually written as follows:

$$\begin{cases} \mathrm{d}v_t \,=\, (N(v_t) \,-\, w_t \,+\, I_{ext})\,\mathrm{d}t \,+\, \sqrt{2}\,\mathrm{d}B_t\,, \\ \mathrm{d}w_t \,=\, A\,(v_t, w_t)\,\mathrm{d}t\,, \end{cases}$$

where the drift N is a confining non-linearity with the following typical form

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

even though a broader class of drifts N is considered here. On the other hand, the drift A is an affine mapping that has the following form

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

where $a,\, c \in \mathbb{R}$ and $b > 0$, which means that A also has some confining properties. Here, the Brownian motion B_t has been added in order to take into account random fluctuations in the dynamics of the membrane potential v_t. Another mathematical reason for looking at this system is that it is a prototypical model of excitable kinetics. Interest in such systems stems from the fact that although the kinetics are relatively simple, couplings between neurons can produce complex dynamics, where well-known examples are the propagation of excitatory pulses, spiral waves in two-dimensions, and spatio-temporal chaos. Here, we introduce coupling through the input current I_{ext}. More specifically, we consider that neurons interact with one another following Ohm's law and that the conductance between two neurons depends on there spatial location $\boldsymbol{x} \in K$, where K is a compact set of $\mathbb{R}^d$. The conductance between two neurons is given by a connectivity kernel $\Phi : K \times K \to \mathbb{R}$. Hence, in the case of a network composed with n interacting neurons described by the triplet voltage-adaptation-position $(v_i, w_i, \boldsymbol{x}_i)_{1\leq i\leq n}$, the current received by neuron i from the other neurons is given by

$$I_{ext} \,=\, -\frac{1}{n}\sum_{j=1}^{n} \Phi(\boldsymbol{x}_i, \boldsymbol{x}_j)\,(v_t^i - v_t^j)\,,$$

where the scaling parameter n is introduced here to re-normalize the contribution of each neuron. According to the former discussion, a neural network of size n is described by the system of equations

$$\begin{cases} \mathrm{d}v_t^i \,=\, \left(N(v_t^i) \,-\, w_t^i \,-\, \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} \Phi(\boldsymbol{x}_i, \boldsymbol{x}_j)\,(v_t^i - v_t^j)\right)\mathrm{d}t \,+\, \sqrt{2}\,\mathrm{d}B_t^i\,, \\ \mathrm{d}w_t^i \,=\, A\left(v_t^i, w_t^i\right)\mathrm{d}t\,, \end{cases}$$

where $i \in \{1, ..., n\}$. In the formal limit $n \to +\infty$, the behavior of the latter system may be described by the evolution of a distribution function $f := f(t, \boldsymbol{x}, \boldsymbol{u})$, with $\boldsymbol{u} = (v, w) \in \mathbb{R}^2$, representing the density of neurons at time t, position $\boldsymbol{x} \in K$ with a membrane potential v and adaptation variable $w \in \mathbb{R}$. It turns out that the distribution function f solves the following mean-field equation

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

where the operator $\mathcal{K}_\Phi[f]$ takes into account spatial interactions and is given by

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

See for instance [1, 5, 19, 18] for more details on the mean field limit for the FitzHugh-Nagumo system and [4] for a related model in collective dynamics.

Various other types of kinetic models have been derived during the past decades depending on the hypotheses assumed for the dynamics of the emission of an action potential. They include for example integrate-and-fire neural networks [7, 8, 9] and time-elapsed neuronal models [21, 11, 10, 12].

Let us be more specific on the modeling of interactions between neurons. A common assumption consists in considering that there of two types: strong short range interactions and weak long range interactions (see [6], [22] & [18]). Here we consider a connectivity kernel of the following type

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

where the Dirac mass δ_0 accounts for strong short range interactions with strength $\varepsilon>0$, whereas the connectivity kernel $\Psi : K\times K\to\mathbb{R}$ is more regular and represents weak long range interactions.

The purpose of this article is to go through the mathematical analysis of the neural network in the regime of strong interactions, that is when $\varepsilon\ll 1$. More precisely, we prove that the voltage distribution concentrates to a Dirac mass by providing a comprehensive description of this concentration phenomenon.

1.2. **Formal derivation.** Our problem is multiscale due to interactions between neurons, which induce macroscopic effects at the mesoscopic level. Consequently, we introduce integrated quantities. First, we consider the spatial distribution of neurons throughout the network

$$\rho_0^\varepsilon(\boldsymbol{x})=\int_{\mathbb{R}^2}f^\varepsilon(t,\,\boldsymbol{x},\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,.$$

It is straightforward to check that ρ_0^ε is indeed time-homogeneous, integrating the mean field equation with respect to $\boldsymbol{u}\in\mathbb{R}^2$. Second, we introduce the averaged voltage and adaptation variable at a spatial location $\boldsymbol{x}$

$$\left\{\begin{array}{lll}\rho_0^\varepsilon\,(\boldsymbol{x})\;\mathcal{V}^\varepsilon(t,\,\boldsymbol{x}) &=\displaystyle\int_{\mathbb{R}^2}v\;f^\varepsilon(t,\,\boldsymbol{x},\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,,\\[2ex]\rho_0^\varepsilon\,(\boldsymbol{x})\;\mathcal{W}^\varepsilon(t,\,\boldsymbol{x}) &=\displaystyle\int_{\mathbb{R}^2}w\;f^\varepsilon(t,\,\boldsymbol{x},\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,.\end{array}\right.\tag{1.1}$$

In the sequel, we use the vector notation $\mathcal{U}^\varepsilon\;=\;(\,\mathcal{V}^\varepsilon,\mathcal{W}^\varepsilon\,)$. At the mesoscopic level, we compare probability density functions using the Wasserstein distances. Hence, we renormalize f^ε as

$$\rho_0^\varepsilon\,\mu^\varepsilon\;=\;f^\varepsilon\,,$$

where μ^ε is a non-negative function which lies in $\mathscr{C}^0\left(\mathbb{R}^+\times K\,,\,L^1\left(\mathbb{R}^2\right)\right)$ and verifies

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

Consequently, we denote $\mu^\varepsilon_{t,\,\boldsymbol{x}}$ the probability density function defined as $\mu^\varepsilon_{t,\,\boldsymbol{x}}\;=\;\mu^\varepsilon\,(t,\,\boldsymbol{x},\cdot)$. With these notations and our modeling assumptions on the connectivity kernel Φ^ε, the mean field equation rewrites

$$\partial_t\,\mu^\varepsilon\;+\;\mathrm{div}_{\boldsymbol{u}}\,[\,\mathbf{b}^\varepsilon\mu^\varepsilon\,]\;-\;\partial_v^2\,\mu^\varepsilon\;=\;\frac{1}{\varepsilon}\,\rho_0^\varepsilon\,\partial_v\,[\,(v-\mathcal{V}^\varepsilon)\,\mu^\varepsilon\,]\,,\tag{1.2}$$

where $\mathbf{b}^\varepsilon$ is defined for all $(t,\, \boldsymbol{x}, \boldsymbol{u}) \in \mathbb{R}^+ \times K \times \mathbb{R}^2$ as

$$\mathbf{b}^\varepsilon(t,\, \boldsymbol{x}, \boldsymbol{u}) \,=\, \begin{pmatrix} N(v) \,-\, w \,-\, \mathcal{K}_\Psi[\rho_0^\varepsilon\, \mu^\varepsilon]\,(t,\, \boldsymbol{x}, v) \\ \\ A(\boldsymbol{u}) \end{pmatrix}.$$

Furthermore, one can notice that the non local term $\mathcal{K}_\Psi[\rho_0^\varepsilon\, \mu^\varepsilon]$ can be expressed in terms of the macroscopic quantities

$$\mathcal{K}_\Psi[\rho_0^\varepsilon\, \mu^\varepsilon](t,\, \boldsymbol{x}, v) \,=\, \Psi *_r \rho_0^\varepsilon(\boldsymbol{x})\, v \,-\, \Psi *_r (\rho_0^\varepsilon\, \mathcal{V}^\varepsilon)(t,\, \boldsymbol{x})\,,$$

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

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

Coming back to the analysis of the strong interaction regime, we look for the leading order in (1.2). In our case, it is induced by strong short range interactions between neurons, and as $\varepsilon \to 0$, we expect

$$(v \,-\, \mathcal{V}^\varepsilon)\, \mu^\varepsilon \,=\, 0\,,$$

which means that μ^ε concentrates around its mean value with respect to the voltage variable at each spatial location $\boldsymbol{x}$ in K, that is, μ^ε converges to a Dirac mass centred in $\mathcal{V}^\varepsilon$. Thus, to quantify the asymptotic behavior of μ^ε when $\varepsilon \ll 1$, we denote by $\mathscr{P}_2(\mathbb{R}^2)$ the set of probability laws with finite second order moments

$$\mathscr{P}_2\left(\mathbb{R}^2\right) \,=\, \left\{\mu \in \mathscr{P}(\mathbb{R}^2), \quad \int_{\mathbb{R}^2} |\boldsymbol{u}|^2 \mathrm{d}\mu(\boldsymbol{u}) \,<\, +\infty\right\}$$

and the Wasserstein distance of order two W_2 defined as follows: for any μ and ν probability measures in $\mathscr{P}_2(\mathbb{R}^2)$,

$$W_2^2\,(\mu, \nu) \,=\, \inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^4} |\boldsymbol{u} - \boldsymbol{u}'|^2 \mathrm{d}\pi(\boldsymbol{u}, \boldsymbol{u}')\,,$$

where $\Pi(\mu, \nu)$ stands for the set of distributions π over $\mathbb{R}^4$ with marginals μ with respect to $\boldsymbol{u}$ and ν with respect to $\boldsymbol{u}'$, that is, for $\pi \in \Pi(\mu, \nu)$

$$\pi\left(A,\, \mathbb{R}^2\right) = \mu(A) \quad \text{and} \quad \pi\left(\mathbb{R}^2,\, A\right) = \nu(A)\,,$$

for any Borel set $A \subset \mathbb{R}^2$. It is easy to quantify the distance between μ^ε and a Dirac mass in this framework since we have

$$W_2^2\left(\mu^\varepsilon_{t,\,\boldsymbol{x}},\, \delta_{\mathcal{V}^\varepsilon} \otimes \bar{\mu}^\varepsilon_{t,\,\boldsymbol{x}}\right) \,=\, \int_{\mathbb{R}^2} |v - \mathcal{V}^\varepsilon|^2\, \mu^\varepsilon_{t,\,\boldsymbol{x}}(\boldsymbol{u})\mathrm{d}\boldsymbol{u}\,,$$

where $\bar{\mu}^\varepsilon$ is defined as the marginal of μ^ε with respect to the voltage variable

$$\bar{\mu}^\varepsilon_{t,\,\boldsymbol{x}}(w) \,=\, \int_{\mathbb{R}} \mu^\varepsilon_{t,\,\boldsymbol{x}}(v, w)\, \mathrm{d}v\,.$$

Hence, we multiply equation (1.2) by $|v - \mathcal{V}^\varepsilon|^2$ and integrate with respect to $\boldsymbol{u} \in \mathbb{R}^2$. We obtain that at each spatial location $\boldsymbol{x}$ in K, we have

$$W_2\left(\mu^\varepsilon_{t,\,\boldsymbol{x}},\, \delta_{\mathcal{V}^\varepsilon(t,\,\boldsymbol{x})} \otimes \bar{\mu}^\varepsilon_{t,\,\boldsymbol{x}}\right) \underset{\varepsilon\to 0}{\sim} \sqrt{\varepsilon}\,.$$

Considering this estimate, we infer that the dynamics of the network when $\varepsilon \ll 1$ are driven by the couple $(\mathcal{V}^\varepsilon, \bar{\mu}^\varepsilon)$, which displays both the macroscopic & the mesoscopic scale. We complete this step of our analysis by deriving the limit of $(\mathcal{V}^\varepsilon, \bar{\mu}^\varepsilon)$. Multiplying equation (1.2) by v (resp. 1) and then integrating over $\boldsymbol{u} \in \mathbb{R}^2$ (resp. $v \in \mathbb{R}$), we obtain that the couple $(\mathcal{V}^\varepsilon, \bar{\mu}^\varepsilon)$ solves the following system

$$\left\{\begin{array}{l} \partial_t \mathcal{V}^\varepsilon \,=\, N(\mathcal{V}^\varepsilon) \,-\, \mathcal{W}^\varepsilon \,-\, \mathcal{L}_{\rho_0^\varepsilon}[\mathcal{V}^\varepsilon] \,+\, \mathcal{E}(\mu^\varepsilon)\,, \\ \\ \partial_t \bar{\mu}^\varepsilon \,+\, \partial_w \left(a \displaystyle\int_{\mathbb{R}} v \mu^\varepsilon \mathrm{d}v \,-\, b\, w\, \bar{\mu}^\varepsilon \,+\, c\, \bar{\mu}^\varepsilon\right) = 0\,, \end{array}\right. \tag{1.3}$$

with

$$\mathcal{W}^\varepsilon \;=\; \int_{\mathbb{R}} w\,\bar{\mu}^\varepsilon\,\mathrm{d}w\,.$$

In equation (1.3), $\mathcal{L}_{\rho_0^\varepsilon}[\mathcal{V}^\varepsilon]$ is a non local operator given by

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

and the error term $\mathcal{E}(\mu^\varepsilon)$ is given by

$$\mathcal{E}(\mu^\varepsilon) \;=\; \int_{\mathbb{R}^2} N(v)\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} \;-\; N(\mathcal{V}^\varepsilon)\,. \tag{1.4}$$

Before computing the limit for $(\mathcal{V}^\varepsilon, \bar{\mu}^\varepsilon)$, we emphasize that multiplying the second equation in (1.3) by w and integrating with respect to $w \in \mathbb{R}$, we get a closed equation for $\mathcal{W}^\varepsilon$ since A is affine,

$$\partial_t \mathcal{W}^\varepsilon \;=\; A\,(\mathcal{V}^\varepsilon, \mathcal{W}^\varepsilon)\,.$$

Both equations on $\mathcal{V}^\varepsilon$ and $\bar{\mu}^\varepsilon$ in (1.3) depend on the distribution function μ^ε. However since our interest here lies in the regime of strong interactions, we replace μ^ε in (1.3) by the *ansatz*

$$\mu^\varepsilon \underset{\varepsilon\to 0}{=} \delta_{\mathcal{V}^\varepsilon} \otimes \bar{\mu}^\varepsilon \;+\; O(\sqrt{\varepsilon})\,.$$

This removes the dependence with respect to μ^ε from the system (1.3). Indeed, we obtain on the one hand (see Proposition 3.5 for more details)

$$(\mathcal{V}^\varepsilon,\, \mathcal{W}^\varepsilon) \underset{\varepsilon\to 0}{=} (\mathcal{V},\, \mathcal{W}) \;+\; O\,(\varepsilon)\,, \quad \text{as} \quad \varepsilon \to 0\,,$$

and on the other hand

$$\bar{\mu}^\varepsilon \underset{\varepsilon\to 0}{=} \bar{\mu} \;+\; O\left(\sqrt{\varepsilon}\right), \quad \text{as} \quad \varepsilon \to 0\,,$$

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

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

with

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

Similar results have already been obtained in a deterministic setting in [13] using relative entropy methods. In the end, it can be proven that μ^ε converges to a mono-kinetic distribution in v with mean $\mathcal{V}$ and we get the following rate of convergence

$$W_2\left(\mu^\varepsilon_{t,\boldsymbol{x}},\; \delta_{\mathcal{V}(t,\boldsymbol{x})} \otimes \bar{\mu}_{t,\boldsymbol{x}}\right) \underset{\varepsilon\to 0}{=} O(\sqrt{\varepsilon})\,.$$

Actually, this latter convergence estimate corresponds to the expansion at order 0 of f^ε in the regime of strong interactions.

In this paper, we investigate the asymptotic profile of the solution f^ε in the regime of strong interactions in order to improve the order of convergence. The strategy consists in finding the concentration's profile with respect to the potential variable v. This leads to considering the following re-scaled version ν^ε of μ^ε

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

where ε^α is the concentration rate of μ^ε around its mean value $\mathcal{V}^\varepsilon$ and α needs to be determined. For a proper choice of α, we expect ν^ε to converge to some limit as ε vanishes. This limit will be interpreted as the concentration profile of the voltage's distribution throughout the network in the regime of strong interaction.

In order to determine the proper concentration exponent α, we derive the equation solved by ν^ε. To this aim, we perform the following change of variable

$$(v,\, w) \mapsto \left(\frac{v - \mathcal{V}^\varepsilon}{\varepsilon^\alpha},\, w - \mathcal{W}^\varepsilon \right) \tag{1.6}$$

in equation (1.2) and use the first equation of (1.3) on $\mathcal{V}^\varepsilon$. It yields

$$\partial_t \nu^\varepsilon \,+\, \mathrm{div}_{\boldsymbol{u}} \left[\mathbf{b}_0^\varepsilon \, \nu^\varepsilon \right] \,=\, \frac{1}{\varepsilon^{2\alpha}} \partial_v \left[\varepsilon^{2\alpha - 1} \, \rho_0^\varepsilon \, v \, \nu^\varepsilon + \partial_v \, \nu^\varepsilon \right] ,$$

where $\mathbf{b}_0^\varepsilon$ is a centered version of $\mathbf{b}^\varepsilon$ and is given by

$$\mathbf{b}_0^\varepsilon \,(t,\, \boldsymbol{x},\, \boldsymbol{u}) \,=\, \begin{pmatrix} \varepsilon^{-\alpha} \left(N(\mathcal{V}^\varepsilon \,+\, \varepsilon^\alpha v) \,-\, N(\mathcal{V}^\varepsilon) \,-\, w \,-\, \varepsilon^\alpha v \, \Psi *_r \rho_0^\varepsilon(\boldsymbol{x}) \,-\, \mathcal{E}\,(\mu^\varepsilon) \right) \\ A_0(\varepsilon^\alpha v,\, w) \end{pmatrix} ,$$

and where A_0 is the linear version of A

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

It turns out that the only suitable value for α is $1/2$. Indeed, when α is lesser than $1/2$, we check that ν^ε converges towards a Dirac mass, which means that the scaling is not precise enough. On the contrary, when $\alpha \,>\, 1/2$, ν^ε converges to 0, which means that we "zoom in" too much. Hence we obtain the following equation

$$\partial_t \, \nu^\varepsilon \,+\, \mathrm{div}_{\boldsymbol{u}} \left[\, \mathbf{b}_0^\varepsilon \, \nu^\varepsilon \, \right] \,=\, \frac{1}{\varepsilon} \, \partial_v \left[\, \rho_0^\varepsilon \, v \, \nu^\varepsilon + \partial_v \, \nu^\varepsilon \, \right] , \tag{1.7}$$

where we take $\alpha = 1/2$ in the definition of $\mathbf{b}_0^\varepsilon$. Keeping only the leading order, it yields that

$$\nu^\varepsilon_{t,\,\boldsymbol{x}}(v, w) \underset{\varepsilon \to 0}{=} \mathcal{M}_{\rho_0^\varepsilon(\boldsymbol{x})}(v) \otimes \bar{\nu}^\varepsilon_{t,\,\boldsymbol{x}}(w) \,+\, O(\sqrt{\varepsilon}) \,,$$

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

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

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

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

At this point, it possible to answer our initial concern: μ^ε concentrates with Gaussian profile as $\varepsilon \to 0$. Then we complete the analysis by deriving the limit of $\bar{\nu}^\varepsilon$: integrating equation (1.7) with respect to the re-scaled voltage variable v, we obtain the equations solved by $\bar{\nu}^\varepsilon$,

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

Once again, the equation still depends on ν^ε through the source term in the right-hand side. However we obtain that it is in fact of order ε when we replace ν^ε with the following *ansatz*

$$\nu^\varepsilon \underset{\varepsilon \to 0}{=} \mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}^\varepsilon \,+\, O(\sqrt{\varepsilon}) \,.$$

Consequently, we expect the following convergence

$$\bar{\nu}^\varepsilon \underset{\varepsilon \to 0}{=} \bar{\nu} \,+ O(\varepsilon) \,,$$

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

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

which corresponds to the same equation as (1.5) for $\bar{\mu}$ after inverting the change of variable (1.6). We come back to our initial problem, which consisted in building a precise model for the dynamics

of μ^ε in the regime of strong interactions. We invert the change of variable (1.6) and obtain in the end

$$W_2\left(\mu^\varepsilon_{t,\boldsymbol{x}},\, \mathcal{M}_{\frac{1}{\varepsilon}\rho_0(\boldsymbol{x}),\mathcal{V}(t,\boldsymbol{x})} \otimes \bar{\mu}_{t,\boldsymbol{x}}\right) \underset{\varepsilon\to 0}{=} O(\varepsilon)\,,$$

where $\mathcal{M}_{\rho_0/\varepsilon,\mathcal{V}}$ is given by $\mathcal{M}_{\rho_0/\varepsilon,\mathcal{V}}(v) = \mathcal{M}_{\rho_0/\varepsilon}(v-\mathcal{V})$. This result should be regarded as the expansion of μ^ε at order 1 in the regime of strong interactions. It may be compared with the expansion of μ^ε at order 0. Furthermore, it enables us to characterize the blow up profile of the distribution function μ^ε and to improve the order of convergence.
Our result is in line with a broader collection of publications, which focus on the mathematical analysis of the dynamics in a FitzHug-Nagumo neural networks with strongly interacting neurons. First, we mention [22], in which similar results are obtained following a Hamilton Jacobi approach. The authors study the so-called Hopf-Cole transform ϕ^ε of μ^ε defined by the following *ansatz*

$$\mu^\varepsilon = \exp\left(\phi^\varepsilon/\varepsilon\right).$$

However, due to this *ansatz*, authors deal with well prepared initial condition in the sense that it is already concentrated at time $t=0$. In our case, we lift this assumption and deal with non-concentrated initial data. Furthermore, the results are stated in a spatially homogeneous setting and the limiting distribution $\bar{\mu}$ of the adaptation variable is not identified in [22]. Secondly, our work follows on from [13], which focuses on the expansion of μ^ε at order 0 in a deterministic setting. On top of that, we cite [4] which deals with mean-field limit in the context of collective dynamics. This article locates itself in a probability framework and the authors develop mathematical methods based on the Wasserstein distance and similar to the ones in the present article. However, the authors adopt a stochastic point of view whereas we focus on the analytic point of view all along this paper. To end with, we mention [16], where methods related to Wasserstein distances are "informally reviewed" for a broad class of models.

