Concentration profiles in Fitzhugh-Nagumo neural networks: A Hopf-Cole approach

Alain Blaustein, Emeric Bouin

CONCENTRATION PROFILES IN FITZHUGH-NAGUMO NEURAL NETWORKS: A HOPF-COLE APPROACH

Alain Blaustein [1] and Emeric Bouin [2]

Abstract. In this paper we focus on a mean-field model for a spatially extended FitzHugh-Nagumo neural network. In the regime where strong and local interactions dominate, we quantify how the probability density of voltage concentrates into a Dirac distribution. Previous work investigating this question have provided relative bounds in integrability spaces. Using a Hopf-Cole framework, we derive precise L^∞ estimates using subtle explicit sub- and super- solutions which prove, with rates of convergence, that the blow up profile is Gaussian.

1. Introduction

The model. Understanding the complex dynamics induced by interactions in large assemblies of neurons constitute one of the great challenge in neuroscience. As described in neuroscience textbooks [38], neurons behave and interact according to intricate chemical and electrical mechanisms. Due to the complexity of these mechanisms, it is mandatory to consider simplified models. A key step in this direction is the pioneering work of A. Hodgkin and A. Huxley [33, 34], who built an accurate model describing the membrane potential dynamics of a single nerve cell submitted to an external current. This model captures the main features of a neuron's membrane potential behaviour such as periodic patterns, relaxation toward equilibrium state as well as spiking behaviour also known as action potential. It falls into the category of so called "voltage-conductance based models" which describe the dynamics of the membrane potential through auxiliary variables taking into account ionic exchanges between a neuron and its extra-cellular environment (see [12, 27, 38] for precise introductions to such models). Due to the complexity of this model, we focus on a simplified version introduced by R. FitzHugh and J. Nagumo [30, 49], which conserves the main features of the Hodgkin-Huxley model while remaining more tractable from a mathematical point of view

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

where $v_t \in \mathbb{R}$ is the membrane potential coupled whereas $w_t \in \mathbb{R}$ is an adaptation variable. The non-linear drift N takes the form $N(v) = v - v^3$ in the original articles of R. FitzHugh and J. Nagumo but we shall consider a broader class of drifts here. Coefficient A is given by

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

where a, $c \in \mathbb{R}$ and $b > 0$. The Brownian motion B_t describes non deterministic fluctuations of the potential which are not taken into account by the model, whereas I_{ext} stands for the current received by the neuron from its environment. Since our purpose is to describe interactions between neurons,

Date: January 20, 2024.
2020 *Mathematics Subject Classification.* Primary: 35Q92, Secondary: 35B40; 35C20; 35K57; 35G20 .
Key words and phrases. Hopf-Cole transform; FitzHugh-Nagumo; mean-field equation; concentration; neural network; asymptotic limit.

[1]Institut de Mathématiques de Toulouse, Université Paul Sabatier, 118 route de Narbonne - F-31062 Toulouse Cedex 9, France. E-mail : alain.blaustein@math.univ-toulouse.fr
[2]CEREMADE - Université Paris-Dauphine, UMR CNRS 7534, Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France. E-mail: bouin@ceremade.dauphine.fr

we introduce coupling through the term I_{ext}: we consider Ohmic interactions between neurons with spatially dependent conductance given by a connectivity kernel $\Phi : K \times K \to \mathbb{R}$, where K is a compact set of $\mathbb{R}^d$. Hence, in the case of m interacting neurons described by the triplet voltage-adaptation-position $(v_i, w_i, \boldsymbol{x}_i)_{1\leq i\leq m}$, the current received by neuron i from its neighbors is given by

$$I_{ext} \,=\, -\frac{1}{m}\sum_{j=1}^{m} \Phi(\boldsymbol{x}_i, \boldsymbol{x}_j)\,(v_t^i - v_t^j)\,. \tag{1.1}$$

Therefore, we obtain the following microscopic description of a FitzHugh-Nagumo neural network

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

where $i \in \{1, ..., m\}$. The mean-field limit, corresponding to $m \to +\infty$ in the latter system, was rigorously analysed in the case of FitzHugh-Nagumo neurons [24, 45] as well as in more general cases, including the Hodkin-Huxley case [1, 9, 41]; we also mention similar works for mean-field limits with non-exchangeable systems [36, 55] and [8] for a related model in collective dynamics. It was proved that the empirical measure associated to the latter system converges towards 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 membrane potential v and adaptation variable $w \in \mathbb{R}$. The dynamics of the distribution function f are prescribed by 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

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

This model is a typical example of McKean-Vlasov equation including voltage and conductance variables; other models of this type are available in the literature [16, 54] as well as other popular family of models including integrate-and-fire neural networks [13, 14, 17] and time-elapsed neuronal models [20, 21, 22, 46, 53]. We mention that the discrete analog of $\mathcal{K}_\Phi[f]$ given by (1.1) may be recovered replacing f in the definition of $\mathcal{K}_\Phi[f]$ by the empirical distribution of the microscopic model.

The question at hand. In the present article, we are interested in the dynamics of the network when short-range interactions dominate: we consider a situation where the connectivity kernel Φ decomposes as follows

$$\Phi(\boldsymbol{x}, \boldsymbol{x}') \,=\, \Psi(\boldsymbol{x}, \boldsymbol{x}') \,+\, \frac{1}{\varepsilon}\,\delta_0(\boldsymbol{x} - \boldsymbol{x}')\,, \tag{1.2}$$

where the Dirac mass δ_0 accounts for short-range interactions whereas the interaction kernel Ψ models long-range interactions. The scaling parameter $\varepsilon > 0$ represents the magnitude of short-range interactions: we now write f^ε instead of f and focus on the regime $\varepsilon \ll 1$. Such decomposition is biologically relevant since visual [35] and sensory [47, 48] cortices are organized into sub-regions called cortical columns within which neuron strongly interact and have thus similar answers to external signals [42]. This structure has been studied both at mathematical and numerical levels [6, 7, 11, 25, 26, 56]. We point out that it is possible to replace the Dirac mass approximation for short range interactions with a localized kernel [11, 26].
Let us present what is known in our particular context. To do so, we introduce the macroscopic

quantities associated to the network: the spatial distribution of neurons (which is time-homogeneous, according to an integration of the mean-field equation with respect to $\boldsymbol{u} \in \mathbb{R}^2$)