The rest of the paper consists in making the asymptotic expansion rigorous. In the next section, we state our assumptions on the parameters of our problem: N, Ψ and $(\mu^\varepsilon_0)_{\varepsilon>0}$. Then we state the main result, Theorem 2.7. Section 3 is devoted to *a priori* estimates on the solutions $(\mu^\varepsilon)_{\varepsilon>0}$, whereas Section 4 contains the proof of Theorem 2.7. Finally, in the Appendix, we give the main ingredients to prove existence and uniqueness of a solution μ^ε to equation (1.2) for any $\varepsilon>0$. We mention that even though this well posedness result is not our main concern here, we develop interesting arguments using a modified relative entropy which might have other applications.

2. Mathematical setting & main results

In this section, we give the precise mathematical setting of our analysis and present our main results on the profile of the distribution function f^ε when $\varepsilon \ll 1$.

2.1. Mathematical setting.

We suppose the drift N to be of class $\mathscr{C}^2$ over $\mathbb{R}$. Then we set $\omega(v) = N(v)/v$ and suppose that the following coupled pair of confining assumptions are met

$$\limsup_{|v|\to+\infty} \omega(v) = -\infty\,, \tag{2.1a}$$

$$\sup_{|v|\geq 1} \frac{|\omega(v)|}{|v|^{p-1}} < +\infty\,, \tag{2.1b}$$

for some $p \geq 2$. Assumption (2.1a) ensures that N has super-linear decay at infinity. It allows us to obtain uniform estimates in time. However, it would have been replaced by

$$\sup_{|v|\geq 1} \omega(v) < +\infty\,,$$

had we worked on finite time intervals. Assumption (2.1b) ensures that the decaying rate of N is controlled by a polynomial. This assumption is merely technical. Indeed, it may be possible to consider exponential control instead, as long as we also suppose exponential moments for the initial condition μ^ε_0. In the end, the assumption only dictates the localization we need on the initial

condition: in our case, we choose polynomial control on the drift N and polynomial moments on the initial condition. A typical choice for N would be any polynomial of the form

$$P(v) \,=\, Q(v) \,-\, Cv|v|^{p-1}\,,$$

for some positive constant $C \,>\, 0$ and where Q has degree less then p.

We turn to the connectivity kernel Ψ. We suppose $\Psi \,\in\, \mathscr{C}^0\left(K_{\boldsymbol{x}},\, L^1\left(K_{\boldsymbol{x}'}\right)\right)$ and assume the following bound to hold

$$\sup_{\boldsymbol{x}'\in K}\int_K |\Psi(\boldsymbol{x},\boldsymbol{x}')|\,\mathrm{d}\boldsymbol{x} \,<\, +\infty\,. \tag{2.2}$$

Moreover, we consider r in $]1,\,+\infty]$, define its conjugate $r' \geq 1$ as $1/r + 1/r' = 1$ and suppose

$$\sup_{\boldsymbol{x}\in K}\int_K |\Psi(\boldsymbol{x},\boldsymbol{x}')|^r\,\mathrm{d}\boldsymbol{x}' \,<\, +\infty\,. \tag{2.3}$$

Our set of assumptions on the connectivity kernel is quite general since we consider non-symmetric interactions between neurons and also authorize the connectivity kernel to follow a power law, a case which is considered in the physical literature (see [18]).

Remark 2.1. *It may be possible to adapt our analysis to the case of a discrete spatial variable. This could be done by replacing the Lebesgue measure* $\mathrm{d}\boldsymbol{x}$ *in the definition of* Ψ *by a positive Borel measure* λ_K *with finite mass over* K *(typically a sum of Dirac mass if* K *is discrete). In this case, we should also suppose that* ρ_0^ε *lies in* $L^1\left(\lambda_K\right)$.

We now state our assumptions on the sequence of initial data $\left(\mu^\varepsilon_{0,\boldsymbol{x}}\right)_{\varepsilon\,>\,0}$. We suppose that for each $\varepsilon \,>\, 0$ we have

$$\left(\boldsymbol{x} \,\mapsto\, \mu^\varepsilon_{0,\boldsymbol{x}}\right) \in \mathscr{C}^0\left(K\,,\,\mathscr{P}^2\left(\mathbb{R}^2\right)\right)\,.$$

In particular, the spatial distribution of the network ρ_0^ε is continuous. We also suppose that it is uniformly bounded from above and below, that is, there exists a constant $m_* \,>\, 0$ such that for all $\varepsilon > 0$

$$m_* \leq \rho_0^\varepsilon \leq 1/m_*\,. \tag{2.4}$$

On top of that, we assume the following condition: there exists two positive constants m_p and $\overline{m}_p$, independent of ε, such that

$$\sup_{\boldsymbol{x}\in K}\int_{\mathbb{R}^2} |\boldsymbol{u}|^{2p}\mu^\varepsilon_{0,\boldsymbol{x}}(\mathrm{d}\boldsymbol{u}) \,\leq\, m_p\,, \tag{2.5}$$

and such that

$$\int_{K\times\mathbb{R}^2} |\boldsymbol{u}|^{2pr'}\,\rho_0^\varepsilon(\boldsymbol{x})\,\mu^\varepsilon_{0,\boldsymbol{x}}(\mathrm{d}\boldsymbol{u})\,\mathrm{d}\boldsymbol{x} \,\leq\, \overline{m}_p\,, \tag{2.6}$$

where p and r' are given in (2.1b) and (2.3).

Now let us define the notion of solution we will consider for equation (1.2).

Definition 2.2. *For all* $\varepsilon > 0$ *we say that* μ^ε *solves* (1.2) *with initial condition* μ_0^ε *if we have*

(1) μ^ε *lies in*

$$\mathscr{C}^0\left(\mathbb{R}^+\times K\,,\,L^1\left(\mathbb{R}^2\right)\right)\cap L^\infty_{loc}\left(\mathbb{R}^+\times K\,,\,\mathscr{P}^2\left(\mathbb{R}^2\right)\right)\,,$$

(2) for all $\boldsymbol{x}\in K$, $t\geq 0$, *and* $\varphi\in\mathscr{C}^\infty_c\left(\mathbb{R}^2\right)$ *, it holds*

$$\begin{aligned}\int_{\mathbb{R}^2}\varphi(\boldsymbol{u})\,\left(\mu^\varepsilon_{t,\boldsymbol{x}}-\mu^\varepsilon_{0,\boldsymbol{x}}\right)(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} \,=\, &\int_0^t\int_{\mathbb{R}^2}\left(\nabla_{\boldsymbol{u}}\varphi(\boldsymbol{u})\cdot\mathbf{b}^\varepsilon(s,\boldsymbol{x},\boldsymbol{u}) \,+\, \partial_v^2\varphi(\boldsymbol{u})\right)\mu^\varepsilon_{s,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}s\\ &-\frac{\rho_0^\varepsilon(\boldsymbol{x})}{\varepsilon}\int_0^t\int_{\mathbb{R}^2}\partial_v\varphi(\boldsymbol{u})\,\left(v-\mathcal{V}^\varepsilon(s,\boldsymbol{x})\right)\mu^\varepsilon_{s,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}s\,,\end{aligned}$$

where $\mathcal{V}^\varepsilon$ *and* $\mathbf{b}^\varepsilon$ *are given by* (1.1) *and* (1.2).

With this notion of solution, equation (1.2) is well-posed. Indeed, we have

Theorem 2.3. *For any $\varepsilon > 0$, suppose that assumptions (2.1a)-(2.1b) and (2.2)-(2.4) are fulfilled and that the initial condition μ_0^ε also verifies*

$$\begin{cases} \sup\limits_{\boldsymbol{x}\in K} \displaystyle\int_{\mathbb{R}^2} e^{|\boldsymbol{u}|^2/2}\, \mu_{0,\boldsymbol{x}}^\varepsilon(\mathrm{d}\boldsymbol{u}) \;\leq\; M^\varepsilon\,, \\ \sup\limits_{\boldsymbol{x}\in K} \displaystyle\int_{\mathbb{R}^2} \ln(\mu_{0,\boldsymbol{x}}^\varepsilon)\, \mu_{0,\boldsymbol{x}}^\varepsilon(\mathrm{d}\boldsymbol{u}) \;\leq\; m^\varepsilon\,, \end{cases} \tag{2.7}$$

and

$$\sup_{\boldsymbol{x}\in K} \left\| \nabla_{\boldsymbol{u}} \sqrt{\mu_{0,\boldsymbol{x}}^\varepsilon} \right\|^2_{L^2(\mathbb{R}^2)} \;\leq\; m^\varepsilon\,, \tag{2.8}$$

where M^ε and m^ε are two positive constant. Then there exists a unique solution μ^ε to equation (1.2) with initial condition μ_0^ε, in the sense of Definition 2.2 which verifies

$$\sup_{(t\,,\,\boldsymbol{x})\in[0,T]\times K} \int_{\mathbb{R}^2} e^{|\boldsymbol{u}|^2/2}\, \mu_{t,\boldsymbol{x}}^\varepsilon(\boldsymbol{u})\, \mathrm{d}\boldsymbol{u} < +\infty\,,$$

for all $T \geq 0$.
Furthermore, the macroscopic quantities $\mathcal{V}^\varepsilon$ and $\mathcal{W}^\varepsilon$ given in (1.1) lie in $\mathscr{C}^0\left(\mathbb{R}^+ \times K\right)$.

We postpone the proof of this result to the Appendix A, which relies on relative entropy estimates. We take advantage of assumption (2.7) to derive continuity estimates for both the time and the spatial variable. More precisely, we apply an abstract result from [3] which ensures that if we suppose some exponential moments such as in (2.7), then the Wasserstein metric is controlled by the relative entropy. We make use of assumption (2.8) in order to obtain strong continuity with respect to the time variable. We emphasize that since we are not able to close the estimates using directly the L^1 norm, we introduce some modified relative entropy, which turns out to be, in some sense, equivalent to the L^1 distance. Our approach is original to our knowledge.

Remark 2.4. *We do not make use of assumptions (2.7) & (2.8) in the analysis of the asymptotic $\varepsilon \to 0$. Therefore, both constants M^ε and m^ε may blow up as ε vanishes.*

We also define solutions for the limiting system (1.5)

Definition 2.5. *We say that $(\mathcal{V}\,,\,\bar{\mu})$ solves (1.5) with initial condition $(\mathcal{V}_0\,,\,\bar{\mu}_0)$ if we have*
(1) $\mathcal{V} \in \mathscr{C}^0\left(\mathbb{R}^+ \times K\right)$ and $\bar{\mu} \in \mathscr{C}^0\left(\mathbb{R}^+ \times K\,,\, L^1\left(\mathbb{R}\right)\right)$,
(2) $\mathcal{V}$ is a mild solution to (1.5),
(3) for all $t \geq 0$, all $\boldsymbol{x} \in K$ and all $\phi \in \mathscr{C}_c^\infty(\mathbb{R})$ we have

$$\int_{\mathbb{R}} \phi(w)\, \bar{\mu}_{t,\boldsymbol{x}}(w)\, \mathrm{d}w \;=\; \int_{\mathbb{R}} \phi(w)\, \bar{\mu}_{0,\boldsymbol{x}}(w)\, \mathrm{d}w \;+\; \int_0^t \int_{\mathbb{R}} A(\mathcal{V}, w)\, \partial_w \phi(w)\, \bar{\mu}_{s,\boldsymbol{x}}(w)\, \mathrm{d}w\, \mathrm{d}s\,.$$

Theorem 2.6. *Under assumptions (2.1a), (2.3)-(2.4), and for any initial condition*

$$(\mathcal{V}_0\,,\,\bar{\mu}_0) \in \mathscr{C}^0\left(K\right) \times \mathscr{C}^0\left(K\,,\, L^1\left(\mathbb{R}\right)\right),$$

there exists a unique solution to (1.5) in the sense of Definition 2.5 with initial condition $(\mathcal{V}_0\,,\,\bar{\mu}_0)$. Furthermore, $\mathcal{V}$ is uniformly bounded over $\mathbb{R}^+ \times K$.

Proof. The key argument is that the system (1.5) may be decoupled through the change of variable (1.6). Indeed, we consider the following system

$$\begin{cases} \partial_t \mathcal{V} = N(\mathcal{V}) - \mathcal{W} - \mathcal{L}_{\rho_0}[\mathcal{V}]\,, \\ \partial_t \mathcal{W} \;=\; A\left(\mathcal{V}\,,\,\mathcal{W}\right)\,, \\ \partial_t \bar{\nu} \;-\; b\, \partial_w\left(w\, \bar{\nu}\right) \;=\; 0\,, \end{cases}$$

which turns out to be equivalent to (1.5) in the sense that it is solved by $(\mathcal{V}\,,\,\mathcal{W}\,,\,\bar{\nu})$ if and only if $(\mathcal{V}\,,\,\bar{\mu})$ solves (1.5), where $\bar{\mu}$ is defined as

$$\bar{\mu}\,(t\,,\,\boldsymbol{x}\,,\,w) \,=\, \bar{\nu}\,(t\,,\,\boldsymbol{x}\,,\,w \,-\, \mathcal{W}(t\,,\,\boldsymbol{x}))\;,$$

for all $(t,\,\boldsymbol{x},\,w)$ in $\mathbb{R}^+ \times K \times \mathbb{R}$. Existence and uniqueness for $\bar{\nu}$ relies on classical arguments and we refer to [13], where one can find the proof of existence and uniqueness for the system $(\mathcal{V},\,\mathcal{W})$. □

2.2. **Main results.** The following theorem is the main result of this article. It states that in the regime of strong interactions, the distribution of the voltage variable concentrates with rate $\sqrt{\varepsilon}$ around $\mathcal{V}$ with Gaussian profile. The distribution of the adaptation variable converges towards $\bar{\mu}$. Hence, the couple $(\mathcal{V},\,\bar{\mu})$, which solves (1.5), encodes the behavior of the system when $\varepsilon \ll 1$. The result is global in time, uniform with respect to the spatial variable $\boldsymbol{x}$ and we obtain explicit convergence rate.

Theorem 2.7. *Under assumptions (2.1a)-(2.1b) on the drift N, (2.2)-(2.3) on Ψ, (2.4)-(2.6) on the initial conditions μ_0^ε and under the additional assumptions of Theorem 2.3, consider the solutions μ^ε & $(\mathcal{V},\,\bar{\mu})$ provided by Theorem 2.3 & 2.6 respectively. Furthermore, define the initial macroscopic and mesoscopic errors as*

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

and

$$\mathcal{E}_{\mathrm{mes}} \,=\, \sup_{\boldsymbol{x}\in K} W_2\left(\bar{\nu}_{0,\boldsymbol{x}}^\varepsilon, \bar{\nu}_{0,\boldsymbol{x}}\right)\;.$$

Then there exists $(C,\varepsilon_0) \in (\mathbb{R}_*^+)^2$ such that the following expansion holds for all $\varepsilon \leq \varepsilon_0$

$$W_2\left(\mu_{t,\,\boldsymbol{x}}^\varepsilon\,,\,\mathcal{M}_{\frac{1}{\varepsilon}\rho_0(\boldsymbol{x})\,,\,\mathcal{V}(t,\,\boldsymbol{x})} \otimes \bar{\mu}_{t,\,\boldsymbol{x}}\right) \,\leq\, C\,\left(\min\left(e^{C\,t}\,(\mathcal{E}_{\mathrm{mac}} \,+\, \varepsilon)\,,\,1\right) \,+\, \mathcal{E}_{\mathrm{mes}}\,e^{-b\,t} \,+\, e^{-\rho_0^\varepsilon(\boldsymbol{x})\,t\,/\,\varepsilon}\right)\;,$$

for all $(t,\boldsymbol{x}) \in \mathbb{R}^+ \times K$.
Moreover, suppose the initial errors to be of order ε, that is

$$\mathcal{E}_{\mathrm{mac}} \,+\, \mathcal{E}_{\mathrm{mes}} \underset{\varepsilon\to 0}{=} O\,(\varepsilon)\;,$$

then the following estimate holds

$$W_2\left(\mu_{t,\,\boldsymbol{x}}^\varepsilon\,,\,\mathcal{M}_{\frac{1}{\varepsilon}\rho_0(\boldsymbol{x})\,,\,\mathcal{V}(t,\,\boldsymbol{x})} \otimes \bar{\mu}_{t,\,\boldsymbol{x}}\right) \,\leq\, C\,\left(\min\left(e^{C\,t}\,\varepsilon\,,\,1\right) \,+\, e^{-\rho_0^\varepsilon(\boldsymbol{x})\,t/\varepsilon}\right)\;,\quad \forall (t,\boldsymbol{x}) \in \mathbb{R}^+ \times K\,.$$

Let us outline the main steps of the proof. First, we obtain in Proposition 3.1 some uniform moment estimates for μ^ε. Second, in Proposition 3.3, we estimate a relative energy, which should be interpreted as the moments of the re-scaled quantity ν^ε after a proper re-normalization, as mentioned in Remark 3.4. The last step consists in the convergence estimate. We focus on the re-scaled quantity ν^ε and develop an analytical coupling method in order to estimate its Wasserstein distance with the limiting profile. The key idea is to consider a coupled equation which is solved by couplings between ν^ε and its limit and then to estimate some energy for the solutions to the coupled equation.
An important feature in our work is that we **do not suppose** the initial condition μ_0^ε to be concentrated, nor the initial profile ν_0^ε to be close to its limit. In particular, this means that the problem on ν^ε is **ill-prepared**. Indeed, performing the change of variable (1.6) in the following *ansatz*

$$\inf_{\boldsymbol{x}\in K} \int_{\mathbb{R}^2} |v \,-\, \mathcal{V}_0^\varepsilon(\boldsymbol{x})|^2\,\mu_{0,\boldsymbol{x}}^\varepsilon(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} \,\geq\, 1\,,$$

we obtain that ν_0^ε blows up in $\mathscr{P}^2$ as ε vanishes

$$\inf_{\boldsymbol{x}\in K} \int_{\mathbb{R}^2} |v|^2\,\nu_{0,\boldsymbol{x}}^\varepsilon(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} \,\geq\, \varepsilon^{-1}\,.$$

This is why we prove in Proposition 3.3 some exponential localizing effects with respect to the variable t/ε, which also appears in Theorem 2.7.

Let us now mention some interpretations and consequences of our result. The first consequence of the latter result is the convergence of the averaged quantities $\mathcal{V}^\varepsilon$ and $\mathcal{W}^\varepsilon$. In fact, we prove a finer result since we obtain that the couple $(\mathcal{V}^\varepsilon, \bar{\mu}^\varepsilon)$ converges towards $(\mathcal{V}, \bar{\mu})$ with rate ε.

Corollary 2.8. *Under the assumptions of Theorem 2.7 and supposing the initial errors to be of order ε, that is*

$$\mathcal{E}_{\text{mac}} + \mathcal{E}_{\text{mes}} \underset{\varepsilon\to 0}{=} O(\varepsilon) ,$$

there exists $(C, \varepsilon_0) \in (\mathbb{R}_*^+)^2$ such that for all $\varepsilon \leq \varepsilon_0$

$$\sup_{\boldsymbol{x}\in K} \left(|\, \mathcal{V}^\varepsilon(t,\boldsymbol{x}) - \mathcal{V}(t,\boldsymbol{x}) \,| + W_2\left(\bar{\mu}^\varepsilon_{t,\boldsymbol{x}}, \bar{\mu}_{t,\boldsymbol{x}}\right)\right) \leq C\left(e^{-m_* t/\varepsilon} + \min\left(e^{C t}\varepsilon, 1\right)\right) ,$$

for all $t \geq 0$, where m_* is given by (2.4).

Proof. Using Jensen's inequality, we check that

$$|\mathcal{V}^\varepsilon(t,\boldsymbol{x}) - \mathcal{V}(t,\boldsymbol{x})| \leq W_2\left(\mu^\varepsilon_{t,\boldsymbol{x}}, \mathcal{M}_{\frac{\rho_0(\boldsymbol{x})}{\varepsilon},\mathcal{V}(t,\boldsymbol{x})} \otimes \bar{\mu}_{t,\boldsymbol{x}}\right) .$$

Furthermore, we use the definition of the Wasserstein distance to obtain

$$W_2\left(\bar{\mu}^\varepsilon_{t,\boldsymbol{x}}, \bar{\mu}_{t,\boldsymbol{x}}\right) \leq W_2\left(\mu^\varepsilon_{t,\boldsymbol{x}}, \mathcal{M}_{\frac{\rho_0(\boldsymbol{x})}{\varepsilon},\mathcal{V}(t,\boldsymbol{x})} \otimes \bar{\mu}_{t,\boldsymbol{x}}\right) .$$

Then we apply Theorem 2.7 and replace ρ_0^ε with its lower bound m_* given in assumption (2.4). □

Another interesting consequence of Theorem 2.7, which may be interpreted as the expansion of f^ε at order 1 when $\varepsilon \ll 1$, consists in recovering order 0. In fact, we prove a stronger result. Indeed, let us make the analogy with other types of expansions, let say with Taylor expansions. The expansion at order 1 yields an equivalence result at order 0. This is exactly what we obtain in our case: we prove that the distance between f^ε and $\delta_{\mathcal{V}} \otimes \bar{\mu}$ is exactly of order $\sqrt{\varepsilon}$. This justifies our approach for two reasons. First, it means that it is not possible to achieve convergence at order ε if restrict the analysis to the convergence towards a Dirac mass. Second, the choice of the Wasserstein metric in our analysis instead of another stronger norm enables to compare easily our result with the convergence of f^ε towards a Dirac mass.

Corollary 2.9. *Under the assumptions of Theorem 2.7 and supposing the initial errors to be of order $\sqrt{\varepsilon}$, that is*

$$\mathcal{E}_{\text{mac}} + \mathcal{E}_{\text{mes}} \underset{\varepsilon\to 0}{=} O\left(\sqrt{\varepsilon}\right) ,$$

there exists $(C, \varepsilon_0) \in (\mathbb{R}_*^+)^2$, such that for all $\varepsilon \leq \varepsilon_0$

$$\sup_{\boldsymbol{x}\in K} \left(W_2\left(\mu^\varepsilon_{t,\boldsymbol{x}}, \delta_{\mathcal{V}(t,\boldsymbol{x})} \otimes \bar{\mu}_{t,\boldsymbol{x}}\right)\right) \leq C\left(e^{-m_* t/\varepsilon} + \min\left(e^{Ct}\sqrt{\varepsilon}, 1\right)\right) ,$$

for all $t \geq 0$, where m_* is given by (2.4).
Moreover, supposing the initial errors to be of order ε, that is

$$\mathcal{E}_{\text{mac}} + \mathcal{E}_{\text{mes}} \underset{\varepsilon\to 0}{=} O(\varepsilon) ,$$

and considering two positive times t_0 and T such that $t_0 < T$, there exists $(C, \varepsilon_0) \in (\mathbb{R}_*^+)^2$ such that for all $\varepsilon \leq \varepsilon_0$

$$C^{-1}\sqrt{\varepsilon} \leq W_2\left(\mu^\varepsilon_{t,\boldsymbol{x}}, \delta_{\mathcal{V}(t,\boldsymbol{x})} \otimes \bar{\mu}_{t,\boldsymbol{x}}\right) \leq C\sqrt{\varepsilon}, \quad \forall (t,\boldsymbol{x}) \in [\,t_0\,, T\,] \times K .$$

Proof. The proof is a direct consequence of Theorem 2.7 and the triangular inequality for W_2. □

3. A priori estimates

In this section, we provide estimates for the moments of μ^ε and for the relative energy given by

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

where $q \geq 2$.

The key point here is to obtain uniform estimates with respect to time for both M_q^ε and D_q^ε using confining properties of N and A. It is actually the only place where we use the super-linear decay of the drift N (2.1a).

Proposition 3.1 (Propagation of moment)**.** *Under assumptions* (2.1a) *on the drift N and* (2.2)-(2.3) *on the interaction kernel Ψ, consider a sequence of solutions $(\mu^\varepsilon)_{\varepsilon>0}$ to* (1.2) *with initial conditions satisfying assumptions* (2.4)-(2.6)*. Then, for all positive ε and all q lying in $[\,2,\,2p\,]$ holds the following estimate*

$$M_q[\,\mu^\varepsilon\,](t,\boldsymbol{x}) \;\leq\; M_q[\,\mu^\varepsilon\,](0,\boldsymbol{x})\,\exp\left(-t/C\right) \,+\, C, \quad \forall\,(t,\boldsymbol{x}) \,\in\, \mathbb{R}^+\times K\,,$$

where $C>0$ is a positive constant which only depends on m_*, $\overline{m}_p$ and the data of the problem: Ψ, A_0 & N.

Proof. Let us choose an exponent $\theta \geq 2$, multiply equation (1.2) by $|\boldsymbol{u}|^\theta/\theta$ and integrate with respect to $\boldsymbol{u}\in\mathbb{R}^2$. Integrating by part, this leads us to the following relation

$$\frac{1}{\theta}\,\frac{\mathrm{d}}{\mathrm{d}t} M_\theta[\mu^\varepsilon](t,\boldsymbol{x}) \;=\; \mathcal{I}\,,$$

where $\mathcal{I}$ splits into $\mathcal{I} = \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3 + \mathcal{I}_4$ with

$$\begin{cases} \mathcal{I}_1 \;=\; -\dfrac{1}{\varepsilon}\,\rho_0^\varepsilon(\boldsymbol{x}) \displaystyle\int_{\mathbb{R}^2} v\,|v|^{\theta-2}\,\left(v \,-\, \mathcal{V}^\varepsilon(t)\right)\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,, \\ \mathcal{I}_2 \;=\; \displaystyle\int_{\mathbb{R}^2} v\,|v|^{\theta-2}\,N(v)\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,, \\ \mathcal{I}_3 \;=\; \displaystyle\int_{\mathbb{R}^2} \left(w\,|w|^{\theta-2}\,A(\boldsymbol{u}) - v|v|^{\theta-2}\,w \,+\, (\theta-1)\,|v|^{\theta-2}\right)\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,, \\ \mathcal{I}_4 \;=\; -\displaystyle\int_{\mathbb{R}^2} v\,|v|^{\theta-2}\,\mathcal{K}_\Psi[\rho_0^\varepsilon\,\mu^\varepsilon]\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,. \end{cases}$$

We first handle the stiff term $\mathcal{I}_1$, which can be simply re-written as

$$\mathcal{I}_1 \;=\; -\frac{\rho_0^\varepsilon(\boldsymbol{x})}{\varepsilon}\int_{\mathbb{R}^2}\left(v\,|v|^{\theta-2} \,-\, \mathcal{V}^\varepsilon\,|\mathcal{V}^\varepsilon|^{\theta-2}\right)\,\left(v\,-\,\mathcal{V}^\varepsilon\right)\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} \;\leq\; 0\,.$$

Then we evaluate $\mathcal{I}_2$, which may involve higher order moments due to the non-linearity N. To overcome this difficulty, we split it in two parts

$$v\,|v|^{\theta-2}\,N(v) \;=\; |v|^\theta\,\frac{N(v)}{v}\,\mathbb{1}_{|v|\geq 1} \;+\; v\,|v|^{\theta-2}\,N(v)\,\mathbb{1}_{|v|<1}\,.$$

According to assumption (2.1a) and since N is continuous, we obtain

$$\mathcal{I}_2 \;\leq\; \int_{\mathbb{R}^2}\left(C \,-\, \omega^-(v)\right)\,|v|^\theta\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} \,+\, C\,,$$

for some constant $C>0$ only depending on N and where ω^- is the following nonnegative function

$$\omega^-(v) \;=\; \left(\omega(v)\;\mathbb{1}_{|v|\geq 1,}\right)^-\,,$$

with $s^- = \max(0,-s)$ and where ω is given by (2.1a).

Now we evaluate $\mathcal{I}_3$, which gathers low order terms. Applying Young's inequality, we obtain the following estimate

$$\mathcal{I}_3 \,\leq\, \frac{C}{\eta^\theta}\int_{\mathbb{R}^2} |v|^\theta\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} \,+\, \int_{\mathbb{R}^2} (C\,\eta \,-\, b)\,|w|^\theta\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} \,+\, \frac{C}{\eta^\theta}\,,$$

for some constant $C>0$ and all $\eta\in]0,1[$.

Finally, to evaluate the non-local term $\mathcal{I}_4$, we estimate $\mathcal{V}^\varepsilon$ by applying Jensen's inequality, which yields

$$|\mathcal{V}^\varepsilon|^\theta \,\leq\, \int_{\mathbb{R}^2} |v|^\theta\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,,$$

hence, applying Young's inequality, we obtain that

$$\mathcal{I}_4 \,\leq\, C\,\left(\|\Psi *_r \rho_0^\varepsilon\|_{L^\infty(K)}\int_{\mathbb{R}^2} |v|^\theta\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} \,+\, \int_{K\times\mathbb{R}^2} |\Psi(\boldsymbol{x},\boldsymbol{x}')|\;|v|^\theta\,\rho_0^\varepsilon(\boldsymbol{x}')\,\mu^\varepsilon_{t,\boldsymbol{x}'}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{x}'\right)\,,$$

where C is a positive constant only depending on θ. Then we use Hölder's inequality and assumption (2.3) & (2.4) (we do not use the constraint $r>1$ here), which yields

$$\|\,\Psi *_r \rho_0^\varepsilon\,\|_{L^\infty(K)} \,\leq\, \|\,\rho_0^\varepsilon\,\|_{L^\infty(K)}\sup_{\boldsymbol{x}\in K}\|\,\Psi(\boldsymbol{x},\cdot)\,\|_{L^r(K)} \,\leq\, C\,,$$

for some $C>0$ independent of ε.

In the former computations we choose η such that $b - C\eta > 0$. With the notation $\alpha = b - C\eta$, this yields

$$\begin{aligned}\frac{1}{\theta}\,\frac{\mathrm{d}}{\mathrm{d}t}\,M_\theta[\,\mu^\varepsilon\,](t,\boldsymbol{x}) \;&\leq\; \int_{\mathbb{R}^2}\left[\,\big(C\,-\,\omega^-(v)\big)\;|v|^\theta \,-\, \alpha\,|w|^\theta\,\right]\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} \,+\, C\\ &\quad + C\int_{K\times\mathbb{R}^2} |\Psi(\boldsymbol{x},\boldsymbol{x}')|\,|v|^\theta\,\rho_0^\varepsilon(\boldsymbol{x}')\,\mu^\varepsilon_{t,\boldsymbol{x}'}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{x}'\,,\end{aligned} \tag{3.1}$$

for another constant $C>0$ depending on θ, m_*, A, N & Ψ but not on $(t,\,\boldsymbol{x})\in\mathbb{R}^+\times K$ nor on ε.

Now, we fix q in $[2,\,2p]$ and proceed in two steps. On the one hand, choosing $\theta = qr'\geq 2$ in (3.1), we evaluate the averaged moments $\overline{M}_{qr'}[\rho_0^\varepsilon\,\mu^\varepsilon]$ given by

$$\overline{M}_{qr'}\,[\,\rho_0^\varepsilon\,\mu^\varepsilon\,]\,(t) \,=\, \int_{K\times\mathbb{R}^2} |\boldsymbol{u}|^{qr'}\,\rho_0^\varepsilon(\boldsymbol{x})\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{x}\,,$$

where r' is given by (2.3).

On the other hand, choosing $\theta = q$ in (3.1), we use the latter estimate to control the non-local contribution on the right hand side of (3.1) and evaluate $M_q[\mu^\varepsilon](t,\boldsymbol{x})$ at each $(t,\boldsymbol{x})\in\mathbb{R}^+\times K$.

Starting from (3.1) with $\theta = qr'\geq 2$, we multiply it by $\rho_0^\varepsilon(\boldsymbol{x})$ and integrate with respect to $\boldsymbol{x}\in K$, which yields

$$\begin{aligned}\frac{1}{qr'}\,\frac{\mathrm{d}}{\mathrm{d}t}\overline{M}_{qr'}\,[\,\rho_0^\varepsilon\,\mu^\varepsilon\,] \,\leq\, &\int_{K\times\mathbb{R}^2}\Big(\big(C\,-\,\omega^-(v)\big)\;|v|^{qr'} \,-\, \alpha\,|w|^{qr'}\Big)\;\rho_0^\varepsilon(\boldsymbol{x})\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{x}\\ &+\, C\int_{K\times K\times\mathbb{R}^2} |\Psi(\boldsymbol{x},\boldsymbol{x}')|\,|v|^{qr'}\,\rho_0^\varepsilon(\boldsymbol{x})\;\rho_0^\varepsilon(\boldsymbol{x}')\,\mu^\varepsilon_{t,\boldsymbol{x}'}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{x}'\,\mathrm{d}\boldsymbol{x} \,+\, C\,.\end{aligned}$$

According to assumption (2.2) & (2.4), we have

$$\int_{K\times K\times\mathbb{R}^2} |\Psi(\boldsymbol{x},\boldsymbol{x}')|\,|v|^{qr'}\,\rho_0^\varepsilon(\boldsymbol{x})\;\rho_0^\varepsilon(\boldsymbol{x}')\,\mu^\varepsilon_{t,\boldsymbol{x}'}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{x}'\,\mathrm{d}\boldsymbol{x} \,\leq\, C\int_{K\times\mathbb{R}^2} |v|^{qr'}\,\rho_0^\varepsilon(\boldsymbol{x})\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{x}\,,$$

for some positive constant C depending on m_* and Ψ. Hence it yields

$$\frac{1}{qr'}\,\frac{\mathrm{d}}{\mathrm{d}t}\,\overline{M}_{qr'}\,[\,\rho_0^\varepsilon\,\mu^\varepsilon\,]\,(t) \leq \int_{K\times\mathbb{R}^2}\Big(\big(C\,-\,\omega^-(v)\big)\;|v|^{qr'} \,-\, \alpha\,|w|^{qr'}\Big)\;\rho_0^\varepsilon(\boldsymbol{x})\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{x} \,+\, C\,.$$

According to assumption (2.1a), $\omega^-(v)$ goes to infinity with $|v|$. Consequently, we are led to

$$\frac{1}{qr'}\,\frac{\mathrm{d}}{\mathrm{d}t}\,\overline{M}_{qr'}\,[\,\rho_0^\varepsilon\,\mu^\varepsilon\,]\,(t) \,\leq\, C \,-\, \frac{1}{C}\,\overline{M}_{qr'}\,[\,\rho_0^\varepsilon\,\mu^\varepsilon\,]\,(t),$$

for $C > 0$ great enough. Using Gronwall's lemma and the assumption (2.6) on the non-local moment of μ_0^ε, we obtain that for all $\varepsilon > 0$

$$\overline{M}_{qr'}[\rho_0^\varepsilon \mu^\varepsilon](t) \leq C\,, \ \forall t \in \mathbb{R}^+\,,$$

where C may depend on $\overline{m}_p$.

Now, we come back to the local estimate on the moment $M_q[\mu^\varepsilon](t,\boldsymbol{x})$. We replace θ with q in (3.1) and estimate the non-local contribution. First, we apply Hölder's inequality and assumptions (2.3) & (2.4) to the non-local contribution in (3.1). This yields

$$\int_{K\times\mathbb{R}^2} |\Psi(\boldsymbol{x},\boldsymbol{x}')|\,|v|^q\,\rho_0^\varepsilon(\boldsymbol{x}')\,\mu^\varepsilon_{t,\boldsymbol{x}'}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{x}' \leq C\,\left|\overline{M}^\varepsilon_{qr'}(t)\right|^{1/r'}$$

We emphasize that we use the constraint $r > 1$ in the former estimate. Then we use the latter bound on $\overline{M}^\varepsilon_{qr'}(t)$ and obtain

$$\int_{K\times\mathbb{R}^2} |\Psi(\boldsymbol{x},\boldsymbol{x}')|\,|v|^q\,\rho_0^\varepsilon(\boldsymbol{x}')\,\mu^\varepsilon_{t,\boldsymbol{x}'}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{x}' \leq C\,.$$

Hence, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}M_q[\mu^\varepsilon](t,\boldsymbol{x}) \leq \int_{\mathbb{R}^2}\left(\left(C - \omega^-(v)\right)\,|v|^q - \alpha|w|^q\right)\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} + C\,.$$

Using the same arguments as before, we obtain some $C > 0$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t}M_q[\mu^\varepsilon](t,\boldsymbol{x}) \leq C - \frac{1}{C}\,M_q[\mu^\varepsilon](t,\boldsymbol{x})\,.$$

We conclude this proof applying Gronwall's lemma. □

As a straightforward consequence of Proposition 3.1, we obtain uniform bounds with respect to time, space and ε for the macroscopic quantities

Corollary 3.2. *Under the assumptions of Proposition 3.1, there exists a constant $C > 0$, independent of ε, such that*

$$|\mathcal{V}^\varepsilon(t,\boldsymbol{x})| + |\mathcal{W}^\varepsilon(t,\boldsymbol{x})| \leq C, \quad \forall\,(t,\boldsymbol{x}) \in \mathbb{R}^+\times K\,,$$

where $\mathcal{V}^\varepsilon$ and $\mathcal{W}^\varepsilon$ are given by (1.1).

Proof. Applying the Cauchy-Schwarz inequality, we have for any $(t,\boldsymbol{x}) \in \mathbb{R}^+\times K$,

$$|\mathcal{V}^\varepsilon(t,\boldsymbol{x})| + |\mathcal{W}^\varepsilon(t,\boldsymbol{x})| \leq 2\,|M_2[\mu^\varepsilon](t,\boldsymbol{x})|^{1/2}\,,$$

hence the result follows on from Proposition 3.1. □

We turn to the estimates for the relative energy $D_q[\mu^\varepsilon]$, which quantifies the convergence of μ^ε towards a Dirac mass centered on $\mathcal{V}^\varepsilon$.