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

as well as the averaged voltage and adaptation variable at spatial location $\boldsymbol{x} \in K$

$$\mathcal{U}^\varepsilon := (\mathcal{V}^\varepsilon, \mathcal{W}^\varepsilon)\ , \quad \text{with} \quad \begin{cases} \rho_0^\varepsilon(\boldsymbol{x})\, \mathcal{V}^\varepsilon(t, \boldsymbol{x}) & = \displaystyle\int_{\mathbb{R}^2} v\ f^\varepsilon(t, \boldsymbol{x}, \boldsymbol{u}) \mathrm{d}\boldsymbol{u} \\ \rho_0^\varepsilon(\boldsymbol{x})\, \mathcal{W}^\varepsilon(t, \boldsymbol{x}) & = \displaystyle\int_{\mathbb{R}^2} w\ f^\varepsilon(t, \boldsymbol{x}, \boldsymbol{u}) \mathrm{d}\boldsymbol{u} \end{cases}. \tag{1.3}$$

We outline that the interaction term $\mathcal{K}_\Phi[f^\varepsilon]$ admits the following simple expression in terms of the macroscopic quantities

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

where $*$ denotes the convolution on the right side of any function g with Φ

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

According to (1.2), we substitute Φ with $\Psi + \delta_0/\varepsilon$ in the latter expression of $\mathcal{K}_\Phi[f^\varepsilon]$; it yields

$$\mathcal{K}_\Phi[f^\varepsilon](t,\ \boldsymbol{x}, v) \ = \ \Psi * \rho_0^\varepsilon(\boldsymbol{x})\, v - \Psi * (\rho_0^\varepsilon\ \mathcal{V}^\varepsilon)(t, \boldsymbol{x}) \ + \ \frac{1}{\varepsilon} \rho_0^\varepsilon(\boldsymbol{x})\, (v - \mathcal{V}^\varepsilon(t, \boldsymbol{x}))\,.$$

Therefore, the mean-field equation may be rewritten

$$\partial_t f^\varepsilon + \mathrm{div}_{\boldsymbol{u}}\left[\mathbf{b}^\varepsilon f^\varepsilon\right] - \partial_v^2 f^\varepsilon = \frac{1}{\varepsilon} \rho_0^\varepsilon \partial_v \left[(v - \mathcal{V}^\varepsilon) f^\varepsilon\right], \tag{1.4}$$

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

$$\mathbf{b}^\varepsilon(t, \boldsymbol{x}, \boldsymbol{u}) := \begin{pmatrix} B^\varepsilon(t, \boldsymbol{x}, \boldsymbol{u}) \\ \\ A(\boldsymbol{u}) \end{pmatrix} = \begin{pmatrix} N(v) - w - v\, \Psi * \rho_0^\varepsilon + \Psi * (\rho_0^\varepsilon\ \mathcal{V}^\varepsilon) \\ \\ av - bw + c \end{pmatrix}. \tag{1.5}$$

To infer the asymptotic behaviour of the network in the regime of strong interactions, we look for the leading order in (1.4): in our case, it is induced by short-range interactions between neurons, and as $\varepsilon \to 0$, we expect

$$(v - \ \mathcal{V}^\varepsilon) f^\varepsilon \underset{\varepsilon \to 0}{\to} 0,$$

to make sure that no term is singular in (1.4). Since (1.4) conserves total mass, this means that f^ε concentrates into a Dirac ditribution centered in $\mathcal{V}^\varepsilon$ with respect to the v-variable, that is

$$f^\varepsilon(t, \boldsymbol{x}, \boldsymbol{u}) \underset{\varepsilon \to 0}{\approx} \delta_0\left(v - \mathcal{V}^\varepsilon(t, \boldsymbol{x})\right) \otimes F^\varepsilon(t, \boldsymbol{x}, w), \tag{1.6}$$

where $\mathcal{V}^\varepsilon$ is given by (1.3) and F^ε is defined as the marginal of f^ε with respect to the voltage variable

$$F^\varepsilon\left(t, \boldsymbol{x}, w\right) = \int_{\mathbb{R}} f^\varepsilon\left(t, \boldsymbol{x}, \boldsymbol{u}\right) \mathrm{d}v.$$

Multiplying equation (1.4) by $\boldsymbol{u}/\rho_0^\varepsilon$ and integrating with respect to $\boldsymbol{u}$, one finds that the couple $(\mathcal{V}^\varepsilon, \mathcal{W}^\varepsilon)$ solves the following system

$$\begin{cases} \partial_t \mathcal{V}^\varepsilon = N(\mathcal{V}^\varepsilon) - \mathcal{W}^\varepsilon - \mathcal{V}^\varepsilon \Psi * \rho_0^\varepsilon + \Psi * (\rho_0^\varepsilon \mathcal{V}^\varepsilon) + \mathcal{E}(f^\varepsilon), \\ \\ \partial_t \mathcal{W}^\varepsilon = A(\mathcal{V}^\varepsilon, \mathcal{W}^\varepsilon), \end{cases} \tag{1.7}$$

where the error term $\mathcal{E}\left(f^\varepsilon\right)$ is given by

$$\mathcal{E}\left(f^\varepsilon\left(t, \boldsymbol{x}, \cdot\right)\right) \ = \ \frac{1}{\rho_0^\varepsilon(\boldsymbol{x})} \int_{\mathbb{R}^2} N(v) f^\varepsilon\left(t, \boldsymbol{x}, \boldsymbol{u}\right) \mathrm{d}\boldsymbol{u} - N(\mathcal{V}^\varepsilon)\,. \tag{1.8}$$

All this, in turn, implies that as ε vanishes, $(\mathcal{V}^\varepsilon, \mathcal{W}^\varepsilon)$ converges to the couple $(\mathcal{V}, \mathcal{W})$, which solves

$$\begin{cases} \partial_t \mathcal{V} = N(\mathcal{V}) - \mathcal{W} - \mathcal{V}\Psi * \rho_0 + \Psi * (\rho_0 \mathcal{V}), \\ \partial_t \mathcal{W} = A(\mathcal{V}, \mathcal{W}), \\ (\mathcal{V}(0,\cdot), \mathcal{W}(0,\cdot)) = (\mathcal{V}_0, \mathcal{W}_0), \end{cases} \tag{1.9}$$

where $\rho_0 = \lim_{\varepsilon\to 0} \rho_0^\varepsilon$. As of the marginal F^ε, it has been shown in [6, 7, 25] that it also converges. The reaction-diffusion system (1.9) has been extensively studied, especially since it features traveling waves solutions [15, 32, 37]. In (1.9), the non-local term may be replaced by a diffusion operator, sometimes also acting on the adaptation variable [19]. Macroscopic systems with diffusion operators may be obtained from (1.4) as $\varepsilon \to 0$ taking localized instead of punctual interactions in (1.2), as shown in [26].