Proposition 3.3 (Relative energy)**.** *Under assumptions (2.1a)-(2.1b) on the drift N and (2.2)-(2.3) on the interaction kernel Ψ, consider a sequence of solutions $(\mu^\varepsilon)_{\varepsilon>0}$ to (1.2) with initial conditions satisfying assumption (2.4)-(2.6). There exists a positive constant $C > 0$, which may depend on m_p and $\overline{m}_p$, such that for all $\varepsilon > 0$ and all q in $[2,\,2p]$ holds the following estimate*

$$D_q[\,\mu^\varepsilon\,](t,\boldsymbol{x}) \leq C\left[D_q[\,\mu^\varepsilon\,](0,\boldsymbol{x})\,\exp\left(-\frac{q\rho_0^\varepsilon(\boldsymbol{x})}{\varepsilon}\,t\right) + \left(\frac{\varepsilon}{\rho_0^\varepsilon(\boldsymbol{x})}\right)^{q/2}\right], \quad \forall (t,\boldsymbol{x}) \in \mathbb{R}^+\times K\,.$$

Proof. We choose some q in $[2,\,2p]$, multiply equation (1.2) by $|v-\mathcal{V}^\varepsilon|^q/q$ and integrate with respect to $\boldsymbol{u} \in \mathbb{R}^2$. After integrating by part and using equation (1.1), it yields

$$\frac{1}{q}\frac{\mathrm{d}}{\mathrm{d}t}D_q[\mu^\varepsilon](t,\boldsymbol{x}) = \mathcal{J}\,,$$

where $\mathcal{J}$ is split as

$$\mathcal{J} \,=\, \mathcal{J}_1 \,+\, \mathcal{J}_2 \,+\, \mathcal{J}_3 \,-\, \frac{\rho_0^\varepsilon}{\varepsilon} D_q[\mu^\varepsilon](t,\boldsymbol{x}) \,+\, (q-1)\, D_{q-2}^\varepsilon[\mu^\varepsilon](t,\boldsymbol{x})\,,$$

with

$$\left\{\begin{array}{l} \displaystyle \mathcal{J}_1 \,=\, -\int_{\mathbb{R}^2} (v \,-\, \mathcal{V}^\varepsilon)\, |v \,-\, \mathcal{V}^\varepsilon|^{q-2} \left(\mathcal{K}_\Psi[\rho_0^\varepsilon\,\mu^\varepsilon] \,-\, \mathcal{L}_{\rho_0^\varepsilon}[\mathcal{V}^\varepsilon]\right)\, \mu_{t,\boldsymbol{x}}^\varepsilon(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,, \\ \displaystyle \mathcal{J}_2 \,=\, \int_{\mathbb{R}^2} (v \,-\, \mathcal{V}^\varepsilon)\, |v \,-\, \mathcal{V}^\varepsilon|^{q-2}\, (N(v) \,-\, N(\mathcal{V}^\varepsilon))\, \mu_{t,\boldsymbol{x}}^\varepsilon(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,, \\ \displaystyle \mathcal{J}_3 \,=\, -\int_{\mathbb{R}^2} (v \,-\, \mathcal{V}^\varepsilon)\, |v \,-\, \mathcal{V}^\varepsilon|^{q-2} \left(w \,-\, \mathcal{W}^\varepsilon \,+\, \mathcal{E}(\mu_{t,\boldsymbol{x}}^\varepsilon)\right)\, \mu_{t,\boldsymbol{x}}^\varepsilon(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,. \end{array}\right.$$

To estimate the non-local term $\mathcal{J}_1$, we observe that

$$\mathcal{K}_\Psi[\rho_0^\varepsilon\,\mu^\varepsilon] \,-\, \mathcal{L}_{\rho_0^\varepsilon}(\mathcal{V}^\varepsilon) \,=\, (v \,-\, \mathcal{V}^\varepsilon)\, \Psi *_r \rho_0^\varepsilon(\boldsymbol{x})\,,$$

hence, using assumptions (2.3) & (2.4), we obtain

$$\mathcal{J}_1 \,=\, -\,\Psi *_r \rho_0^\varepsilon(\boldsymbol{x})\; D_q[\mu^\varepsilon](t,\boldsymbol{x}) \,\le\, C\, D_q[\mu^\varepsilon](t,\boldsymbol{x})\,.$$

We turn to $\mathcal{J}_2$, which involves higher order moments since it displays the non-linearity N. On the one hand, Corollary 3.2 ensures that $\mathcal{V}^\varepsilon$ is uniformly bounded. On the other hand N lies in $\mathscr{C}^1(\mathbb{R})$ and meets the confining assumption (2.1a). Hence, there exists a constant $C>0$ independent of ε such that

$$(v-\mathcal{V}^\varepsilon)\, (N(v) \,-\, N(\mathcal{V}^\varepsilon)) \,\le\, C\, |v-\mathcal{V}^\varepsilon|^2\,, \quad \forall (t,\,\boldsymbol{x},\,v) \in \mathbb{R}^+ \times K \times \mathbb{R}\,.$$

Hence, we obtain

$$\mathcal{J}_2 \,\le\, C\, D_q[\mu^\varepsilon](t,\boldsymbol{x})\,.$$

Finally we estimate $\mathcal{J}_3$, which gathers the low order terms. According to assumption (2.1b) on N and applying Proposition 3.1, we have for all positive ε

$$\int_{\mathbb{R}^2} |w \,-\, \mathcal{W}^\varepsilon|^q\, \mu_{t,\boldsymbol{x}}^\varepsilon(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} \,+\, |\mathcal{E}(\mu_{t,\boldsymbol{x}}^\varepsilon)| \,\le\, C\,, \quad \forall\,(t,\,\boldsymbol{x}) \in \mathbb{R}^+ \times K\,,$$

for some constant C that may depend on m_p and $\overline{m}_p$. Hence, applying Hölder's inequality, it yields

$$\mathcal{J}_3 \,\le\, C\, D_q[\mu^\varepsilon]^{(q-1)/q}(t,\boldsymbol{x}) \,\le\, C\, \left(D_q[\mu^\varepsilon](t,\boldsymbol{x}) \,+\, D_q[\mu^\varepsilon]^{(q-2)/q}(t,\boldsymbol{x})\right)\,.$$

Gathering these computations and applying Hölder's inequality to D_{q-2}^ε, we obtain

$$\frac{1}{q}\,\frac{\mathrm{d}}{\mathrm{d}t} D_q[\mu^\varepsilon](t,\boldsymbol{x}) \,+\, \frac{\rho_0^\varepsilon}{\varepsilon} D_q[\mu^\varepsilon](t,\boldsymbol{x}) \,\le\, C\, \left(D_q[\mu^\varepsilon](t,\boldsymbol{x}) \,+\, D_q[\mu^\varepsilon]^{(q-2)/q}(t,\boldsymbol{x})\right)\,.$$

To estimate $D_q[\mu^\varepsilon]$, we introduce the function $u \,=\, (D_q[\mu^\varepsilon])^{2/q}$, which satisfies the following differential inequality

$$\frac{1}{2}\,\frac{\mathrm{d}u}{\mathrm{d}t} \,+\, \frac{\rho_0^\varepsilon}{\varepsilon}\, u \,\le\, C\, (u \,+\, 1)\,.$$

Applying Proposition 3.1, we get a first bound on u since

$$D_q[\mu^\varepsilon](t,\boldsymbol{x}) \,\le\, C\; M_q[\mu^\varepsilon](t,\boldsymbol{x}) \,\le\, C\,,$$

hence, we substitute this estimate on the right hand side of the former differential inequality and obtain

$$\frac{\mathrm{d}u}{\mathrm{d}t} \,+\, \frac{2\,\rho_0^\varepsilon}{\varepsilon}\, u \,\le\, C\,,$$

which implies that

$$u(t) \,\le\, u(0)\, \exp\left(-\frac{2\rho_0^\varepsilon\, t}{\varepsilon}\right) \,+\, C\,\frac{\varepsilon}{\rho_0^\varepsilon}\, \left(1 \,-\, \exp\left(-\frac{2\,\rho_0^\varepsilon\, t}{\varepsilon}\right)\right)\,.$$

The result follows from replacing u by $(D_q[\mu^\varepsilon])^{2/q}$. □

Both Propositions 3.1 & 3.3 may be interpreted in terms of the re-scaling ν^ε according to the following remark

Remark 3.4. *Performing the change of variable (1.6) in the expression of $M_q[\mu^\varepsilon]$ and $D_q[\mu^\varepsilon]$, we obtain the following relations*

$$\begin{cases} \displaystyle\int_{\mathbb{R}^2} |v|^q \, \nu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u} \;=\; \varepsilon^{-q/2} D_q[\mu^\varepsilon](t,\boldsymbol{x}) \,, \\ \displaystyle\int_{\mathbb{R}^2} |w|^q \, \nu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u} \;=\; \int_{\mathbb{R}^2} |w - \mathcal{W}^\varepsilon|^q \, \mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u} \;\leq\; C \, M_q[\mu^\varepsilon](t,\boldsymbol{x}) \,. \end{cases}$$

To conclude this section, we deduce from Propositions 3.1 & 3.3 the following error estimate

Proposition 3.5. *Under assumptions (2.1a)-(2.1b) on the drift N, (2.2)-(2.3) on the interaction kernel Ψ consider a sequence of solutions $(\mu^\varepsilon)_{\varepsilon>0}$ to (1.2) with initial conditions satisfying assumption (2.4)-(2.6). There exists a constant $C>0$ such that for all $\varepsilon>0$ we have*

$$\left|\mathcal{E}(\mu^\varepsilon_{t,\boldsymbol{x}})\right| \leq C \left(e^{-2\rho_0^\varepsilon(\boldsymbol{x})t/\varepsilon} + \varepsilon\right) \,, \quad \forall (t,\boldsymbol{x}) \in \mathbb{R}^+ \times K \,.$$

Proof. We rewrite the error term $\mathcal{E}(\mu^\varepsilon_{t,\boldsymbol{x}})$ given by (1.4) as follows

$$\mathcal{E}(\mu^\varepsilon_{t,\boldsymbol{x}}) \;=\; \mathcal{E}_1 \,+\, \mathcal{E}_2 \,+\, \mathcal{E}_3 \,,$$

where

$$\begin{cases} \mathcal{E}_1 \;=\; N'\left(\mathcal{V}^\varepsilon\right) \displaystyle\int_{\mathbb{R}^2} \left(v - \mathcal{V}^\varepsilon\right) \mathbb{1}_{|v-\mathcal{V}^\varepsilon|\leq 1} \, \mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u} \,, \\ \mathcal{E}_2 \;=\; \displaystyle\int_{\mathbb{R}^2} \left(N\left(v\right) - N\left(\mathcal{V}^\varepsilon\right) - N'\left(\mathcal{V}^\varepsilon\right)\left(v - \mathcal{V}^\varepsilon\right)\right) \mathbb{1}_{|v-\mathcal{V}^\varepsilon|\leq 1} \, \mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u} \,, \\ \mathcal{E}_3 \;=\; \displaystyle\int_{\mathbb{R}^2} \left(N\left(v\right) - N\left(\mathcal{V}^\varepsilon\right)\right) \mathbb{1}_{|v-\mathcal{V}^\varepsilon|> 1} \, \mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u} \,. \end{cases}$$

We start with $\mathcal{E}_1$. According to the definition of $\mathcal{V}^\varepsilon$, we have

$$\mathcal{E}_1 = 0 \;-\; N'\left(\mathcal{V}^\varepsilon\right) \int_{\mathbb{R}^2} \left(v - \mathcal{V}^\varepsilon\right) \mathbb{1}_{|v-\mathcal{V}^\varepsilon|>1} \, \mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u} \,.$$

Since $N \in \mathscr{C}^1(\mathbb{R})$ and applying Corollary 3.2 & Proposition 3.3, it yields

$$|\mathcal{E}_1| \;\leq\; C \left(e^{-2\rho_0^\varepsilon t/\varepsilon} \,+\, \varepsilon\right) \,.$$

We turn to $\mathcal{E}_2$. Since N is of class $\mathscr{C}^2$, and using Corollary 3.2, we have

$$\left|N\left(v\right) - N\left(\mathcal{V}^\varepsilon\right) - N'\left(\mathcal{V}^\varepsilon\right)\left(v - \mathcal{V}^\varepsilon\right)\right| \mathbb{1}_{|v-\mathcal{V}^\varepsilon|\leq 1} \;\leq\; C \left|v - \mathcal{V}^\varepsilon\right|^2 \,,$$

for some constant $C>0$ independent of $(t,\boldsymbol{x},v) \in \mathbb{R}^+ \times K \times \mathbb{R}$ and $\varepsilon>0$. Hence, we obtain the following estimate applying Proposition 3.3

$$|\mathcal{E}_2| \;\leq\; C \left(e^{-2\rho_0^\varepsilon t/\varepsilon} \,+\, \varepsilon\right) \,.$$

To end with, we estimate $\mathcal{E}_3$ using the assumption (2.1b) on N and Corollary 3.2. Indeed, we have

$$\left|N\left(v\right) - N\left(\mathcal{V}^\varepsilon\right)\right| \mathbb{1}_{|v-\mathcal{V}^\varepsilon|>1} \;\leq\; C \left|v - \mathcal{V}^\varepsilon\right|^p \,.$$

Hence, according to Proposition 3.3, we obtain

$$|\mathcal{E}_3| \leq C \left(e^{-p\rho_0^\varepsilon t/\varepsilon} \,+\, \varepsilon^{p/2}\right) \,.$$

Gathering these computations, we obtain the expected result. □

In the present section, we derived pointwise estimates for the moments of μ^ε, a relative energy which corresponds to the moments of ν^ε and the macroscopic error term $\mathcal{E}(\mu^\varepsilon)$. We build on these results to prove Theorem 2.7.

4. Proof of Theorem 2.7

To achieve the proof of Theorem 2.7, we quantify the concentration of $(\mu^\varepsilon)_{\varepsilon>0}$ around its asymptotic profile $\mathcal{M}_{\rho_0/\varepsilon,\mathcal{V}} \otimes \bar{\mu}$ when $\varepsilon \to 0$. Therefore, this section consists in estimating the error term $W_2\left(\mu^\varepsilon, \mathcal{M}_{\rho_0/\varepsilon,\mathcal{V}} \otimes \bar{\mu}\right)$, which we decompose as follows: for any $(t,\boldsymbol{x}) \in \mathbb{R}^+ \times K$,

$$W_2^2\left(\mu^\varepsilon_{t,\boldsymbol{x}},\, \mathcal{M}_{\rho_0/\varepsilon,\mathcal{V}} \otimes \bar{\mu}_{t,\boldsymbol{x}}\right) \,\leq\, \mathcal{D}_1 \,+\, \mathcal{D}_2 \,+\, \mathcal{D}_3,$$

where $\mathcal{D}_1$, $\mathcal{D}_2$ and $\mathcal{D}_3$ are defined as

$$\begin{cases} \mathcal{D}_1 \,:=\, W_2^2\left(\mu^\varepsilon_{t,\,\boldsymbol{x}}\,,\, \mathcal{M}_{\rho_0^\varepsilon/\varepsilon,\,\mathcal{V}^\varepsilon} \otimes \bar{\nu}_{t,\,\boldsymbol{x}}\left(\cdot \,-\, \mathcal{W}^\varepsilon\right)\right)\,, \\ \mathcal{D}_2 \,:=\, W_2^2\left(\mathcal{M}_{\rho_0^\varepsilon/\varepsilon,\mathcal{V}^\varepsilon} \otimes \bar{\nu}_{t,\boldsymbol{x}}\left(\cdot - \mathcal{W}^\varepsilon\right)\,,\, \mathcal{M}_{\rho_0^\varepsilon/\varepsilon,\mathcal{V}} \otimes \bar{\mu}_{t,\boldsymbol{x}}\right)\,, \\ \mathcal{D}_3 \,:=\, W_2^2\left(\mathcal{M}_{\rho_0^\varepsilon/\varepsilon,\mathcal{V}} \otimes \bar{\mu}_{t,\boldsymbol{x}}\,,\, \mathcal{M}_{\rho_0/\varepsilon,\mathcal{V}} \otimes \bar{\mu}_{t,\boldsymbol{x}}\right)\,. \end{cases}$$

The term $\mathcal{D}_1$ quantifies the distance between the concentration profile ν^ε and the Gaussian distribution since from the change of variable (1.6), we have

$$\mathcal{D}_1 \,=\, \inf_{\pi^\varepsilon \in \Pi^\varepsilon} \int_{\mathbb{R}^4} \left(\varepsilon\,|v-v'|^2 + |w-w'|^2\right) \mathrm{d}\pi^\varepsilon(\boldsymbol{u},\boldsymbol{u}')\,,$$

where Π^ε stands for the set of distributions over $\mathbb{R}^4$ with marginals $\nu^\varepsilon_{t,\boldsymbol{x}}$ and $\mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}_{t,\boldsymbol{x}}$, whereas $\mathcal{D}_2$ quantifies the error between the macroscopic quantities. Indeed, with the same change of variable, we obtain

$$\mathcal{D}_2 = |\mathcal{V}^\varepsilon - \mathcal{V}|^2 + |\mathcal{W}^\varepsilon - \mathcal{W}|^2\,.$$

Finally, the term $\mathcal{D}_3$ is simply the error on the macroscopic density. In the following we give an estimate for each term $(\mathcal{D}_i)_{1\leq i\leq 3}$.

Estimate for $\mathcal{D}_1$. Our strategy to estimate $\mathcal{D}_1$ consists in introducing a coupled equation on $\pi^\varepsilon \in \Pi^\varepsilon$. Let us consider the following problem,

$$\begin{aligned} &\partial_t \pi^\varepsilon \,+\, \mathrm{div}_{\boldsymbol{u}}\left[\mathbf{b}_0^\varepsilon(t,\boldsymbol{x},\boldsymbol{u})\,\pi^\varepsilon\right] \,+\, \partial_{w'}\left[A_0\,(0,w')\,\pi^\varepsilon\right] \\ &\quad=\, \frac{1}{\varepsilon}\,\mathrm{div}_{v,v'}\left[\rho_0^\varepsilon(\boldsymbol{x})\,(v,\,v')^\top\,\pi^\varepsilon \,+\, D\cdot\nabla_{v,v'}\pi^\varepsilon\right], \end{aligned} \tag{4.1}$$

where $\mathbf{b}_0^\varepsilon$ is given in (1.7) and the diffusion matrix D is given by

$$D \,=\, \begin{pmatrix} 1 & \beta \\ \beta & 1 \end{pmatrix}\,,$$

for some $\beta \in [-1,1]$. On the one hand, integrating the former equation with respect to $\boldsymbol{u}' = (v',w')$, we obtain equation (1.7) on ν^ε. On the other hand, equation (4.1) integrated with respect to $\boldsymbol{u}$ is given by

$$\partial_t\,\nu \,+\, \partial_{w'}\left[A_0\,(0,w')\,\nu\right] \,=\, \frac{1}{\varepsilon}\partial_{v'}\left(\rho_0^\varepsilon v'\,\nu + \partial_{v'}\,\nu\right)\,,$$

which is solved by $\mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}$. Consequently, if we take some initial data π_0^ε such that

$$\pi_0^\varepsilon \in \Pi\left(\nu^\varepsilon_{0,\boldsymbol{x}}\,,\, \mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}_{0,\boldsymbol{x}}\right)\,,$$

we obtain that the solution π^ε to (4.1) has marginals $\nu^\varepsilon_{t,\boldsymbol{x}}$ with respect to $\boldsymbol{u}$ and $\mathcal{M}_{\rho_0^\varepsilon} \otimes \bar{\nu}_{t,\boldsymbol{x}}$ with respect to $\boldsymbol{u}'$ at all time $t \geq 0$. Hence, according to the definition of the Wasserstein metric, the following inequality holds

$$\mathcal{D}_1 \,\leq\, \int_{\mathbb{R}^4} \left(\varepsilon\,|v-v'|^2 \,+\, |w-w'|^2\right) \mathrm{d}\pi^\varepsilon(\boldsymbol{u},\boldsymbol{u}')\,.$$

Moreover, we say that the equation is coupled because of the diffusion matrix D. Condition $\beta \in [-1,1]$ ensures that the matrix is positive and hence is required in order for equation (4.1) to be well-posed. In the case $\beta = 0$, there is no coupling between variables $\boldsymbol{u}$ and $\boldsymbol{u}'$. However,

we can see that there is only one suitable choice for the parameter β. Indeed, considering only the leading order in equation (4.1), the equation rewrites

$$\partial_t \pi^\varepsilon \;=\; \frac{1}{\varepsilon}\,\mathrm{div}_{v,v'}\left[\rho_0^\varepsilon\,(v,\,v')^\top\,\pi^\varepsilon \;+\; D\cdot\nabla_{v,v'}\pi^\varepsilon\right].$$

Multiplying the former equation by $|v-v'|^2/2$ and integrating by part, we obtain

$$\frac{1}{2}\,\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^4}|v-v'|^2\,\mathrm{d}\pi^\varepsilon(\boldsymbol{u},\boldsymbol{u}') \;+\; \frac{\rho_0^\varepsilon}{\varepsilon}\int_{\mathbb{R}^4}|v-v'|^2\,\mathrm{d}\pi^\varepsilon(\boldsymbol{u},\boldsymbol{u}') \;=\; 2\,\frac{1-\beta}{\varepsilon}\,.$$

Consequently, $\beta=1$ is the unique suitable choice in order to avoid blow up in the asymptotic $\varepsilon\to 0$. Hence, **we fix parameter** β to 1 in what follows. On a probabilistic point of view, this choice corresponds to taking the same Brownian motion for the processes associated to ν^ε and $\mathcal{M}_{\rho_0^\varepsilon}\otimes\bar\nu$.
Before estimating $\mathcal{D}_1$, we precise the nature of solutions we consider for equation (4.1)

Definition 4.1. *For any $\boldsymbol{x}$ in K and some $\varepsilon>0$, we say that π^ε solves (4.1) with initial condition π_0^ε if we have*

(1) π^ε lies in

$$\mathscr{C}^0\left(\mathbb{R}^+\,,\,\mathcal{D}'\left(\mathbb{R}^4\right)\right)\cap L^\infty_{loc}\left(\mathbb{R}^+\,,\,L\,\log L\left(\mathbb{R}^4\right)\right),$$

where $\mathcal{D}'(\mathbb{R}^4)$ stands for the set of distributions over $\mathbb{R}^4$ and $L\log L$ stands for the set of L^1 functions with finite entropy over $\mathbb{R}^4$.

(2) for all $t\geq 0$, and $\varphi\in\mathscr{C}^\infty_c\left(\mathbb{R}^4\right)$, it holds

$$\int_{\mathbb{R}^4}\varphi\,(\pi_t^\varepsilon-\pi_0^\varepsilon)\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{u}' \;=\; \int_0^t\int_{\mathbb{R}^4}\left(\nabla_{\boldsymbol{u}}\varphi\cdot\mathbf{b}_0^\varepsilon(s,\boldsymbol{x},\boldsymbol{u}) \;+\; \partial_{w'}\varphi\,A_0(0,\,w')\right)\pi_s^\varepsilon\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{u}'\,\mathrm{d}s$$

$$-\,\frac{1}{\varepsilon}\int_0^t\int_{\mathbb{R}^4}\left(\rho_0^\varepsilon(\boldsymbol{x})\nabla_{v,v'}\varphi\cdot(v,v')^\top \;-\; \mathrm{div}_{(v,v')}\left(D\cdot\nabla_{v,v'}\varphi\right)\right)\pi_s^\varepsilon\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{u}'\,\mathrm{d}s\,,$$

where $\mathcal{V}^\varepsilon$ and $\mathbf{b}_0^\varepsilon$ are given by (1.1) and (1.2).

We prove the following result for equation (4.1)

Proposition 4.2. *Under the assumptions of Theorem 2.3, consider a positive ε, some $\boldsymbol{x}$ lying in K and a coupling*

$$\pi_0^\varepsilon \;\in\; \Pi\left(\nu^\varepsilon_{0,\boldsymbol{x}}\,,\,\mathcal{M}_{\rho_0^\varepsilon}\otimes\bar\nu_{0,\boldsymbol{x}}\right).$$

There exists a solution π^ε to equation (4.1) with initial condition π_0^ε and parameter $\beta=1$ in the sense of Definition 4.1. Furthermore, we have

$$\pi^\varepsilon(t,\cdot)\in\Pi\left(\nu^\varepsilon_{t,\boldsymbol{x}}\,,\,\mathcal{M}_{\rho_0^\varepsilon}\otimes\bar\nu_{t,\boldsymbol{x}}\right),\quad\forall t\in\mathbb{R}^+.$$

We postpone the proof to Appendix B. It is mainly technical since equation (4.1) is linear.

We come back to the main concern of this section which consists in estimating $\mathcal{D}_1$ and prove the following estimate

Proposition 4.3. *Under the assumptions of Theorem 2.7, there exist two positive constants C & ε_0 such that for all $\varepsilon\leq\varepsilon_0$, holds the following estimate*

$$\mathcal{D}_1\;\leq\;C\,\left(W_2^2\left(\bar\nu^\varepsilon_{0,\boldsymbol{x}}\,,\,\bar\nu_{0,\boldsymbol{x}}\right)e^{-2\,b\,t}\;+\;e^{-2\,\rho_0^\varepsilon\,t/\varepsilon}\;+\;\varepsilon^2\right),\quad\forall(t,\boldsymbol{x})\in\mathbb{R}^+\times K\,.$$

Proof. We consider $\pi_0^\varepsilon\in\Pi\left(\nu^\varepsilon_{0,\boldsymbol{x}}\,,\,\mathcal{M}_{\rho_0^\varepsilon}\otimes\bar\nu_{0,\boldsymbol{x}}\right)$ and the associated solution π^ε to equation (4.1) given by Proposition 4.2. When the context is clear, we omit the dependence with respect to t. To evaluate the term $\mathcal{D}_1$, we introduce the following quantities

$$\begin{cases}\mathcal{A}(t)\;:=\;\displaystyle\int_{\mathbb{R}^4}|v-v'|^2\;\mathrm{d}\pi^\varepsilon(\boldsymbol{u},\boldsymbol{u}')\,,\\[2ex]\mathcal{B}(t)\;:=\;\displaystyle\int_{\mathbb{R}^4}|w-w'|^2\;\mathrm{d}\pi^\varepsilon(\boldsymbol{u},\boldsymbol{u}')\,.\end{cases}$$

Observing that $\mathcal{D}_1(t) \leq \varepsilon\,\mathcal{A}(t) + \mathcal{B}(t)$, this proof consists in showing that $\mathcal{A}(t)$ is of order ε whereas $\mathcal{B}(t)$ is of order ε^2.
We begin with $\mathcal{A}$ and multiply equation (4.1) by $|v-v'|^2/2$ and integrate by part with respect to $(\boldsymbol{u},\boldsymbol{u}')$, it yields

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{A} + \frac{\rho_0^\varepsilon}{\varepsilon}\mathcal{A} = \mathcal{A}_1\,,$$

where

$$\mathcal{A}_1 = \frac{1}{\sqrt{\varepsilon}}\int_{\mathbb{R}^4}(v-v')\left(N(\mathcal{V}^\varepsilon+\sqrt{\varepsilon}\,v) - w - \sqrt{\varepsilon}\,v\,\Psi *_r \rho_0^\varepsilon(\boldsymbol{x})\right)\mathrm{d}\pi^\varepsilon(\boldsymbol{u},\boldsymbol{u}')\,.$$

We apply Cauchy-Schwarz inequality, assumptions (2.1b), (2.3) & (2.4) and Corollary 3.2 to estimate $\mathcal{A}_1$, it yields

$$\mathcal{A}_1 \leq \frac{C}{\sqrt{\varepsilon}}\,\mathcal{A}(t)^{1/2}\left(\int_{\mathbb{R}^2}\left(1+|w|^2+|\sqrt{\varepsilon}\,v|^{2p}\right)\mathrm{d}\nu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\right)^{1/2}.$$

Then we perform the change of variable (1.6) in the latter integral and apply Propositions 3.1 & 3.3 and obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}(t) + \frac{\rho_0^\varepsilon}{\varepsilon}\,\mathcal{A}(t) \leq \frac{C}{\sqrt{\varepsilon}}\,\mathcal{A}(t)^{1/2}\,.$$

Dividing the former estimate by $\mathcal{A}(t)^{1/2}$ and applying Gronwall's lemma, this yields

$$\mathcal{A}(t)^{1/2} \leq \mathcal{A}(0)^{1/2}\exp\left(-\rho_0^\varepsilon\,t/\varepsilon\right) + C\,\varepsilon^{1/2}\,,$$

where the constant $C>0$ may depend on the lower bound m_* of the spatial distribution ρ_0^ε (see assumption (2.4)), m_p and $\bar{m}_p$. We point out that since we do not prepare the initial condition, $\mathcal{A}(0)$ may blow up as ε vanishes. Indeed, we have

$$\mathcal{A}(0) \leq 2\left(\int_{\mathbb{R}^2}|v|^2\,\mathrm{d}\nu^\varepsilon_{0,\boldsymbol{x}}(\boldsymbol{u}) + \frac{1}{\rho_0^\varepsilon(\boldsymbol{x})}\right).$$

Hence, operating the change of variable (1.6) in the latter integral and applying assumptions (2.4) & (2.5), the former estimate becomes

$$\mathcal{A}(0) \leq \frac{C}{\varepsilon}\,.$$

Therefore, for all ε less than 1, we obtain

$$\mathcal{A}(t) \leq \frac{C}{\varepsilon}\left(\exp\left(-2\,\rho_0^\varepsilon(\boldsymbol{x})\,t/\varepsilon\right) + \varepsilon^2\right).$$

We turn to $\mathcal{B}$ and prove that it is of order ε^2 using the previous estimate on $\mathcal{A}$. Indeed, we compute the derivative of $\mathcal{B}$ multiplying equation (4.1) by $|w-w'|^2/2$ and integrating by part with respect to $(\boldsymbol{u},\boldsymbol{u}')$, it yields

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{B}(t) = -\,b\,\mathcal{B}(t) + a\sqrt{\varepsilon}\,\left(\mathcal{B}_1(t)+\mathcal{B}_2(t)\right),$$

where

$$\begin{cases}\mathcal{B}_1(t) = \displaystyle\int_{\mathbb{R}^4}(v-v')\,(w-w')\,\mathrm{d}\pi^\varepsilon(\boldsymbol{u},\boldsymbol{u}')\,,\\[2ex] \mathcal{B}_2(t) = \displaystyle\int_{\mathbb{R}^4}v'\,(w-w')\,\mathrm{d}\pi^\varepsilon(\boldsymbol{u},\boldsymbol{u}')\,.\end{cases}$$

In order to estimate $\mathcal{B}_1$, we apply Cauchy-Schwarz inequality, use the estimate on $\mathcal{A}(t)$ and assumption (2.4) on ρ_0^ε, which yields

$$a\sqrt{\varepsilon}\,\mathcal{B}_1 \leq C\,\left(\exp\left(-m_*\,t/\varepsilon\right)+\varepsilon\right)\mathcal{B}(t)^{1/2}\,,$$

for some $C > 0$ and where m_* is given in (2.4). The next step consists in proving that $\mathcal{B}_2(t)$ is of order $\varepsilon^{3/2}$. We compute the derivative of $\mathcal{B}_2(t)$ multiplying equation (4.1) by $v'(w - w')$ and integrating by part with respect to $(\boldsymbol{u}, \boldsymbol{u}')$, it yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{B}_2(t) \,=\, -\left(\frac{\rho_0^\varepsilon}{\varepsilon} \,+\, b\right)\mathcal{B}_2(t) \,+\, a\,\sqrt{\varepsilon}\int_{\mathbb{R}^4} v\,v'\,\mathrm{d}\pi^\varepsilon(\boldsymbol{u},\boldsymbol{u}').$$

Then we apply Young's inequality, invert the change of variable (1.6) and apply Proposition 3.3. It yields

$$a\,\sqrt{\varepsilon}\int_{\mathbb{R}^4} |v\,v'|\;\mathrm{d}\pi^\varepsilon\,(\boldsymbol{u},\,\boldsymbol{u}') \,\leq\, C\,\left(\varepsilon^{-1/2}\,e^{-2\rho_0^\varepsilon t/\varepsilon} \,+\, \varepsilon^{1/2}\right),$$

where the positive constant C may depend on m_*, m_p and $\bar{m}_p$. Then, we multiply the equation on $\mathcal{B}_2$ by its sign and apply Gronwall's lemma. In the end, this leads to

$$|\mathcal{B}_2(t)| \leq |\mathcal{B}_2(0)|\exp\left(-\rho_0^\varepsilon t/\varepsilon\right) + C\left(\varepsilon^{1/2}\exp\left(-\rho_0^\varepsilon t/\varepsilon\right) + \varepsilon^{3/2}\right).$$

for any ε less than $m_*/(2b)$. Furthermore, applying Cauchy-Schwarz inequality in $\mathcal{B}_2(0)$ and using assumption (2.4), we obtain

$$|\mathcal{B}_2(0)| \,\leq\, C\mathcal{B}(0)^{1/2}.$$

Gathering the former computations and applying assumption (2.4), we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\,\mathcal{B}(t) + b\,\mathcal{B}(t) \,\leq\, C\,\left(e^{-m_*t/\varepsilon} \,+\, \varepsilon\,\right)\mathcal{B}(t)^{1/2} \,+\, C\,\left(\varepsilon^{1/2}\,\mathcal{B}(0)^{1/2}\,e^{-m_*t/\varepsilon} \,+\, \varepsilon\,e^{-m_*t/\varepsilon} \,+\, \varepsilon^2\,\right),$$

To estimate $\mathcal{B}(t)$, we will construct an upper-bound $\mathcal{B}_+(t)$ by considering the following ODE

$$\begin{cases} \dfrac{1}{2}\dfrac{\mathrm{d}}{\mathrm{d}t}\,\mathcal{B}_+(t) \,+\, b\,\mathcal{B}_+(t) \,=\, 2\,C\,\left(e^{(b\,-\,m_*/\varepsilon)\,t} \,+\, \varepsilon\right)\mathcal{B}_+(t)^{1/2}\,, \\[2ex] \mathcal{B}_+(0)^{1/2} \,=\, \mathcal{B}(0)^{1/2} \,+\, \dfrac{2\,C\,\varepsilon}{b}\,, \end{cases}$$

whose exact solution is given by

$$\mathcal{B}_+(t)^{1/2} \,=\, \mathcal{B}_+(0)^{1/2}\,e^{-bt} \,+\, \frac{2\,C\,\varepsilon}{b}\left(1 \,-\, e^{-bt}\right) \,+\, \frac{2\,C\,\varepsilon}{m_* \,-\, 2b\varepsilon}\,\left(e^{-bt} \,-\, e^{-(m_*/\varepsilon\,-\,b)\,t}\right).$$

We check that the following condition is fulfilled at all time t

$$\mathcal{B}_+(t)^{1/2} \,\geq\, \varepsilon^{1/2}\,\mathcal{B}(0)^{1/2}\,e^{-b\,t} \,+\, \varepsilon\,,$$

as long as $2\,C \,\geq\, b$ and $\varepsilon \,<\, \min\left\{\, m_*\,/\,(2b)\,,\, 1\,\right\}$. Making use of the latter inequality, we get

$$\varepsilon^{1/2}\,\mathcal{B}(0)^{1/2}\,e^{-m_*t/\varepsilon} \,+\, \varepsilon\,e^{-m_*t/\varepsilon} \,+\, \varepsilon^2 \,\leq\, \left(e^{(b\,-\,m_*/\varepsilon)t} \,+\, \varepsilon\right)\,\mathcal{B}_+(t)^{1/2}\,.$$

Therefore, defining $u(t) = \mathcal{B}_+(t) \,-\, \mathcal{B}(t)$, we check that the following inequality holds

$$\frac{1}{2}\frac{\mathrm{d}u}{\mathrm{d}t} \,+\, b\,u \,\geq\, C\,\frac{e^{-m_*t/\varepsilon} \,+\, \varepsilon}{\mathcal{B}(t)^{1/2} \,+\, \mathcal{B}_+(t)^{1/2}}\,u\,.$$

Since $\mathcal{B}$ is non negative and $\mathcal{B}_+(t)$ stays lower bounded by ε^2, we apply Gronwall's lemma to the latter estimate and noticing that $u(0) \,\geq\, 0$, it yields that $u(t)$ is non-negative. Then, we deduce

$$\mathcal{B}(t) \,\leq\, \mathcal{B}_+(t) \,\leq\, 2\,\mathcal{B}(0)\,e^{-2bt} \,+\, C\,\varepsilon^2\,,$$

for all $t \geq 0$, as long as $2\,C \,\geq\, b$ and $\varepsilon \,<\, \min\left\{\, m_*\,/\,(2b)\,,\, 1\,\right\}$.
Gathering the former computations, we obtain the following estimate for $\mathcal{D}_1$

$$\mathcal{D}_1 \,\leq\, C\,\left(\mathcal{B}(0)\,e^{-2bt} \,+\, e^{-2\,\rho_0^\varepsilon\,t/\varepsilon} \,+\, \varepsilon^2\right).$$

We conclude the proof taking the infimum over all π_0^ε. □

Estimate for $\mathcal{D}_2$. We mentioned that the change of variables (1.6) in $\mathcal{D}_2$ yields

$$\mathcal{D}_2 \;=\; |\mathcal{V}-\mathcal{V}^\varepsilon|^2 \,+\, |\mathcal{W}-\mathcal{W}^\varepsilon|^2\,.$$

Hence the proof consists in injecting the error estimate obtained in Proposition 3.5 in equations (1.3) & (1.5)

Proposition 4.4. *Under the assumptions of Theorem 2.7, there exist two positive constants C & ε_0 such that for all $\varepsilon \le \varepsilon_0$, holds the following estimate*

$$\mathcal{D}_2^{1/2} \;\le\; C\,\min\left(e^{C\,t}\,(\mathcal{E}_{\rm mac}\,+\,\varepsilon)\,,1\right), \quad \forall\,(t,\boldsymbol{x})\in\mathbb{R}^+\times K\,,$$

where $\mathcal{E}_{\rm mac}$ is defined in Theorem 2.7.