In this article, we refine convergence (1.6) by investigating the concentration profile of the solution f^ε when ε goes to 0. To this end, we perform the so-called Hopf-Cole transform of f^ε

$$\phi^\varepsilon(t, \boldsymbol{x}, \boldsymbol{u}) := \varepsilon \ln\left(\sqrt{\frac{2\pi\varepsilon}{\rho_0}} f^\varepsilon(t, \boldsymbol{x}, \boldsymbol{u})\right), \qquad \forall (t, \boldsymbol{x}, \boldsymbol{u}) \in \mathbb{R}^+ \times K \times \mathbb{R}^2, \tag{1.10}$$

and study the convergence of ϕ^ε as ε goes to zero. This approach has been widely followed to study concentration phenomena occurring in selection-mutation models in population dynamics [3, 5, 10, 18, 29, 39, 43, 44]. Quininao and Touboul have shown in [56] that it can lead to fruitful results in the present context. Indeed, inverting (1.10), we obtain

$$f^\varepsilon = \sqrt{\frac{\rho_0}{2\pi\varepsilon}} \exp\left(\frac{\phi^\varepsilon}{\varepsilon}\right),$$

and therefore expect $\phi^\varepsilon \leq 0$ as $\varepsilon \to 0$. Furthermore, we see that the concentration points of f^ε are characterized by the level sets $\{\phi^\varepsilon = 0\}$. Here, we are specifically interested in the behavior of ϕ^ε at $v \to +\infty$, which describes precisely the asymptotic tail of f^ε with respect to v.

Heuristics. Let us now formally justify the convergence of ϕ^ε: injecting *ansatz* (1.10) in equation (1.4), we find that ϕ^ε solves the following Hamilton-Jacobi equation

$$\partial_t \phi^\varepsilon + \nabla_{\boldsymbol{u}} \phi^\varepsilon \cdot \mathbf{b}^\varepsilon + \varepsilon \operatorname{div}_{\boldsymbol{u}}[\mathbf{b}^\varepsilon] - \partial_v^2 \phi^\varepsilon - \rho_0^\varepsilon = \frac{1}{\varepsilon} \partial_v \left(\frac{1}{2}\rho_0^\varepsilon |v - \mathcal{V}^\varepsilon|^2 + \phi^\varepsilon\right) \partial_v \phi^\varepsilon. \tag{1.11}$$

Keeping only the leading order in equation (1.11), we expect

$$\phi^\varepsilon \underset{\varepsilon\to 0}{\approx} \phi,$$

where ϕ satisfies

$$\partial_v \left(\frac{1}{2}\rho_0^\varepsilon |v - \mathcal{V}^\varepsilon|^2 + \phi\right) \partial_v \phi = 0.$$

To determine ϕ, we reformulate the latter equation in the following equivalent form

$$\partial_v \phi(t, \boldsymbol{x}, v) = -\delta(t, \boldsymbol{x}, v)\, \rho_0^\varepsilon(\boldsymbol{x})\, (v - \mathcal{V}^\varepsilon(t, \boldsymbol{x})), \quad \forall (t, \boldsymbol{x}, v) \in \mathbb{R}^+ \times K \times \mathbb{R}, \tag{1.12}$$

where $\delta(t, \boldsymbol{x}, v)$ takes values in $\{0, 1\}$. We fix some $(t, \boldsymbol{x}) \in \mathbb{R}^+ \times K$ and suppose on the one hand $\partial_v \phi(t, \boldsymbol{x}, \cdot)$ to be smooth with respect to v and on the other hand $\rho_0^\varepsilon(\boldsymbol{x}) > 0$. Since $\partial_v \phi(t, \boldsymbol{x}, \mathcal{V}^\varepsilon) = 0$, we may divide (1.12) by $\rho_0^\varepsilon(\boldsymbol{x})\,(v - \mathcal{V}^\varepsilon(t, \boldsymbol{x}))$ and deduce that $\delta(t, \boldsymbol{x}, \cdot)$ is a smooth function of v. Together with the fact that it takes discrete values, this implies that it does not depend on v. After integrating (1.12) with respect to v and passing to the limit $\varepsilon \to 0$, this yields

$$\phi^\varepsilon(t, \boldsymbol{x}, v) \underset{\varepsilon\to 0}{\longrightarrow} -\frac{\delta}{2}(t, \boldsymbol{x})\, \rho_0(\boldsymbol{x})\, |v - \mathcal{V}(t, \boldsymbol{x})|^2 + c(t, \boldsymbol{x}).$$

Furthermore, since our problem conserves mass, we expect for each $(t, \boldsymbol{x}) \in \mathbb{R}^+ \times K$

$$\int_{v\in\mathbb{R}} \exp\left(-\frac{\delta}{2\,\varepsilon}\,\rho_0\,|v-\mathcal{V}|^2+\frac{c}{\varepsilon}\right)(t,\boldsymbol{x})\,\mathrm{d}v \;=\; \sqrt{\frac{2\pi\varepsilon}{\rho_0(\boldsymbol{x})}}\,,$$

for all ε. This forces $\delta(t,\boldsymbol{x}) = 1$ and $c(t,\boldsymbol{x}) = 0$; thus we obtain

$$\phi^\varepsilon(t,\boldsymbol{x},v) \underset{\varepsilon\to 0}{\longrightarrow} -\frac{1}{2}\,\rho_0(\boldsymbol{x})\,|v-\mathcal{V}(t,\boldsymbol{x})|^2\,,$$

where $\mathcal{V}$ is the solution to (1.9). This convergence is the object of our main result, Theorem 1.5 below, in which we provide explicit convergence rates with respect ε. Before going further, we shall be precise about the mathematical framework of this article.