Proof. We omit the dependence with respect to $(t,\boldsymbol{x})$ when the context is clear and write $\|\cdot\|_\infty$ instead of $\|\cdot\|_{L^\infty(K)}$ in this proof. According to equations (1.3) & (1.5), we have

$$\left\{\begin{array}{ll}\dfrac{\rm d}{{\rm d}t}\,|\mathcal{V}-\mathcal{V}^\varepsilon| & =\,{\rm sgn}\,(\mathcal{V}-\mathcal{V}^\varepsilon)\left[N(\mathcal{V})-N(\mathcal{V}^\varepsilon)-\left(\mathcal{W}-\mathcal{W}^\varepsilon+\mathcal{L}_{\rho_0}\,(\mathcal{V})-\mathcal{L}_{\rho_0^\varepsilon}\,(\mathcal{V}^\varepsilon)\right)-\mathcal{E}(\mu^\varepsilon)\right]\,,\\ \dfrac{\rm d}{{\rm d}t}\,|\mathcal{W}-\mathcal{W}^\varepsilon| & =\,{\rm sgn}\,(\mathcal{W}-\mathcal{W}^\varepsilon)\,A_0\,(\,\mathcal{V}\,-\,\mathcal{V}^\varepsilon\,,\,\mathcal{W}\,-\,\mathcal{W}^\varepsilon\,)\,,\end{array}\right.$$

where ${\rm sgn}\,(v)=v\,/\,|\,v\,|$ for all $v\in\mathbb{R}^*$. According to Corollary 3.2 and Theorem 2.6, $(\mathcal{V},\mathcal{V}^\varepsilon)$ is uniformly bounded by a constant $R>0$ with respect to ε, time & space. Consequently, we have

$$(\mathcal{V}\,-\,\mathcal{V}^\varepsilon)\,(N(\mathcal{V})\,-\,N(\mathcal{V}^\varepsilon))\;\le\;C\,|\mathcal{V}\,-\,\mathcal{V}^\varepsilon|^2,$$

where C stands for the Lipschitz constant of N over the ball of radius R. We now estimate the contribution of the non-local terms. Using the linearity of $\mathcal{L}$ we split the term as follows

$$\mathcal{L}_{\rho_0}\,(\mathcal{V})\,-\,\mathcal{L}_{\rho_0^\varepsilon}\,(\mathcal{V}^\varepsilon)\;=\;\mathcal{L}_{\rho_0}\,(\mathcal{V}\,-\,\mathcal{V}^\varepsilon)\,+\,\mathcal{L}_{(\rho_0-\rho_0^\varepsilon)}\,(\mathcal{V}^\varepsilon)\,.$$

According to assumption (2.3) (we do not use the constraint $r>1$ here), and since $\mathcal{V}^\varepsilon$ is uniformly bounded (see Corollary 3.2), we have

$$-{\rm sgn}\;(\mathcal{V}\,-\,\mathcal{V}^\varepsilon)\;\mathcal{L}_{(\rho_0-\rho_0^\varepsilon)}\,(\mathcal{V}^\varepsilon)\;\le\;C\;\|\rho_0\,-\,\rho_0^\varepsilon\|_\infty\,.$$

Furthermore, according to assumptions (2.3) & (2.4) (we do not use the constraint $r>1$ here), we obtain

$$-{\rm sgn}\;(\mathcal{V}\,-\,\mathcal{V}^\varepsilon)\;\mathcal{L}_{\rho_0}\,(\mathcal{V}\,-\,\mathcal{V}^\varepsilon)\;\le\;C\;|\mathcal{V}-\mathcal{V}^\varepsilon|\,+\,|\Psi|\,*_r\,(|\mathcal{V}-\mathcal{V}^\varepsilon|\,\rho_0)\,.$$

We estimate the non-local term using assumptions (2.3) & (2.4) (we do not use the constraint $r>1$ here). It yields

$$|\Psi|\,*_r\,(|\mathcal{V}-\mathcal{V}^\varepsilon|\,\rho_0)\;\le\;C\;\|\mathcal{V}-\mathcal{V}^\varepsilon\|_\infty\,,$$

for some constant C only depending on m_* and Ψ. Then we gather the former computations and replace $\mathcal{E}(\mu^\varepsilon)$ by the bound obtained in Proposition 3.5. It yields

$$\frac{\rm d}{{\rm d}t}\,|\mathcal{U}\,-\,\mathcal{U}^\varepsilon\,|\;\le\;C\left(\|\,\mathcal{U}\,-\,\mathcal{U}^\varepsilon\,\|_\infty\,+\,\varepsilon\,+\,e^{-2\,m_*\,t/\varepsilon}\,+\,\|\rho_0\,-\,\rho_0^\varepsilon\|_\infty\right)\,,$$

for some positive constant C which may depend on m_p, $\overline{m}_p$ and m_* but not on ε and $(t,\boldsymbol{x})$. Integrating the latter inequality between 0 and t and taking the supremum over all $\boldsymbol{x}$ in K, we end up with the following inequality

$$\|\,\mathcal{U}(t)\,-\,\mathcal{U}^\varepsilon(t)\,\|_\infty\;\le\;\|\,\mathcal{U}_0\,-\,\mathcal{U}_0^\varepsilon\,\|_\infty\,+\,C\int_0^t\|\mathcal{U}(s)\,-\,\mathcal{U}^\varepsilon(s)\|_\infty\,+\,\varepsilon\,+\,e^{-m_*\,s/\varepsilon}\,+\,\|\rho_0\,-\,\rho_0^\varepsilon\|_\infty\,{\rm d}s\,.$$

We conclude the proof applying Gronwall's lemma and using that $\mathcal{U}$ and $\mathcal{U}^\varepsilon$ are uniformly bounded according to Theorem 2.6 and Corollary 3.2. □

Estimate for $\mathcal{D}_3$. We now turn to the last term $\mathcal{D}_3$. This section is only technical and we prove that $\mathcal{D}_3$ is negligible in comparison to $\mathcal{D}_1$ and $\mathcal{D}_2$. Indeed, changing variables and applying the gluing Lemma (see [24,] for more details), we obtain

$$\mathcal{D}_3 \,\leq\, \frac{2\,\varepsilon}{\rho_0^\varepsilon}\, W_2^2\,(\mathcal{M}_1, \mathcal{M}_1) \,+\, 2\,\varepsilon\,\left(\frac{1}{\sqrt{\rho_0^\varepsilon}} \,-\, \frac{1}{\sqrt{\rho_0}}\right)^2 .$$

Hence, according to assumption (2.4), we obtain

$$\mathcal{D}_3 \,\leq\, \frac{\varepsilon}{2\,m_*^3}\,\|\rho_0 \,-\, \rho_0^\varepsilon\|_{L^\infty(K)}^2 .$$

Therefore applying Propositions 4.3 and 4.4 together with the latter estimate on $\mathcal{D}_3$, we have proven Theorem 2.7.

5. Conclusion & Perspectives

In this paper, we have characterized the blow-up profile (Gaussian distribution) of the voltage distribution in the regime of strong & short-range coupling between neurons and have computed the limiting distribution for the adaptation variable as well. Our result should be interpreted as the first order expansion of the network's distribution as ε vanishes. Indeed, it allows to improve on former convergence result in the sense that we gain an order $\sqrt{\varepsilon}$ in our convergence rate. On top of that, we benefit from the first order expansion to derive an asymptotic equivalent of the distribution at order zero (see Corollary 2.9).

Let us also mention a few natural questions that arise from this work. The natural continuation of this article consists in obtaining a strong convergence result towards the concentration profile (see [2]). It appears that the relative entropy approach we developed in Appendix A may adapt to the analysis of the asymptotic $\varepsilon \ll 1$. To end with, we point out that it should be possible to derive the next corrective terms in the expansion of the macroscopic quantities thanks to this work. Indeed, based on our result with a slight improvement, it may be possible to prove

$$(\,\mathcal{V}^\varepsilon,\, \mathcal{W}^\varepsilon\,) \underset{\epsilon\to 0}{=} \left(\,\overline{\mathcal{V}}^\varepsilon,\, \overline{\mathcal{W}}^\varepsilon\,\right) \,+\, O\left(\varepsilon^{3/2}\right),$$

where the limiting macroscopic system $\left(\,\overline{\mathcal{V}}^\varepsilon,\, \overline{\mathcal{W}}^\varepsilon\,\right)$ solves the following system

$$\begin{cases} \partial_t\,\overline{\mathcal{V}}^\varepsilon \,=\, N\left(\,\overline{\mathcal{V}}^\varepsilon\,\right) \,-\, \overline{\mathcal{W}}^\varepsilon \,-\, \mathcal{L}_{\rho_0}[\,\overline{\mathcal{V}}^\varepsilon\,] \,+\, \dfrac{\varepsilon}{2}\,\rho_0\,N''\left(\,\overline{\mathcal{V}}^\varepsilon\,\right), \\ \partial_t\,\overline{\mathcal{W}}^\varepsilon \,=\, A\left(\overline{\mathcal{V}}^\varepsilon\,,\, \overline{\mathcal{W}}^\varepsilon\right). \end{cases}$$

The corrective term adds a dependence with respect to the spatial distribution of neurons and it might add some complexity to the dynamics of the limiting macroscopic system.

Acknowledgment

Francis Filbet gratefully acknowledges the support of ANITI (Artificial and Natural Intelligence Toulouse Institute).

Appendix A. Proof of Theorem 2.3

In this section, we prove existence and uniqueness of solution for equation (1.2) in the sense of Definition 2.2. The main difficulty is to prove uniqueness and continuity of the solution

$$\mu^\varepsilon \in \mathscr{C}^0\left(\mathbb{R}^+ \times K\,,\, L^1\left(\mathbb{R}^2\right)\right).$$

To this aim, we first establish *a priori* estimates on the solution by propagating exponential moments in $\boldsymbol{u} \in \mathbb{R}^2$. Then the key point of the proof is to define a modified relative entropy $H_\alpha\,[\,\mu\,|\,\nu\,]$, given for any α in $]0,1[$ as

$$H_\alpha\,[\,\mu\,|\,\nu\,] \,=\, \int_{\mathbb{R}^2} \mu\,\ln\left(\frac{\mu}{\alpha\mu + (1-\alpha)\nu}\right)\mathrm{d}\boldsymbol{u}\,, \tag{A.1}$$

for two non-negative functions $\mu, \nu \in L^1(\mathbb{R}^2)$ with mass equal to one. Relative entropy has several technical advantages in comparison to the L^1 norm. On the one hand, it makes explicit the dissipation due to the Laplace operator, which we refer as the Fisher information in what follows. On the other hand, the latter modified relative entropy, where the denominator is a convex combination of μ & ν is well defined for all positive functions μ and ν with mass 1, with no further condition on the domain on which they vanish. Indeed, under these conditions, holds the following inequality

$$H_\alpha[\mu \,|\, \nu] \leq -\ln(\alpha) < +\infty.$$

We even prove the following stronger result.

Lemma A.1. *For any two non-negative functions $\mu, \nu \in L^1(\mathbb{R}^2)$ with integral equal to one, the following estimate holds*

$$\frac{(1-\alpha)^2}{2} \|\mu - \nu\|^2_{L^1(\mathbb{R}^2)} \leq H_\alpha[\mu \,|\, \nu] \leq \frac{1-\alpha}{\alpha} \|\mu - \nu\|_{L^1(\mathbb{R}^2)} . \tag{A.2}$$

Proof. The first inequality is a straightforward consequence of the Csizár-Kullback inequality. For the second inequality, we write

$$\kappa^\alpha = \alpha\mu + (1-\alpha)\nu$$

and we notice that $|\mu/\kappa^\alpha| \leq \alpha^{-1}$. Then we apply the convex inequality $\ln(1+x) \leq x$ to the following relation

$$H_\alpha[\mu \,|\, \nu] = \int_{\mathbb{R}^2} \mu \ln\left(1 + \frac{\mu - \kappa^\alpha}{\kappa^\alpha}\right) \mathrm{d}\boldsymbol{u}\,.$$

□

From this modified relative entropy functional, we prove the continuity of the solution with respect to both the time and the spatial variable in the functional space L^1. Now, we fix $\alpha = 1/2$ and denote by $I_{1/2}$ the associated relative Fisher information

$$I_{1/2}[\mu \,|\, \nu] := \int_{\mathbb{R}^2} \left|\partial_v \ln\left(\frac{2\mu}{\nu+\mu}\right)\right|^2 \mu \,\mathrm{d}\boldsymbol{u}\,. \tag{A.3}$$

To deal with existence issues, we provide an entropy estimate

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

and follow a classical regularization argument.

A.1. **A priori estimates.** In this section, we suppose that we have a smooth solution $\mu^\varepsilon_{t,\boldsymbol{x}}$ to equation (1.2) and provide exponential moment, entropy and continuity estimates for $\mu^\varepsilon_{t,\boldsymbol{x}}$.

We first define $J[\mu^\varepsilon_{t,\boldsymbol{x}}]$ as

$$J[\mu^\varepsilon_{t,\boldsymbol{x}}] := \int_{\mathbb{R}^2} e^{|\boldsymbol{u}|^2/2}\, \mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,,$$

and prove the following result.

Proposition A.2 (Exponential moments)**.** *For any $\varepsilon > 0$, suppose that assumptions (2.1a) and (2.2)-(2.4) are fulfilled whereas the initial condition μ^ε_0 verifies the first condition of (2.7). Then, for any solution μ^ε to equation (1.2), there exists a positive constant C that may depend on ε such that*

$$J[\mu^\varepsilon_{t,\boldsymbol{x}}] \leq C\,, \quad \forall\,(t,\boldsymbol{x}) \in \mathbb{R}^+ \times K\,.$$

Proof. We multiply equation (1.2) by $e^{|\boldsymbol{u}|^2/2}$ and integrate with respect to $\boldsymbol{u} \in \mathbb{R}^2$, after an integration by part, it yields

$$\frac{\mathrm{d}}{\mathrm{d}t} J[\mu^\varepsilon_{t,\boldsymbol{x}}] = \mathcal{J}\,,$$

where $\mathcal{J}$ is split into $\mathcal{J} = \mathcal{J}_1 + \mathcal{J}_2 + \mathcal{J}_3$, with

$$\begin{cases} \mathcal{J}_1 &= \left(\dfrac{1}{\varepsilon}\,(\rho_0^\varepsilon\,\mathcal{V}^\varepsilon) \,+\, \Psi *_r (\rho_0^\varepsilon\,\mathcal{V}^\varepsilon)\right)(t,\boldsymbol{x}) \displaystyle\int_{\mathbb{R}^2} v\, e^{|\boldsymbol{u}|^2/2}\, \mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,, \\ \mathcal{J}_2 &= \displaystyle\int_{\mathbb{R}^2} \left(w\,A(\boldsymbol{u}) \,-\, v\,w \,-\, v^2\left(\frac{\rho_0^\varepsilon(\boldsymbol{x})}{\varepsilon} + \Psi *_r \rho_0^\varepsilon(\boldsymbol{x}) - 1\right) + 1\right) e^{|\boldsymbol{u}|^2/2}\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,, \\ \mathcal{J}_3 &= \displaystyle\int_{\mathbb{R}^2} v\,N(v)\,e^{|\boldsymbol{u}|^2/2}\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,. \end{cases}$$

The proof follows the same lines as the one given in Proposition 3.1 on the moment estimates except that here the non-local term $\mathcal{J}_1$ can be roughly estimated. Indeed, applying Corollary 3.2 and assumptions (2.3) & (2.4), we first obtain

$$\mathcal{J}_1 \,\le\, C \int_{\mathbb{R}^2} |v|\, e^{|\boldsymbol{u}|^2/2}\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,,$$

for some positive constant $C > 0$ that may depend on ε. Then we estimate $\mathcal{J}_2$ and $\mathcal{J}_3$ following the computations in the proof of Proposition 3.1 and taking advantage of the confinement property (2.1a) on N. In the end, we get

$$\frac{\mathrm{d}}{\mathrm{d}t} J[\mu^\varepsilon_{t,\boldsymbol{x}}] \,\le\, \int_{\mathbb{R}^2} \left(\left(C \,-\, \omega^-(v)\right)|v|^2 \,-\, \alpha|w|^2 + C\right) e^{|\boldsymbol{u}|^2/2}\,\mu^\varepsilon_{t,\boldsymbol{x}}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,,$$

where C and α are two positive constants that may depend on ε and where ω^- is defined as

$$\omega^-(v) \,=\, \left(\omega(v)\,\mathbb{1}_{|v|\ge 1}\right)^-\,,$$

where ω is given in (2.1a). Finally, according to (2.1a), we have

$$\lim_{|\boldsymbol{u}|\to+\infty} \left(C \,-\, \omega^-(v)\right)|v|^2 \,-\, \alpha|w|^2 + C \,=\, -\infty\,,$$

hence it gives

$$\frac{\mathrm{d}}{\mathrm{d}t} J[\mu^\varepsilon_{t,\boldsymbol{x}}] \,\le\, C \,-\, \frac{1}{C}\, J[\mu^\varepsilon_{t,\boldsymbol{x}}]\,.$$

We conclude the proof applying Gronwall's lemma and using the first assumption in (2.7). □

Then we provide an entropy estimate of the solution $\mu^\varepsilon_{t,\boldsymbol{x}}$.

Proposition A.3 (Entropy estimates)**.** *For any $\varepsilon > 0$, suppose that assumptions (2.1a)-(2.1b) and (2.2)-(2.4) are fulfilled whereas the initial condition μ_0^ε verifies the second condition in (2.7). Then, for any solution μ^ε to equation (1.2), there exists a positive constant $C > 0$ that may depend on ε such that*

$$H[\mu^\varepsilon_{t,\boldsymbol{x}}] \,\le\, H[\mu^\varepsilon_{0,\boldsymbol{x}}] \,+\, C\,t\,, \quad \forall\,(t,\boldsymbol{x}) \,\in\, \mathbb{R}^+ \times K\,.$$

Proof. We multiply equation (1.2) by $\ln\left(\mu^\varepsilon_{t,\boldsymbol{x}}\right)$ and integrate with respect to $\boldsymbol{u} \in \mathbb{R}^2$. Then we use conservation of mass for (1.2) and integrate by part. It yields

$$\frac{\mathrm{d}}{\mathrm{d}t} H[\mu^\varepsilon_{t,\boldsymbol{x}}] \,+\, I\left[\mu^\varepsilon_{t,\,\boldsymbol{x}}\right] \,=\, \frac{\rho_0^\varepsilon}{\varepsilon} \,+\, \Psi *_r \rho_0^\varepsilon \,+\, b \,+\, \int_{\mathbb{R}^2} N(v)\,\partial_v\left(\ln \mu^\varepsilon_{t,\boldsymbol{x}}\right)\,\mu^\varepsilon_{t,\boldsymbol{x}}\,\mathrm{d}\boldsymbol{u}\,,$$

where the Fisher information I is given by

$$I\left[\mu^\varepsilon_{t,\,\boldsymbol{x}}\right] \,=\, \int_{\mathbb{R}^2} \left|\partial_v \ln\left(\mu^\varepsilon_{t,\,\boldsymbol{x}}\right)\right|^2 \mu^\varepsilon_{t,\,\boldsymbol{x}}\,\mathrm{d}\boldsymbol{u}\,.$$

According to Young's inequality, we have

$$\int_{\mathbb{R}^2} N(v)\,\partial_v\left(\ln \mu^\varepsilon_{t,\boldsymbol{x}}\right)\,\mu^\varepsilon_{t,\boldsymbol{x}}\,\mathrm{d}\boldsymbol{u} \,\le\, \frac{C}{\eta}\int_{\mathbb{R}^2} |N(v)|^2\,\mu^\varepsilon_{t,\boldsymbol{x}}\,\mathrm{d}\boldsymbol{u} \,+\, \eta\, I\left[\mu^\varepsilon_{t,\boldsymbol{x}}\right]\,,$$

for all positive η. We choose $\eta = 1/2$, apply Proposition 3.1, and assumptions (2.1b), (2.3) & (2.4). This yields

$$\frac{\mathrm{d}}{\mathrm{d}t} H[\mu^\varepsilon_{t,\boldsymbol{x}}] \leq C\,,$$

for some $C > 0$ that may depend on ε. Then we integrate the former inequality with respect to time and obtain the result. □

We now turn to continuity estimates with respect to the spatial variable and close the estimates in L^1 making use of the modified relative entropy $H_\alpha[\,\mu\,|\,\nu\,]$ given in (A.1).

Proposition A.4 (Continuity with respect to $\boldsymbol{x}$)**.** *For any $\varepsilon > 0$, suppose that assumption (2.1a) and (2.2)-(2.4) are fulfilled whereas the initial condition μ^ε_0 verifies the first condition of (2.7). Then, for any solution μ^ε to equation (1.2), the following estimate holds*

$$\left\|\mu^\varepsilon_{t,\boldsymbol{x}} - \mu^\varepsilon_{t,\boldsymbol{y}}\right\|_{L^1(\mathbb{R}^2)} \leq C\, e^{Ct}\, \gamma(\boldsymbol{x},\boldsymbol{y}), \quad \forall\,(\boldsymbol{x}\,,\,\boldsymbol{y}) \in K^2\,,$$

for some positive constant C which may depend on ε and where γ is the following continuous function

$$\gamma(\boldsymbol{x},\boldsymbol{y}) = \left\|\mu^\varepsilon_{0,\boldsymbol{x}} - \mu^\varepsilon_{0,\boldsymbol{y}}\right\|^{1/2}_{L^1(\mathbb{R}^2)} + |\rho^\varepsilon_0(\boldsymbol{x}) - \rho^\varepsilon_0(\boldsymbol{y})| + \|\Psi(\boldsymbol{x},\cdot) - \Psi(\boldsymbol{y},\cdot)\|_{L^1(K)}\,.$$

Proof. Our strategy consists in estimating $H_{1/2}$ instead of the L^1 error in order to take advantage of the entropy dissipation, then Lemma A.1 will ensure the continuity in L^1.

We choose some $\boldsymbol{x}$ and $\boldsymbol{y}$ in K all along this proof and consider $\kappa = \left(\mu^\varepsilon_{t,\boldsymbol{x}} + \mu^\varepsilon_{t,\boldsymbol{y}}\right)/2$. It satisfies the following equation

$$\partial_t\,\kappa + \partial_v\left(\,(N(v) - w)\,\kappa\,\right) + \partial_w\left(\,A(\boldsymbol{u})\,\kappa\,\right) - \partial^2_v\,\kappa - \frac{1}{2}\partial_v\left(\mathcal{N}_{\boldsymbol{x}}\,\mu^\varepsilon_{t,\boldsymbol{x}} + \mathcal{N}_{\boldsymbol{y}}\,\mu^\varepsilon_{t,\boldsymbol{y}}\right) = 0\,,$$

where $\mathcal{N}$ gathers the non-linear terms and is given by

$$\mathcal{N}_{\boldsymbol{x}}(t,v) := \frac{\rho^\varepsilon_0(\boldsymbol{x})}{\varepsilon}(v - \mathcal{V}^\varepsilon(t,\,\boldsymbol{x})) + \left(\Psi *_r \rho^\varepsilon_0(\boldsymbol{x})\,v - \Psi *_r \left(\rho^\varepsilon_0 \mathcal{V}^\varepsilon\right)(t,\,\boldsymbol{x})\right)\,.$$

We compute the time derivative of $H_{1/2}\left[\,\mu^\varepsilon_{t,\,\boldsymbol{x}}\,|\,\mu^\varepsilon_{t,\,\boldsymbol{y}}\,\right]$ using the former equation, equation (1.2) and conservation of mass for (1.2). After an integration by part, all the terms associated to linear transport cancel and we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} H_{1/2}\left[\,\mu^\varepsilon_{t,\,\boldsymbol{x}}\,|\,\mu^\varepsilon_{t,\,\boldsymbol{y}}\,\right] + I_{1/2}\left[\,\mu^\varepsilon_{t,\,\boldsymbol{x}}\,|\,\mu^\varepsilon_{t,\,\boldsymbol{y}}\,\right] = \mathcal{H}_1\,,$$

where $I_{1/2}[\,\mu\,|\,\nu\,]$ is given by (A.3) and corresponds to the dissipation due to the second order term whereas $\mathcal{H}_1$ is given by

$$\mathcal{H}_1 = -\int_{\mathbb{R}^2} \partial_v\left(\frac{\mu^\varepsilon_{t,\boldsymbol{x}}}{\kappa}\right)\left(\mathcal{N}_{\boldsymbol{x}}\,\kappa - \frac{1}{2}\left(\mathcal{N}_{\boldsymbol{x}}\,\mu^\varepsilon_{t,\boldsymbol{x}} + \mathcal{N}_{\boldsymbol{y}}\,\mu^\varepsilon_{t,\boldsymbol{y}}\right)\right)\mathrm{d}\boldsymbol{u}\,.$$

Exact computations yield

$$\mathcal{H}_1 = -\frac{1}{2}\int_{\mathbb{R}^2} \partial_v\left(\ln\left(\frac{\mu^\varepsilon_{t,\boldsymbol{x}}}{\kappa}\right)\right)(\mathcal{N}_{\boldsymbol{x}} - \mathcal{N}_{\boldsymbol{y}})\,\frac{\mu^\varepsilon_{t,\boldsymbol{y}}}{\kappa}\,\mu^\varepsilon_{t,\boldsymbol{x}}\,\mathrm{d}\boldsymbol{u}\,.$$

Then, from Proposition A.2, we get that $\mu^\varepsilon_{t,\boldsymbol{x}}$ has exponential moments, hence we can apply Corollary 2.4 in [3], which yields

$$|\,\mathcal{V}^\varepsilon(t,\boldsymbol{x}) - \mathcal{V}^\varepsilon(t,\boldsymbol{y})\,| \leq 2\,W_1\left(\mu^\varepsilon_{t,\boldsymbol{x}},\kappa\right) \leq C\,H_{1/2}\left[\mu^\varepsilon_{t,\boldsymbol{x}}\,|\,\mu^\varepsilon_{t,\boldsymbol{y}}\right]^{1/2}\,,$$

for some positive constant $C > 0$ which may depend on ε. Consequently, using assumption (2.4) and Corollary 3.2, we obtain the following bound for $\mathcal{H}_1$

$$|\mathcal{H}_1| \leq C\left(\beta(\boldsymbol{x},\boldsymbol{y}) + H_{1/2}\left[\mu^\varepsilon_{t,\boldsymbol{x}}\,|\,\mu^\varepsilon_{t,\boldsymbol{y}}\right]^{1/2}\right)\int_{\mathbb{R}^2}\left|\partial_v \ln\left(\frac{\mu^\varepsilon_{t,\boldsymbol{x}}}{\kappa}\right)\right|(1 + |v|)\,\frac{\mu^\varepsilon_{t,\boldsymbol{y}}}{\kappa}\,\mu^\varepsilon_{t,\boldsymbol{x}}\,\mathrm{d}\boldsymbol{u}\,,$$

where C is a positive constant that may depend on ε and where β is given by

$$\beta(\boldsymbol{x},\boldsymbol{y}) \,=\, |\rho_0^\varepsilon(\boldsymbol{x}) \,-\, \rho_0^\varepsilon(\boldsymbol{y})| \,+\, \|\Psi(\boldsymbol{x},\,\cdot) \,-\, \Psi(\boldsymbol{y},\,\cdot)\|_{L^1(K)}\,.$$

Then we apply Young's inequality, Proposition 3.1 and the bound $\left|\mu_{t,\boldsymbol{y}}^\varepsilon \,/\, \kappa\right| \,\leq\, 2$. It yields

$$|\mathcal{H}_1| \,\leq\, C\,\eta^{-1}\left(\beta(\boldsymbol{x},\boldsymbol{y})^2 \,+\, H_{1/2}\left[\mu_{t,\boldsymbol{x}}^\varepsilon \,|\, \mu_{t,\boldsymbol{y}}^\varepsilon\right]\right) \,+\, \eta\, I_{1/2}\left[\mu_{t,\boldsymbol{x}}^\varepsilon \,|\, \mu_{t,\boldsymbol{y}}^\varepsilon\right]\,,$$

for all positive η and where C may depend on ε. Therefore, taking $\eta \,=\, 1/2$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\, H_{1/2}\left[\mu_{t,\boldsymbol{x}}^\varepsilon \,|\, \mu_{t,\boldsymbol{y}}^\varepsilon\right] \,+\, \frac{1}{2}\, I_{1/2}\left[\mu_{t,\boldsymbol{x}}^\varepsilon \,|\, \mu_{t,\boldsymbol{y}}^\varepsilon\right] \,\leq\, C\left(\beta(\boldsymbol{x},\boldsymbol{y})^2 \,+\, H_{1/2}\left[\mu_{t,\boldsymbol{x}}^\varepsilon \,|\, \mu_{t,\boldsymbol{y}}^\varepsilon\right]\right)\,.$$

We apply Gronwall's lemma and Lemma A.1. It yields

$$\left\|\mu_{t,\boldsymbol{x}}^\varepsilon \,-\, \mu_{t,\boldsymbol{y}}^\varepsilon\right\|_{L^1(\mathbb{R}^2)} \,\leq\, C\, e^{Ct}\left(\left\|\mu_{0,\boldsymbol{x}}^\varepsilon \,-\, \mu_{0,\boldsymbol{y}}^\varepsilon\right\|_{L^1(\mathbb{R}^2)}^{1/2} \,+\, \beta(\boldsymbol{x},\boldsymbol{y})\right)\,.$$

□

We now turn to continuity with respect to the time variable and split the proof into two steps. First we prove continuity at time $t = 0$ and then deduce continuity for all time $t > 0$.

Proposition A.5 (Continuity at time $t = 0$)**.** *Under the assumptions of Theorem 2.3, the following estimate holds*

$$\sup_{\boldsymbol{x}\in K}\, \left\|\, \mu_{t,\boldsymbol{x}}^\varepsilon \,-\, \mu_{0,\boldsymbol{x}}^\varepsilon \,\right\|_{L^1(\mathbb{R}^2)} \,\leq\, C\,\sqrt{t}\,,\quad \forall\, t\in\mathbb{R}_+\,,$$

for some positive constant C which may depend on ε.

Remark A.6. *It is the only time that we use assumption (2.8).*

Proof. All along this proof, we set $\kappa \,=\, \left(\mu_{t,\boldsymbol{x}}^\varepsilon \,+\, \mu_{0,\boldsymbol{x}}^\varepsilon\right)/2$. In order to simplify notation, we also define B^ε as follows

$$B^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) \,=\, N(v) \,-\, w \,-\, \mathcal{K}_\Psi[\rho_0^\varepsilon\,\mu^\varepsilon]\,(t,\boldsymbol{x},v) \,-\, \frac{1}{\varepsilon}\,\rho_0^\varepsilon\,(v - \mathcal{V}^\varepsilon)\,,$$

and we point out that according to Corollary 3.2, assumptions (2.1b) on N, (2.3) & (2.4), there exists a positive constant $C > 0$ that may depend on ε such that

$$|B^\varepsilon(t,\boldsymbol{x},\boldsymbol{u})| \,\leq\, C\,(\,|\boldsymbol{u}|^p \,+\, 1\,)\,,\quad \forall\,(t,\boldsymbol{x},\boldsymbol{u})\in\mathbb{R}_+\times K\times\mathbb{R}^2\,.$$

We compute the derivative of $H_{1/2}\left[\,\mu_{0,\boldsymbol{x}}^\varepsilon \,|\, \mu_{t,\boldsymbol{x}}^\varepsilon\,\right]$ using equation (1.2). It yields

$$\frac{\mathrm{d}}{\mathrm{d}t} H_{1/2}\left[\mu_{0,\boldsymbol{x}}^\varepsilon \,|\, \mu_{t,\boldsymbol{x}}^\varepsilon\right] \,=\, \mathcal{H}_1 \,+\, \mathcal{H}_2 \,+\, \mathcal{H}_3\,,$$

where

$$\left\{\begin{aligned}
\mathcal{H}_1 \,&=\, -\frac{1}{2}\int_{\mathbb{R}^2} \partial_v^2\, \mu_{t,\boldsymbol{x}}^\varepsilon\, \frac{\mu_{0,\boldsymbol{x}}^\varepsilon}{\kappa}\,\mathrm{d}\boldsymbol{u}\,,\\
\mathcal{H}_2 \,&=\, \frac{1}{2}\int_{\mathbb{R}^2} \partial_v\left[B^\varepsilon\,\mu_{t,\boldsymbol{x}}^\varepsilon\right]\frac{\mu_{0,\boldsymbol{x}}^\varepsilon}{\kappa}\,\mathrm{d}\boldsymbol{u}\,,\\
\mathcal{H}_3 \,&=\, \frac{1}{2}\int_{\mathbb{R}^2} \partial_w\left[A\,\mu_{t,\boldsymbol{x}}^\varepsilon\right]\frac{\mu_{0,\boldsymbol{x}}^\varepsilon}{\kappa}\,\mathrm{d}\boldsymbol{u}\,.
\end{aligned}\right.$$

We start with $\mathcal{H}_1$. First, we integrate by part and rewrite the term as follows

$$\mathcal{H}_1 \,=\, -\,I_{1/2}\left[\mu_{0,\boldsymbol{x}}^\varepsilon \,|\, \mu_{t,\boldsymbol{x}}^\varepsilon\right] \,+\, \int_{\mathbb{R}^2} \partial_v\left(\ln\frac{\mu_{0,\boldsymbol{x}}^\varepsilon}{\kappa}\right)\left(1 \,-\, \frac{1}{2}\frac{\mu_{0,\boldsymbol{x}}^\varepsilon}{\kappa}\right)\partial_v\left(\ln\mu_{0,\boldsymbol{x}}^\varepsilon\right)\mu_{0,\boldsymbol{x}}^\varepsilon\,\mathrm{d}\boldsymbol{u}\,.$$

Second, we apply the inequality $|\mu_0^\varepsilon \,/\, \kappa| \,\leq\, 2$, Young's inequality and assumption (2.8) on μ_0^ε. It yields for all positive η

$$\mathcal{H}_1 \,\leq\, -\left(1 - \frac{\eta}{2}\right)\, I_{1/2}\left[\mu_{0,\boldsymbol{x}}^\varepsilon \,|\, \mu_{t,\boldsymbol{x}}^\varepsilon\right] \,+\, 8\,\frac{m^\varepsilon}{\eta},$$

We turn to $\mathcal{H}_2$. After an integration by part, it rewrites as follows

$$\mathcal{H}_2 \,=\, -\frac{1}{2}\int_{\mathbb{R}^2} B^\varepsilon\,\partial_v\left(\ln\frac{\mu^\varepsilon_{0,\boldsymbol{x}}}{\kappa}\right)\frac{\mu^\varepsilon_{t,\boldsymbol{x}}}{\kappa}\,\mu^\varepsilon_{0,\boldsymbol{x}}\,\mathrm{d}\boldsymbol{u}\,.$$

Then we apply the inequality $|\mu^\varepsilon\,/\,\kappa| \,\leq\, 2$, Young's inequality, the bound on B^ε and the first assumption in (2.7). We obtain for some positive constant C and for all positive η

$$\mathcal{H}_2 \,\leq\, \eta\, I_{1/2}\left[\mu^\varepsilon_{0,\boldsymbol{x}}\,|\,\mu^\varepsilon_{t,\boldsymbol{x}}\right] \,+\, \frac{C}{\eta}\,.$$

To end with, we estimate $\mathcal{H}_3$. This term is a little trickier to estimate since we do not have dissipation with respect to the adaptation variable. We rewrite $\mathcal{H}_3 \,=\, \mathcal{H}_{31} \,+\, \mathcal{H}_{32}$, where

$$\begin{cases} \mathcal{H}_{31} \,=\, \displaystyle\int_{\mathbb{R}^2} \partial_w\left(A(\boldsymbol{u})\right)\,\psi\left(\frac{\mu^\varepsilon_{0,\boldsymbol{x}}}{\kappa}\right)\frac{\mu^\varepsilon_{0,\boldsymbol{x}}}{\kappa}\,\kappa\,\mathrm{d}\boldsymbol{u}\,, \\ \mathcal{H}_{32} \,=\, -\displaystyle\int_{\mathbb{R}^2} A(\boldsymbol{u})\,\psi\left(\frac{\mu^\varepsilon_{0,\boldsymbol{x}}}{\kappa}\right)\,\partial_w\left(\ln\mu^\varepsilon_{0,\boldsymbol{x}}\right)\,\mu^\varepsilon_{0,\boldsymbol{x}}\,\mathrm{d}\boldsymbol{u}\,, \end{cases}$$

where $\psi(x) \,=\, \ln(x) - x/2$. We notice that $\mu^\varepsilon_{0,\boldsymbol{x}}/\kappa$ lies in $[0,\,2]$ and that $(x \mapsto x\,\psi(x))$ is bounded on $[0,\,2]$. Hence, using conservation of mass for equation (1.2), we obtain

$$\mathcal{H}_{31} \,\leq\, b\,\sup_{x\in[0,2]}\left|x\,\psi(x)\right|\,.$$

We turn to $\mathcal{H}_{32}$. We apply Young's inequality and use assumption (2.8) on μ^ε_0. It yields

$$\mathcal{H}_{32} \,=\, \frac{1}{2}\int_{\mathbb{R}^2}|A(\boldsymbol{u})|^2\left|\psi\left(\frac{\mu^\varepsilon_{0,\boldsymbol{x}}}{\kappa}\right)\right|^2\frac{\mu^\varepsilon_{0,\boldsymbol{x}}}{\kappa}\,\kappa\,\mathrm{d}\boldsymbol{u} \,+\, 2\,m^\varepsilon\,.$$

We notice that $\left(x \mapsto x\,|\psi(x)|^2\right)$ is bounded on $[0,2]$ an that $\mu^\varepsilon_{0,\boldsymbol{x}}/\kappa$ lies in $[0,\,2]$. Furthermore, we apply Proposition 3.1 and assumption (2.5). It yields

$$\mathcal{H}_3 \,\leq\, C\,.$$

Gathering former computations and taking η small enough, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}H_{1/2}\left[\mu^\varepsilon_{0,\boldsymbol{x}}\,|\,\mu^\varepsilon_{t,\boldsymbol{x}}\right] \,+\, \frac{1}{2}\,I_{1/2}\left[\mu^\varepsilon_{0,\boldsymbol{x}}\,|\,\mu^\varepsilon_{t,\boldsymbol{x}}\right] \,\leq\, C\,.$$

We integrate the former relation between 0 and t and apply Lemma A.1. Since the constant C does not depend on $\boldsymbol{x}$, we take the supremum over K. In the end, we obtain

$$\sup_{\boldsymbol{x}\in K}\left\|\,\mu^\varepsilon_{t,\boldsymbol{x}} \,-\, \mu^\varepsilon_{0,\boldsymbol{x}}\,\right\|_{L^1(\mathbb{R}^2)} \,\leq\, C\,\sqrt{t}\,.$$

□

Using Proposition A.5, we deduce strong continuity at all times for μ^ε.

Proposition A.7 (Continuity at time $t > 0$)**.** *For any $\varepsilon > 0$, suppose that assumptions (2.1a) and (2.2)-(2.4) are fulfilled whereas the initial condition μ^ε_0 verifies the first condition of (2.7), the following estimate holds*

$$\sup_{\boldsymbol{x}\in K}\left\|\,\mu^\varepsilon_{t\,,\,\boldsymbol{x}} \,-\, \mu^\varepsilon_{t+h\,,\,\boldsymbol{x}}\,\right\|_{L^1(\mathbb{R}^2)} \,\leq\, e^{Ct}\,\sup_{\boldsymbol{x}\in K}\left\|\,\mu^\varepsilon_{0\,,\,\boldsymbol{x}} \,-\, \mu^\varepsilon_{h\,,\,\boldsymbol{x}}\,\right\|^{1/2}_{L^1(\mathbb{R}^2)}\,,\quad \forall\,(t\,,\,h)\in\left(\mathbb{R}_+\right)^2\,,$$

for some positive constant C which may depend on ε.