Mathematical framework. In this paragraph, we first state and motivate our assumptions on the data of the problem: N, Ψ and f_0^ε. Then we precise the notion of solution we consider for equation (1.4). We suppose that the drift $N \in \mathscr{C}^2(\mathbb{R})$ satisfies

$$\limsup_{|v|\to+\infty} \frac{N(v)}{\mathrm{sgn}(v)|v|^p} < 0, \qquad \sup_{|v|\geq 1}\left|\frac{N(v)}{|v|^p}\right| < +\infty, \tag{1.13}$$

for some $p \geq 2$, and

$$\sup_{|v|\geq 1}\left(\left|N''(v)\right|+\left|N'(v)\right|\right)|v|^{-p'} < +\infty, \tag{1.14}$$

for some $p' \geq 0$.

Remark 1.1. *On the one hand, assumption (1.13) is a key feature in the model proposed by R. FitzHugh and J. Nagumo: it says that N has super-linear confining properties in the sense that it decays super-linearly at infinity. On the other hand, assumption (1.14) is technical yet not restrictive in our case since it is satisfied when N is given by*

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

which is the original choice in the FitzHugh-Nagumo model. More generally, (1.13)-(1.14) are satisfied by all drifts $P(v)$ given by

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

for some positive constant $C > 0$ and where Q is a polynomial function with degree less then p.

The connectivity kernel satisfies

$$(\Psi : \boldsymbol{x} \longmapsto \Psi(\boldsymbol{x}, \cdot)) \in \mathscr{C}^0\left(K, L^1(K)\right)\,, \tag{1.15}$$

and

$$\sup_{\boldsymbol{x}'\in K}\int_K \left|\Psi(\boldsymbol{x},\boldsymbol{x}')\right|\mathrm{d}\boldsymbol{x} < \;+\infty, \qquad \sup_{\boldsymbol{x}\in K}\int_K \left|\Psi(\boldsymbol{x},\boldsymbol{x}')\right|^r\mathrm{d}\boldsymbol{x}' < +\infty, \tag{1.16}$$

for some $r > 1$; in the sequel we define r' by $\frac{1}{r}+\frac{1}{r'} = 1$. This set of assumptions on Ψ allows for non-symmetric interactions between neurons [11], interactions following negative power law [23, 31, 41, 50] as well as "nearest-neighbor" type interactions [41, 51, 52].

We now specify the notion of solution we consider for equation (1.4). To this end, we state our assumptions on f_0^ε. We suppose, for each $\varepsilon > 0$

$$f_0^\varepsilon \in \mathscr{C}^0\left(K, L^1\left(\mathbb{R}^2\right)\right), \quad f^\varepsilon \geq 0 \quad \text{and} \quad \int_{\mathbb{R}^2} f_0^\varepsilon(\boldsymbol{x},\boldsymbol{u})\mathrm{d}\boldsymbol{u}\mathrm{d}\boldsymbol{x} = 1, \tag{1.17}$$

which ensures $\rho_0^\varepsilon \in \mathscr{C}^0(K)$. We also suppose

$$m_* \leq \rho_0^\varepsilon \leq 1/m_*, \tag{1.18}$$

for some positive constant m_* independent of ε. On top of that, we assume the following condition: there exist two positive constants m_p and $\overline{m}_p$, independent of ε, such that

$$\sup_{\boldsymbol{x}\in K}\int_{\mathbb{R}^2}|\boldsymbol{u}|^{2(p+p')}f_0^\varepsilon(\boldsymbol{x},\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\leq m_p, \tag{1.19}$$

and such that

$$\int_{K\times\mathbb{R}^2}|\boldsymbol{u}|^{2(p+p')r'}f_0^\varepsilon(\boldsymbol{x},\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\mathrm{d}\boldsymbol{x}\leq\overline{m}_p, \tag{1.20}$$

where p, p' and r' are given in (1.13), (1.14) and (1.16).

Remark 1.2. *Moment assumptions such as (1.19)-(1.20) are common in the literature of mean-field description for neural networks (see* [28, Assumption 2], [13, Theorem 2.2 and Corollary 2.4], [54, Section 5.1]*). In our context, it allows to propagate moments of the solution f^ε to (1.4), which in turn provides control over macroscopic quantities $(\mathcal{V}^\varepsilon,\mathcal{W}^\varepsilon)$ and over the error term $\mathcal{E}(f^\varepsilon)$ given by (1.8) (see the second item of Theorem 2.4 for more details). The order of moment required is related to the drift N through exponents p and p' defined in (1.13)-(1.14): this is natural in order to control $\mathcal{E}(f^\varepsilon)$ which displays the drift N. These assumptions might not be optimal in the sense that the order of moment may be lowered; however we do not investigate any further this technical aspect since it is not our main interest here.*

We consider the following solutions to (1.4)

Definition 1.3. *For all $\varepsilon>0$, we say that f^ε solves (1.4) with initial condition f_0^ε if $f^\varepsilon\in\mathscr{C}^0\left(\mathbb{R}^+\times K,L^1\left(\mathbb{R}^2\right)\right)$ and for all $\boldsymbol{x}\in K$, $t\geq0$, and $\varphi\in\mathscr{C}_c^\infty\left(\mathbb{R}^2\right)$, it holds*

$$\begin{aligned}\int_{\mathbb{R}^2}\varphi(\boldsymbol{u})\left(f^\varepsilon(t,\boldsymbol{x},\boldsymbol{u})-f_0^\varepsilon(\boldsymbol{x},\boldsymbol{u})\right)\mathrm{d}\boldsymbol{u}&=\int_0^t\int_{\mathbb{R}^2}\left[\left(\nabla_{\boldsymbol{u}}\varphi\cdot\mathbf{b}^\varepsilon+\partial_v^2\varphi\right)f^\varepsilon\right](s,\boldsymbol{x},\boldsymbol{u})\mathrm{d}\boldsymbol{u}\mathrm{d}s\\&\quad-\frac{\rho_0^\varepsilon(\boldsymbol{x})}{\varepsilon}\int_0^t\int_{\mathbb{R}^2}\left[\partial_v\varphi\left(v-\mathcal{V}^\varepsilon\right)f^\varepsilon\right](s,\boldsymbol{x},\boldsymbol{u})\mathrm{d}\boldsymbol{u}\mathrm{d}s,\end{aligned}$$

where $\mathcal{V}^\varepsilon$ and $\mathbf{b}^\varepsilon$ are given by (1.3) and (1.4) respectively.

With this notion of solution, equation (1.4) is well-posed, the following result being proved in [7].

Theorem 1.4 ([7])**.** *For any $\varepsilon>0$, suppose that assumptions (1.13) on N, (1.16) on Ψ and (1.17)-(1.18) on the initial condition are fulfilled and that f_0^ε also verifies*

$$\begin{cases}\displaystyle\sup_{\boldsymbol{x}\in K}\int_{\mathbb{R}^2}e^{|\boldsymbol{u}|^2/2}f_0^\varepsilon(\boldsymbol{x},\boldsymbol{u})\mathrm{d}\boldsymbol{u}<+\infty,\\[2ex]\displaystyle\sup_{\boldsymbol{x}\in K}\int_{\mathbb{R}^2}\ln\left[f_0^\varepsilon(\boldsymbol{x},\boldsymbol{u})\right]f_0^\varepsilon(\boldsymbol{x},\boldsymbol{u})\mathrm{d}\boldsymbol{u}<+\infty,\end{cases}$$

and

$$\sup_{\boldsymbol{x}\in K}\left\|\nabla_{\boldsymbol{u}}\sqrt{f_0^\varepsilon}\right\|_{L^2(\mathbb{R}^2)}^2<+\infty.$$

Then there exists a unique solution f^ε to equation (1.4) with initial condition f_0^ε, in the sense of Definition 1.3 which verifies

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

for all $T\geq0$.

Let us now state our main result.

Main result. The following theorem states that in the regime of strong local interactions, the voltage distribution of the neural network described by (1.4) blows up into a Dirac distribution and that concentration occurs with Gaussian profile. More specifically, we prove that the Hopf-Cole transform ϕ^ε of f^ε defined by (1.10) converges to $-\rho_0 \left|v-\mathcal{V}\right|^2/2$ uniformly with respect to all variables $(t,\boldsymbol{x},\boldsymbol{u})$ as ε vanishes. Furthermore, we prove that convergence occurs at rate $O(\varepsilon)$, which is (at least formally) optimal.

Theorem 1.5. *Assume (1.13)-(1.20) and the additional assumptions of Theorem 1.4 and Proposition 2.2. Suppose that there exists a positive constant C independent of ε such that the following compatibility assumption holds*

$$\|\mathcal{U}_0-\mathcal{U}_0^\varepsilon\|_{L^\infty(K)}+\|\rho_0-\rho_0^\varepsilon\|_{L^\infty(K)}\leq C\varepsilon, \tag{1.21}$$

as well as the following set of "smallness assumptions"

$$\left|\phi_0^\varepsilon+\frac{1}{2}\rho_0\left|v-\mathcal{V}_0\right|^2-\varepsilon n\right|\leq \varepsilon C\left(1+|\boldsymbol{u}|^2\right),\quad \forall(\boldsymbol{x},\boldsymbol{u})\in K\times\mathbb{R}^2, \tag{1.22a}$$

$$\int_{\mathbb{R}^2}\left(|v-\mathcal{V}_0^\varepsilon|^2+|v-\mathcal{V}_0^\varepsilon|^{p'+1}\right)f_0^\varepsilon(\boldsymbol{x},\boldsymbol{u})\mathrm{d}\boldsymbol{u}\leq C\varepsilon,\quad \forall \boldsymbol{x}\in K, \tag{1.22b}$$

for all $\varepsilon>0$, where n is a primitive of N: $n'(v)=N(v)$. Then the sequence $(\phi^\varepsilon)_{\varepsilon>0}$ of Hopf-Cole transforms of the solutions $(f^\varepsilon)_{\varepsilon>0}$ to (1.4) is well defined and it converges locally uniformly on $\mathbb{R}^+\times K\times\mathbb{R}^2$ to $-\rho_0\left|v-\mathcal{V}\right|^2/2$ with rate ε, where $\mathcal{V}$ is given by (1.9). More precisely, there exist two positive constants C and ε_0 such that for all $\varepsilon\leq\varepsilon_0$,

$$\left|\phi^\varepsilon+\frac{1}{2}\rho_0\left|v-\mathcal{V}\right|^2-\varepsilon n\right|(t,\boldsymbol{x},\boldsymbol{u})\leq \varepsilon Ce^{Ct}\left(1+|\boldsymbol{u}|^2\right),\quad \forall(t,\boldsymbol{x})\in\mathbb{R}^+\times K,\ a.e.\ in\ \boldsymbol{u}\in\mathbb{R}^2.$$

As a consequence, f^ε converges uniformly to 0 on the compact subsets of $\mathbb{R}^+\times K\times\mathbb{R}^2\backslash\{v\neq\mathcal{V}(t,\boldsymbol{x})\}$. In the latter results, constants C and ε_0 only depend on the data of our problem: f_0^ε (only through the constants appearing in assumptions (1.18)-(1.20) and (1.21)-(1.22b)), N, A and Ψ.

Before going further into our analysis, let us comment on our result. We first emphasize that it deals with uniform convergence with respect to all variables, which is a great improvement in comparison to former results obtained in [6, 7], where L^1, L^2 and weak convergence estimates were obtained.

We also point out that the present article is in line with [56], where a compactness argument was used to obtain the convergence of ϕ^ε for truncated versions of the drift N. In the present article, we take advantage of the particular structure of the problem, specifically of the confining properties of N provided by assumption (1.13) to obtain a rate of convergence with respect to ε.