Proof. All along this proof, we consider some $\boldsymbol{x} \in K$ and $h > 0$. We introduce the following notation $\kappa = \left(\mu^\varepsilon_{t\,,\boldsymbol{x}} + \mu^\varepsilon_{t+h\,,\boldsymbol{x}}\right)/2$, which satisfies the following equation

$$\partial_t \kappa = -\partial_v \left(\left(N(v) - w - \left(\frac{\rho^\varepsilon_0}{\varepsilon} + \Psi *_r \rho^\varepsilon_0 \right) v \right) \kappa \right) - \partial_w \left(A(\boldsymbol{u})\, \kappa \right) + \partial_v^2 \kappa$$

$$- \frac{1}{2} \partial_v \left(\mathcal{N}_t\, \mu^\varepsilon_{t,\boldsymbol{x}} + \mathcal{N}_{t+h}\, \mu^\varepsilon_{t+h,\boldsymbol{x}} \right) ,$$

where $\mathcal{N}$ gathers the non-linear terms with respect to the time variable and is given by

$$\mathcal{N}_t(\boldsymbol{x}) = \frac{\rho^\varepsilon_0(\boldsymbol{x})}{\varepsilon} \mathcal{V}^\varepsilon(t,\, \boldsymbol{x}) + \Psi *_r \left(\rho^\varepsilon_0 \mathcal{V}^\varepsilon\right)(t,\, \boldsymbol{x}) .$$

We compute the derivative of $H_{1/2}\left[\, \mu^\varepsilon_{t,\,\boldsymbol{x}} \,|\, \mu^\varepsilon_{t+h,\,\boldsymbol{x}} \,\right]$ using the former equation, equation (1.2) and conservation of mass for (1.2). After an integration by part, all the terms associated to linear transport cancel and we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} H_{1/2}\left[\, \mu^\varepsilon_{t,\,\boldsymbol{x}} \,|\, \mu^\varepsilon_{t+h,\,\boldsymbol{x}} \,\right] + I_{1/2}\left[\, \mu^\varepsilon_{t,\,\boldsymbol{x}} \,|\, \mu^\varepsilon_{t+h,\,\boldsymbol{x}} \,\right] = \mathcal{H}_1 ,$$

where $\mathcal{H}_1$ is given by

$$\mathcal{H}_1 = \left(\mathcal{N}_t - \mathcal{N}_{t+h}\right) \int_{\mathbb{R}^2} \partial_v \left(\ln \frac{\mu^\varepsilon_{t,\,\boldsymbol{x}}}{\kappa} \right) \frac{\mu^\varepsilon_{t,\,\boldsymbol{x}}\, \mu^\varepsilon_{t+h\,,\boldsymbol{x}}}{\kappa} \,\mathrm{d}\boldsymbol{u} .$$

We use Young's inequality, assumptions (2.3) & (2.4) and the inequality $|\,\mu^\varepsilon / \kappa\,| \leq 2$. It yields

$$|\mathcal{H}_1| \leq \eta\, I_{1/2}\left[\, \mu^\varepsilon_{t,\,\boldsymbol{x}} \,|\, \mu^\varepsilon_{t+h,\,\boldsymbol{x}} \,\right] + C\eta^{-1} \sup_{\boldsymbol{x}\in K} |\mathcal{V}(t,\boldsymbol{x}) - \mathcal{V}(t+h,\boldsymbol{x})|^2 ,$$

for all positive η and some constant C that may depend on ε. Then from Proposition A.2, we get exponentional moments on κ and apply Corollary 2.4 in [3], which yields

$$\sup_{\boldsymbol{x}\in K} |\mathcal{V}(t,\boldsymbol{x}) - \mathcal{V}(t+h,\boldsymbol{x})|^2 \leq C \sup_{\boldsymbol{x}\in K} H_{1/2}\left[\, \mu^\varepsilon_{t,\,\boldsymbol{x}} \,|\, \mu^\varepsilon_{t+h,\,\boldsymbol{x}} \,\right] ,$$

for some constant $C > 0$ which may depend on ε. Gathering the former computations and taking $\eta = 1/2$, it yields

$$\frac{\mathrm{d}}{\mathrm{d}t} H_{1/2}\left[\, \mu^\varepsilon_{t,\,\boldsymbol{x}} \,|\, \mu^\varepsilon_{t+h,\,\boldsymbol{x}} \,\right] + \frac{1}{2}\, I_{1/2}\left[\, \mu^\varepsilon_{t,\,\boldsymbol{x}} \,|\, \mu^\varepsilon_{t+h,\,\boldsymbol{x}} \,\right] \leq C \sup_{\boldsymbol{x}\in K} H_{1/2}\left[\, \mu^\varepsilon_{t,\,\boldsymbol{x}} \,|\, \mu^\varepsilon_{t+h,\,\boldsymbol{x}} \,\right] .$$

We integrate this relation between 0 and t and take the supremum over all $\boldsymbol{x}$ in K. It yields

$$\sup_{\boldsymbol{x}\in K} H_{1/2}\left[\, \mu^\varepsilon_{t,\,\boldsymbol{x}} \,|\, \mu^\varepsilon_{t+h,\,\boldsymbol{x}} \,\right] \leq \sup_{\boldsymbol{x}\in K} H_{1/2}\left[\, \mu^\varepsilon_{0,\,\boldsymbol{x}} \,|\, \mu^\varepsilon_{h,\,\boldsymbol{x}} \,\right] + C \int_0^t \sup_{\boldsymbol{x}\in K} H_{1/2}\left[\, \mu^\varepsilon_{s,\,\boldsymbol{x}} \,|\, \mu^\varepsilon_{s+h,\,\boldsymbol{x}} \,\right] \mathrm{d}s .$$

We apply Gronwall's lemma to the former inequality and use Lemma A.1. We obtain

$$\sup_{\boldsymbol{x}\in K} \left\| \mu^\varepsilon_{t,\,\boldsymbol{x}} - \mu^\varepsilon_{t+h\,,\boldsymbol{x}} \right\|_{L^1(\mathbb{R}^2)} \leq C e^{Ct} \sup_{\boldsymbol{x}\in K} \left\| \mu^\varepsilon_{0,\,\boldsymbol{x}} - \mu^\varepsilon_{h,\,\boldsymbol{x}} \right\|^{1/2}_{L^1(\mathbb{R}^2)} .$$

□

A.2. **Uniqueness.** We turn to the proof of uniqueness for equation (1.2). We follow the same method as in the proof of Proposition A.7. We consider two solutions μ^1 and μ^2 to equation (1.2) in the sense of Definition 2.2 and with the same initial condition μ^ε_0. Then we take some $T > 0$ and prove that the following relative entropy

$$H_{1/2}\left[\, \mu^1_{t,\,\boldsymbol{x}} \,|\, \mu^2_{t,\boldsymbol{x}} \,\right] ,$$

is zero for all $(t, \boldsymbol{x}) \in [0\,,\, T] \times K$. All along this proof, we take some $\boldsymbol{x} \in K$ and omit the dependence with respect to $\boldsymbol{x}$ and t when the context is clear. Furthermore we write

$$\left(\mathcal{V}^1\,,\, \mathcal{V}^2\right) = \left(\int_{\mathbb{R}^2} v\, \mu^1 \,\mathrm{d}\boldsymbol{u}\,,\, \int_{\mathbb{R}^2} v\, \mu^2 \,\mathrm{d}\boldsymbol{u} \right) ,$$

and $\kappa = \left(\mu^1 + \mu^2\right)/2$. Since μ^1 and μ^2 are solution to (1.2), κ solves the following equation

$$\partial_t \kappa + \partial_v \left(\left(N(v) - w - \left(\frac{\rho_0^\varepsilon}{\varepsilon} + \Psi *_r \rho_0^\varepsilon \right) \right) \kappa \right) + \partial_w \left(A(\boldsymbol{u})\, \kappa \right) - \partial_v^2 \kappa + \frac{1}{2} \partial_v \left(\mathcal{N}^1 \mu^1 + \mathcal{N}^2 \mu^2 \right) = 0 \,,$$

where $\mathcal{N}^i$, for $i \in \{1\,,\,2\}$, gathers the non-linear terms and is given by

$$\mathcal{N}^i(t,\, \boldsymbol{x}) = \frac{\rho_0^\varepsilon(\boldsymbol{x})}{\varepsilon} \mathcal{V}^i(t,\, \boldsymbol{x}) \,+\, \Psi *_r \left(\rho_0^\varepsilon \mathcal{V}^i \right)(t,\, \boldsymbol{x}) \,.$$

We compute the derivative of $H_{1/2}\left[\,\mu^1 \,|\, \mu^2\,\right]$ using the former equation, equation (1.2) and conservation of mass for (1.2). After an integration by part, all the terms associated to linear transport cancel and we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} H_{1/2}\left[\,\mu^1 \,|\, \mu^2\,\right] \,+\, I_{1/2}\left[\,\mu^1 \,|\, \mu^2\,\right] \,=\, \mathcal{H}_1 \,,$$

where $\mathcal{H}_1$ is given by

$$\mathcal{H}_1 \,=\, \frac{\mathcal{N}^1 \,-\, \mathcal{N}^2}{2} \int_{\mathbb{R}^2} \partial_v \left(\ln\left(\frac{\mu^1}{\kappa} \right) \right) \frac{\mu^2}{\kappa}\, \mu^1 \,\mathrm{d}\boldsymbol{u} \,.$$

According to Theorem 2.3, μ^1 and μ^2 both have uniformly bounded exponential moments on $[0, T]$, hence applying Corollary 2.4 in [3], we obtain that

$$\left| \mathcal{V}^1 \,-\, \mathcal{V}^2 \right| \,\leq\, W_1\left(\,\mu^1\,,\,\mu^2\,\right) \,\leq\, C\, H_{1/2}\left[\,\mu^1 \,|\, \mu^2\,\right]^{1/2} \,,$$

for some positive constant $C > 0$ that may depend on T. Consequently, we use Young's inequality, the estimate $\left|\mu^2/\kappa\right| \,\leq\, 2$, assumptions (2.3) & (2.4) and we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} H_{1/2}\left[\,\mu^1_{t,\boldsymbol{x}} \,|\, \mu^2_{t,\boldsymbol{x}}\,\right] \,+\, \frac{1}{2}\, I_{1/2}\left[\,\mu^1_{t,\boldsymbol{x}} \,|\, \mu^2_{t,\boldsymbol{x}}\,\right] \,\leq\, C \sup_{\boldsymbol{x} \in K} \left(H_{1/2}\left[\,\mu^1_{t,\boldsymbol{x}} \,|\, \mu^2_{t,\boldsymbol{x}}\,\right] \right) \,.$$

We integrate the former relation between 0 and t and take the supremum over all $\boldsymbol{x}$ in K in the left-hand side. It yields

$$\sup_{\boldsymbol{x} \in K} \left(H_{1/2}\left[\,\mu^1_{t,\boldsymbol{x}} \,|\, \mu^2_{t,\boldsymbol{x}}\,\right] \right) \,\leq\, C \int_0^t \sup_{\boldsymbol{x} \in K} \left(H_{1/2}\left[\,\mu^1_{s,\boldsymbol{x}} \,|\, \mu^2_{s,\boldsymbol{x}}\,\right] \right) \mathrm{d}s \,.$$

We apply Gronwall's lemma and conclude the proof.

A.3. **Existence.** In this section, we outline the main ideas in order to construct the solution to equation (1.2) given by Theorem 2.3. Let us first point out that the continuity of the macroscopic quantities $\mathcal{V}^\varepsilon$ and $\mathcal{W}^\varepsilon$ may be deduced from the continuity and the exponential moments of the solution μ^ε. Indeed, according to [3] (see Corollary 2.4), we have

$$(1-\alpha)\left| \mathcal{V}^\varepsilon(t,\, \boldsymbol{x}) \,-\, \mathcal{V}^\varepsilon(s,\, \boldsymbol{y}) \right| \,\leq\, C\, H_{1/2}\left[\mu^\varepsilon_{t,\boldsymbol{x}},\, \mu^\varepsilon_{s,\boldsymbol{y}} \right]^{\frac{1}{2}} \,,$$

as soon as $\mu^\varepsilon_{t,\boldsymbol{x}}$ and $\mu^\varepsilon_{s,\boldsymbol{y}}$ both have exponential moments. Then we apply Lemma A.1 and deduce continuity.

Hence, it is sufficient to prove that μ^ε exists in order to complete the proof. For that matter, we consider the following regularized equation

(A.4)

$$\partial_t\, \mu^R \,+\, \partial_v \left(\left(N^R(v) - w - \frac{\rho_0^\varepsilon}{\varepsilon}(v - \mathcal{V}^R) - \mathcal{K}_\Psi\left[\rho_0^\varepsilon\, \mu^R\right] \right) \mu^R \right) \,+\, \partial_w \left(A(\boldsymbol{u})\, \mu^R \right) \,-\, \partial_v^2\, \mu^R \,=\, 0 \,,$$

where

$$\mathcal{V}^R(t, \boldsymbol{x}) \,=\, \int_{\mathbb{R}^2} v\; \mu^R_{t,\boldsymbol{x}}(\boldsymbol{u})\, \mathrm{d}\boldsymbol{u} \,,$$

and where $\left(N^R\right)_{R>0}$ is a suitable sequence of smooth coefficients. We construct solutions to (A.4) with an iterative scheme. In order to prove that the scheme converges, we use the exact same method as in the proof for uniqueness (see Section A.2).

Then we let the truncation parameter R grow to infinity. Making use of the continuity estimates in Section A.1, we check that Ascoli theorem applies here and prove that the sequence $\left(\mu^R\right)_{R>0}$

is relatively compact. Furthermore, we prove that the limit of any subsequence of $\left(\mu^R\right)_{R>0}$ which converges in $\mathscr{C}^0\left([0,T]\times K\,,\,L^1\left(\mathbb{R}^2\right)\right)$ and which has uniformly bounded exponential moments is a solution to (1.2).

Appendix B. Proof of Proposition 4.2

We provide an entropy estimate in order to ensure existence and a uniqueness estimate in order to ensure that

$$\pi^\varepsilon(t,\cdot)\in\Pi\left(\nu^\varepsilon_{t,\boldsymbol{x}},\,\mathcal{M}_{\rho_0^\varepsilon}\otimes\bar\nu_{t,\boldsymbol{x}}\right),\quad\forall\,t\in\mathbb{R}^+\,.$$

We start with the uniqueness result.

Lemma B.1. *Under the assumptions of Theorem 2.3, consider $\boldsymbol{x}\in K$, $\varepsilon>0$ and a solution π^ε to equation (4.1) in the sense of Definition 4.1 and with initial condition π_0^ε lying in $\Pi\left(\nu^\varepsilon_{0,\boldsymbol{x}},\,\mathcal{M}_{\rho_0^\varepsilon}\otimes\bar\nu_{0,\boldsymbol{x}}\right)$. Then we have*

$$\pi^\varepsilon(t,\cdot)\in\Pi\left(\nu^\varepsilon_{t,\boldsymbol{x}},\,\mathcal{M}_{\rho_0^\varepsilon}\otimes\bar\nu_{t,\boldsymbol{x}}\right),\quad\forall\,t\in\mathbb{R}^+\,.$$

Proof. We define π^1 (resp. π^2) the marginal of π^ε with respect to $\boldsymbol{u}$ (resp. $\boldsymbol{u}'$) and we drop the dependence with respect to time and space when the context is clear. We integrate equation (4.1) with respect to $\boldsymbol{u}'$. We obtain that the difference between π^1 and $\nu^\varepsilon_{t,\,\boldsymbol{x}}$ solves

$$\partial_t\left(\pi^1-\nu^\varepsilon\right)+\operatorname{div}_{\boldsymbol{u}}\left[\mathbf{b}_0^\varepsilon\left(\pi^1-\nu^\varepsilon\right)\right]=\frac{1}{\varepsilon}\,\partial_v\left[\rho_0^\varepsilon\,v\left(\pi^1-\nu^\varepsilon\right)+\partial_v\left(\pi^1-\nu^\varepsilon\right)\right].$$

Multiplying the former equation by $\operatorname{sign}\left(\pi^1-\nu^\varepsilon\right)$, we obtain that $|\pi^1-\nu^\varepsilon|$ is a sub-solution to the former equation. Then we integrate with respect to $\boldsymbol{u}$ and obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\|\pi^1-\nu^\varepsilon\right\|_{L^1(\mathbb{R}^2)}\le 0\,.$$

We follow the method same for π^2 and obtain the expected result. □

We end this section with an entropy estimate, which ensures existence for solutions to (4.1) in the sense of Definition 4.1. We define the Fisher information associated to (4.1)

$$I_{v,v'}\left[\,\pi\,\right]=\int_{\mathbb{R}^4}\left|\,(\,\partial_v+\partial_{v'}\,)\ln\pi\,\right|^2\,\pi\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{u}'\,.$$

Lemma B.2. *Under the assumptions of Theorem 2.3, consider $\boldsymbol{x}\in K$, $\varepsilon>0$ and a solution π^ε to equation (4.1) with some initial condition π_0^ε lying in $\Pi\left(\nu^\varepsilon_{0,\boldsymbol{x}},\mathcal{M}_{\rho_0^\varepsilon}\otimes\bar\nu_{0,\boldsymbol{x}}\right)$. There exists a positive constant $C>0$ which may depend on ε, m_*, m_p and $\overline{m}_p$ such that*

$$H\left[\,\pi_t^\varepsilon\,\right]\le H\left[\,\pi_0^\varepsilon\,\right]+C\,t\,.$$

Proof. We integrate equation (4.1) with respect to $\boldsymbol{u}$ and $\boldsymbol{u}'$ and deduce that mass is conserved through time. Hence, we compute the derivative of $H\left[\,\pi_t^\varepsilon\,\right]$ multiplying equation (4.1) by $\ln\left(\pi_t^\varepsilon\right)$. After an integration by part, it yields

$$\frac{\mathrm{d}}{\mathrm{d}t}H\left[\,\pi_t^\varepsilon\,\right]+\frac{1}{\varepsilon}\,I_{v,v'}\left[\,\pi_t^\varepsilon\,\right]=\Psi*_r\rho_0^\varepsilon(\boldsymbol{x})+2\left(b+\frac{\rho_0^\varepsilon}{\varepsilon}\right)+\mathcal{A}\,,$$

where $\mathcal{A}$ is given by

$$\mathcal{A}=\frac{1}{\sqrt{\varepsilon}}\int_{\mathbb{R}^4}N\left(\mathcal{V}^\varepsilon+\sqrt{\varepsilon}\,v\right)\,\partial_v\left(\ln\pi_t^\varepsilon\right)\,\pi_t^\varepsilon\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{u}'\,.$$

Since the Fisher information $I_{v,v'}$ is somehow degenerate, we can not apply directly Young's inequality to $\mathcal{A}$. Hence, we re-write $\mathcal{A}$ as follows

$$\mathcal{A}=\frac{1}{\sqrt{\varepsilon}}\int_{\mathbb{R}^4}N\left(\mathcal{V}^\varepsilon+\sqrt{\varepsilon}\,v\right)\,(\,\partial_v+\partial'_v\,)\,\left(\ln\pi_t^\varepsilon\right)\,\pi_t^\varepsilon\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{u}'\,.$$

Then, we apply Young's inequality to $\mathcal{A}$ and obtain

$$\mathcal{A} \leq \frac{C\eta}{\varepsilon} I_{v,v'}[\pi_t^\varepsilon] + C\eta \int_{\mathbb{R}^4} \left|N\left(\mathcal{V}^\varepsilon + \sqrt{\varepsilon}\, v\right)\right|^2 \pi_t^\varepsilon \,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{u}',$$

for all η in $]0,1[$. Applying Lemma B.1, assumption (2.1b), Proposition 3.3 and Corollary 3.2, it yields

$$\int_{\mathbb{R}^4} \left|N\left(\mathcal{V}^\varepsilon + \sqrt{\varepsilon}\, v\right)\right|^2 \pi_t^\varepsilon \,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{u}' \leq C.$$

Hence, taking η small enough, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} H[\pi_t^\varepsilon] + \frac{1}{2\,\varepsilon} I_{v,v'}[\pi_t^\varepsilon] \leq C,$$

where C may depend on ε, m_p and $\overline{m}_p$. We integrate the former inequality between 0 and t and obtain the result. □

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 — *Concentration phenomena in Fitzhugh-Nagumo's equations : Part II strong convergence,* in preparation.

[3] F. Bolley and C. Villani — *Weighted Csiszár-Kullback-Pinsker inequalities and applications to transportation inequalities,* Ann. Fac. Sci. Toulouse Math. **14**, (2005), pp. 331—352.

[4] F. Bolley, J. A. Cañizo and J. A. Carrillo — *Stochastic Mean-Field Limit: Non-Lipschitz Forces & Swarming,* Mathematical Models and Methods in Applied Sciences **21**, (2011), pp. 2179–2210.

[5] 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).

[6] P.C. Bressloff — *Spatially periodic modulation of cortical patterns by long-range horizontal connections,* Physica D: Nonlinear Phenomena, **185**, (2003), pp. 131-–157.

[7] N. Brunel and V. Hakim — *Fast global oscillations in networks of integrate-and-fire neurons with low firing rates,* Neural Computation **11**, (1999), pp. 1621–1671.

[8] 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, (2011), pp. 1–7.

[9] 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.

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

[11] 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.

[12] 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.

[13] 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.

[14] J. Crevat — *Mean-field limit of a spatially-extended FitzHugh-Nagumo neural network,* Kinetic & Related Models, **12**, (2019), pp. 1329–1358.

[15] R. FitzHugh — *Impulses and physiological sates in theoretical models of nerve membrane,* Biophysical Journal, **1**, (1961), pp. 445-–466.

[16] N. Fournier and B. Perthame — *Transport distances for PDEs: the coupling method,* EMS Surv. Math. Sci. **7**, (2020), pp. 1–31.

[17] 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).

[18] E. Luçon and W. Stannat — *Mean-field limit for disordered diffusions with singular interactions,* The Annals of Applied Probability, **24**, (2014), pp. 1946—1993.

[19] S. Mischler, C. Quiñinao and J. Touboul — *On a kinetic FitzHugh-Nagumo model of neuronal network,* Comm. Math. Phys. **342**, (2015), pp. 1001—1042.

[20] J. Nagumo, S. Arimoto and S. Yoshizawa — *An active pulse transmission line simulating nerve axon,* Proceedings of the IRE **50**, (1962), pp. 2061—2070.

[21] K. Pakdaman, B. Perthame and D. Salort *Dynamics of a structured neuron population,* Nonlinearity **23**, (2010), pp. 55–75.
[22] C. Quiñinao and J. Touboul — *Clamping and synchronization in the strongly coupled FitzHugh-Nagumo model,* SIAM J. Appl. Dyn. Syst. **19**, (2020), pp 788-–827.
[23] N. Torres and D. Salort — *Dynamics of neural networks with elapsed time model and learning processes,* Acta Appl. Math. **170**, (2020), pp. 1065-–1099.
[24] C. Villani — *Optimal transport, old and new,* Springer Berlin Heidelberg, (2009).