We emphasize that our result holds in a perturbative setting. To illustrate this remark, we consider the following initial condition, which satisfies the assumptions of Theorem 1.5

$$f_0^\varepsilon(\boldsymbol{x},\boldsymbol{u})=c_\varepsilon(\boldsymbol{x})\frac{\rho_0(\boldsymbol{x})^{\frac{3}{2}}}{\pi\sqrt{2\varepsilon}}\exp\left(-\frac{\rho_0(\boldsymbol{x})}{2\,\varepsilon}\left|v-\mathcal{V}_0(\boldsymbol{x})\right|^2-\left|w-\mathcal{W}_0(\boldsymbol{x})\right|^2+n(v)\right),$$

where c_ε is a normalizing constant such that for all $\boldsymbol{x}\in K$, it holds

$$\int_{\mathbb{R}^2\times K}f_0^\varepsilon(\boldsymbol{u},\boldsymbol{x})\mathrm{d}\boldsymbol{u}=\rho_0(\boldsymbol{x}),$$

In the latter example, we see that f_0^ε is concentrated with respect to v. This restriction appears in most of our references [3, 10, 29, 43, 44, 56] using the Hopf-Cole transform. We also outline that, since $N(v)$ decays super-linearly at $\pm\infty$, its primitive $n(v)$ induces fast decay of f_0^ε as $v\to\pm\infty$. This condition is natural since $n(v)$ appears in the formal expansion of ϕ^ε as $\varepsilon\to0$ (see (3.3)).

To conclude, diffusion with respect to w in (1.4) is sometimes taken into account [40, 41, 56]. Whether our analysis extend to this case may depend on the scaling with respect to ε of the additional term. For example, adding $\varepsilon\partial_w^2 f^\varepsilon$ to (1.4) should simplify the analysis, as stated in [56, page

5].

Comments on the strategy. Let us outline our strategy and the challenges in order to prove Theorem 1.5. As identified in [56], the main difficulty is induced by the drift N, which is not globally Lipschitz according to assumption (1.13). To bypass this difficulty, authors in [56] rely on a regularization argument: they consider truncated versions N^R, for $R > 0$, of N, in order to recover global Lipschitz properties, allowing to prove uniform estimates in ε on the derivatives $\nabla_{\boldsymbol{u}}\phi^{\varepsilon,R}$ of the truncated problem thanks to the Bernstein method (see [2] for a general description of the method and [3] for another application). Then, they conclude that $\phi^{\varepsilon,R}$ converges as ε goes to zero relying a compactness argument. The argument is robust and may apply to a wider range of problems. An alternate mean to carry out the proof would be to use the method of half-relaxed limits (see [4] for a general introduction to this method and [43] for an application) which applies without requiring any regularity estimates, at the cost of loosing continuity and therefore uniqueness in the limit $\varepsilon \to 0$. To recover uniqueness, we may add the additional constraint

$$\phi(t, \boldsymbol{x}, \mathcal{V}, w) \;=\; 0$$

on the limiting problem. However, proving that the limit provided by this method satisfies this constraint also requires uniform regularity estimates on the derivatives of ϕ^ε and we are back to our initial problem, since N is not globally Lipschitz.
In this article, we take advantage of the particular structure of the problem, specifically of the confining properties of N, to build a method which does not require regularity estimates and which has the advantage of providing explicit convergence rates: instead of proving uniform estimates on the derivatives of ϕ^ε, we prove uniform estimates on the first term in the expansion of ϕ^ε with respect to ε. This is made possible since this first term takes into account the non-linear fluctuations induced by N. Indeed, these non-linear variations induced by N are expected to be perturbations of order ε, as it may be seen rewriting equation (1.11) on ϕ^ε as follows

$$\frac{1}{\varepsilon}\,\partial_v\left(-\frac{1}{2}\rho_0^\varepsilon\,|v-\mathcal{V}^\varepsilon|^2+\varepsilon n(v)-\phi^\varepsilon\right)\partial_v\phi^\varepsilon+\ldots=0,$$

where the correction $n(v)$ is such that $n'(v) = N(v)$ and where "..." gathers the lower order terms with respect to v and w. Hence, at least formally as ε goes to zero, we expect

$$\phi^\varepsilon \underset{\varepsilon\to 0}{\approx} -\frac{1}{2}\rho_0^\varepsilon\,|v-\mathcal{V}^\varepsilon|^2+\varepsilon n(v).$$

Therefore, we consider the first term ϕ_1^ε in the expansion of ϕ^ε with respect to ε, that is

$$\phi^\varepsilon = -\frac{1}{2}\rho_0^\varepsilon\,|v-\mathcal{V}^\varepsilon|^2+\varepsilon\phi_1^\varepsilon,$$

and identify its formal equivalent $\overline{\phi_1^\varepsilon}$ as ε goes to zero, which displays $n(v)$ and which depends on ε only through the macroscopic quantities $(\mathcal{V}^\varepsilon, \mathcal{W}^\varepsilon)$ (see Section 3). In Lemma 3.1, we look for super- and sub-solutions to the equation solved by ϕ_1^ε with the form $\chi_\pm = \overline{\phi_1^\varepsilon} \pm \psi$. Once this is done, we apply a comparison principle in order to obtain $\chi_- \le \phi_1^\varepsilon \le \chi_+$, which in turns ensures

$$-\frac{1}{2}\rho_0^\varepsilon\,|v-\mathcal{V}^\varepsilon|^2+\varepsilon\,\chi_- \;\le\; \phi^\varepsilon \;\le\; -\frac{1}{2}\rho_0^\varepsilon\,|v-\mathcal{V}^\varepsilon|^2+\varepsilon\,\chi_+.$$

The last step consists in proving that $-\frac{1}{2}\rho_0^\varepsilon\,|v-\mathcal{V}^\varepsilon|^2$ and $\overline{\phi_1^\varepsilon}$ converge as $\varepsilon \to 0$. This is done relying on previous results which ensure that the macroscopic quantities converge (see Theorem 2.4).

Comments and perspectives. Several perspectives arise from our work. First, our result holds in a perturbative setting in the sense that we need the initial data to be concentrated in order for our result to hold true. It would be interesting to treat general initial data without requiring any well-preparedness condition. To achieve this, one possibility would be to adapt the strategy adopted in [6], where was introduced a time dependent scaling, taking into account the initial layer induced by ill-prepared initial data.

Another natural continuation is to describe the limiting dynamics of the Hopf-Cole exponent ϕ^ε with respect to the adaptation variable w. This is challenging since equation (1.4) is not singular in ε with respect to w, meaning that limiting dynamics with respect to the adaptation variable correspond to fluctuations of ϕ^ε of order $O(\varepsilon)$. However, this article provides a glimpse of the difficulties and potential outcomes related to such analysis. Indeed, referring to the beginning of Section 3, the ε-correction of ϕ^ε should be close to

$$\overline{\phi_1^\varepsilon}(t,\boldsymbol{x},\boldsymbol{u}) + \overline{\psi_1^\varepsilon}(t,\boldsymbol{x},w),$$

as ε vanishes, where $\overline{\phi_1^\varepsilon}$ is defined by (3.3). The term $\overline{\psi_1^\varepsilon}$ gathers ε-fluctuations depending only on w, it was already derived in [6, 7, Sections 1.2] where both weak and Lebesgue convergence were rigorously proved. However, we surprisingly obtain additional cross terms between v and w as well as higher order terms with respect to v, gathered in $\overline{\phi_1^\varepsilon}$, which at least to our knowledge, were not known in the literature. This assesses the potential of the Hopf-Cole approach to provide new insights for the model at hand. However, it also highlights the difficulties attached to this approach: its precision and rigidity lead to intricate and technical analysis whereas other methods focusing rather on the convergence at the level of densities yield coarser convergence result but are easier to implement.

Structure of the paper. The remaining part of this article is organized as follows: in Section 2, we prove some regularity estimates for equation (1.4) in order to make our further computations rigorous: this is the object of Lemma 2.1 and Proposition 2.2. In Theorem 2.4 and Corollary 2.5, we also we recall and prove some convergence results on the macroscopic quantities $\mathcal{U}^\varepsilon$ and $\mathcal{E}(f^\varepsilon)$. Then we pass to Section 3, which is dedicated to the proof of Theorem 1.5. The proof relies on the key Lemma 3.1, in which we construct sub- and super-solution for equation (1.11) on ϕ^ε.

2. Preliminary estimates

First, we prove non-uniform in ε regularity estimates for the solutions to (1.4) in order to make later computations rigorous. Second, we recall uniform in ε moment estimates for the solutions to (1.4). We start with the following lemma, in which we prove regularity results for the macroscopic quantities $\mathcal{U}^\varepsilon$ and $\mathcal{E}(f^\varepsilon)$

Lemma 2.1. *Consider the solution f^ε to (1.4) provided by Theorem 1.4. For all function $\varphi \in \mathscr{C}^2\left(\mathbb{R}^2\right)$ with polynomial growth of order $l \geq 0$, that is*

$$|\varphi(\boldsymbol{u})| + |\nabla_{\boldsymbol{u}}\varphi(\boldsymbol{u})| + \left|\nabla^2_{\boldsymbol{u}}\varphi(\boldsymbol{u})\right| \underset{|\boldsymbol{u}|\to+\infty}{=} O\left(|\boldsymbol{u}|^l\right),$$

the function $\left((t,\boldsymbol{x}) \mapsto \displaystyle\int_{\mathbb{R}^2} \varphi(\boldsymbol{u}) f^\varepsilon(t,\boldsymbol{x},\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\right)$ is continuous and has continuous time derivative over $\mathbb{R}^+ \times K$. In particular, the macroscopic quantities $\mathcal{V}^\varepsilon$ and $\mathcal{W}^\varepsilon$ given by (1.3) and the error $\mathcal{E}(f^\varepsilon)$ given by (1.8) are continuous and have continuous time derivatives.

Proof. Consider such function φ. To simplify notations we write

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

We start by proving that $\varphi(f^\varepsilon)$ is continuous. To do so, we fix some $(t,\boldsymbol{x}) \in \mathbb{R}^+ \times K$ and prove

$$\lim_{(s,\boldsymbol{y})\to(t,\boldsymbol{x})} \varphi(f^\varepsilon)(s,\boldsymbol{y}) = \varphi(f^\varepsilon)(t,\boldsymbol{x}).$$

For all $(s,\boldsymbol{y}) \in \mathbb{R}^+ \times K$, it holds

$$|\varphi(f^\varepsilon)(t,\boldsymbol{x}) - \varphi(f^\varepsilon)(s,\boldsymbol{y})| \leq \int_{\mathbb{R}^2} |\varphi(\boldsymbol{u})|\,|f^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) - f^\varepsilon(s,\boldsymbol{y},\boldsymbol{u})|^{1/2}\,|f^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) - f^\varepsilon(s,\boldsymbol{y},\boldsymbol{u})|^{1/2}\,\mathrm{d}\boldsymbol{u}.$$

Applying Cauchy-Schwarz inequality to the latter estimate, we deduce

$$|\varphi\left(f^\varepsilon\right)(t,\boldsymbol{x})-\varphi\left(f^\varepsilon\right)(s,\boldsymbol{y})|\leq\left(\varphi^2\left(f^\varepsilon\right)(t,\boldsymbol{x})+\varphi^2\left(f^\varepsilon\right)(s,\boldsymbol{y})\right)^{1/2}\left\|f^\varepsilon(t,\boldsymbol{x})-f^\varepsilon(s,\boldsymbol{y})\right\|_{L^1(\mathbb{R}^2)}^{1/2}.$$

According to Theorem 1.4, $\varphi^2\left(f^\varepsilon\right)$ is locally bounded over $\mathbb{R}^+\times K$ since f^ε has exponential moments and φ has polynomial growth. Therefore, we obtain the result since f^ε lies in $\mathscr{C}^0\left(\mathbb{R}^+\times K,L^1\left(\mathbb{R}^2\right)\right)$ according to Definition 1.3.

We now prove that $\varphi\left(f^\varepsilon\right)$ has continuous time derivative. We multiply equation (1.4) by φ and integrate with respect to $\boldsymbol{u}$. After an integration by part, this yields

$$\partial_t\varphi\left(f^\varepsilon\right)=\xi_1\left(f^\varepsilon\right)+\left(\frac{1}{\varepsilon}\rho_0^\varepsilon\mathcal{V}^\varepsilon+\Psi*\left(\rho_0^\varepsilon\,\mathcal{V}^\varepsilon\right)\right)\partial_v\varphi\left(f^\varepsilon\right)-\left(\frac{1}{\varepsilon}\rho_0^\varepsilon+\Psi*\rho_0^\varepsilon\right)\xi_2\left(f^\varepsilon\right),$$

with

$$\begin{cases}\xi_1(\boldsymbol{u})=\partial_v\varphi(\boldsymbol{u})\left(N(v)-w\right)+\partial_w\varphi(\boldsymbol{u})A(\boldsymbol{u})+\partial_v^2\varphi(\boldsymbol{u}),\\ \\ \xi_2(\boldsymbol{u})=\partial_v\varphi(\boldsymbol{u})v.\end{cases}$$

Functions $\xi_1\left(f^\varepsilon\right),\xi_2\left(f^\varepsilon\right),\partial_v\varphi\left(f^\varepsilon\right),\mathcal{V}^\varepsilon$ are continuous according to the previous step. Furthermore, we obtain that functions $\Psi*\left(\rho_0^\varepsilon\,\mathcal{V}^\varepsilon\right)$ and $\Psi*\rho_0^\varepsilon$ are continuous using continuity of $\mathcal{V}^\varepsilon$ and ρ_0^ε and assumption (1.15) on Ψ. This yields the result. □

We prove that when the initial data f_0^ε is smooth, the associated solution f^ε to (1.4) is regular.

Proposition 2.2. *Under the assumptions of Theorem 1.4, suppose in addition that f_0^ε lies in $\mathscr{C}^0\left(K,\mathscr{C}_c^\infty\left(\mathbb{R}^2\right)\right)$ and that N meets the following assumptions*

$$\sup_{|v|\geq1}\left|N'(v)\right||v|^{1-p}<+\infty,\qquad\sup_{|v|\geq1}\left|N'''(v)\right||v|^{-p'}<+\infty,\tag{2.1}$$

where p is given in assumption (1.13) and p' in assumption (1.14). Then the solution f^ε to equation (1.4) provided by Theorem 1.4 verifies

$$f^\varepsilon\in L^\infty_{loc}\left(\mathbb{R}^+\times K,W^{2,1}\left(\mathbb{R}^2\right)\right),\quad\partial_t f^\varepsilon\in L^\infty_{loc}\left(\mathbb{R}^+\times K,L^1\left(\mathbb{R}^2\right)\right).$$

We postpone the proof to Appendix B: it is mainly technical and relies on moment estimates on the derivatives of f^ε.

Remark 2.3. *Assumption (2.1) on N is purely technical but it does not constitute a limitation in our context since it is satisfied by the general class of drifts described below assumptions (1.13)-(1.14), which includes the original FitzHugh-Nagumo model.*

For self-consistency, we recall a result from [7] about the control of the macroscopic quantities $(\mathcal{V}^\varepsilon,\mathcal{W}^\varepsilon)$ defined by (1.3) and the error term $\mathcal{E}\left(f^\varepsilon\right)$ defined by (1.8). We also provide uniform estimates with respect to ε for the moments of f^ε and for the relative energy given by

$$\begin{cases}M_q\left[f^\varepsilon\right](t,\boldsymbol{x}):=\dfrac{1}{\rho_0^\varepsilon(\boldsymbol{x})}\displaystyle\int_{\mathbb{R}^2}|\boldsymbol{u}|^q f^\varepsilon(t,\boldsymbol{x},\boldsymbol{u})\mathrm{d}\boldsymbol{u},\\ \\ D_q\left[f^\varepsilon\right](t,\boldsymbol{x}):=\dfrac{1}{\rho_0^\varepsilon(\boldsymbol{x})}\displaystyle\int_{\mathbb{R}^2}|v-\mathcal{V}^\varepsilon(t,\boldsymbol{x})|^q f^\varepsilon(t,\boldsymbol{x},\boldsymbol{u})\mathrm{d}\boldsymbol{u},\end{cases}\tag{2.2}$$

where $q\geq2$.

Theorem 2.4 ([7])**.** *Under assumptions (1.13)-(1.20) and under the additional assumptions of Theorem 1.4, consider the solutions f^ε to (1.4) provided by Theorem 1.4 and the solution $\mathcal{U}$ to (1.9). Furthermore, define the initial macroscopic error as*

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

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

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

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

where $\mathcal{U}^\varepsilon$ and $\mathcal{U}$ are respectively given by (1.3) and (1.9).

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

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

where exponent p is given in assumption (1.13) and p' in assumption (1.14). In particular, $\mathcal{U}^\varepsilon$, $\partial_t \mathcal{U}^\varepsilon$ and $\mathcal{E}(f^\varepsilon)$ are uniformly bounded with respect to both $(t, \boldsymbol{x}) \in \mathbb{R}^+ \times K$ and ε, where $\mathcal{E}$ is defined by (1.8).

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

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

In this theorem, constants C and ε_0 only depend on m_p, $\overline{m}_p$, m_* (see (1.18)-(1.20)) and on the data of the problem N, A and Ψ.

We deduce from this result an estimate on the derivative of the error term, that will be used later in our proof.

Corollary 2.5. *Under the assumptions of Theorem 2.4 and the additional assumption (1.22b) on the sequence $(f_0^\varepsilon)_{\varepsilon>0}$, there exists a constant $C > 0$ such that*

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(f^\varepsilon(t, \boldsymbol{x}, \cdot))\right| \leq C,$$

for all $\varepsilon > 0$ and all $(t, \boldsymbol{x}) \in \mathbb{R}^+ \times K$, where $\mathcal{E}(f^\varepsilon)$ is the macroscopic error given by (1.8). In this result, constant C only depends on m_p, $\overline{m}_p$, m_* (see (1.18)-(1.20)), on the data of the problem N, A and Ψ and on the constant in assumption (1.22b).

Proof. We compute the time derivative of $\mathcal{E}(f^\varepsilon)$ taking the difference between equation (1.4) multiplied by N/ρ_0^ε and integrated with respect to $\boldsymbol{u}$, and the first line of (1.7) multiplied by $N'(\mathcal{V}^\varepsilon)$. After integrating by part with respect to v, it yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(f^\varepsilon(t, \boldsymbol{x}, \cdot)) = \mathcal{A} + \mathcal{B}, \tag{2.3}$$

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

$$\begin{cases} \mathcal{A} = -\dfrac{1}{\varepsilon}\displaystyle\int_{\mathbb{R}^2} N'(v)\left(v - \mathcal{V}^\varepsilon(t, \boldsymbol{x})\right) f^\varepsilon(t, \boldsymbol{x}, \boldsymbol{u})\mathrm{d}\boldsymbol{u}, \\ \mathcal{B} = \dfrac{1}{\rho_0^\varepsilon}\displaystyle\int_{\mathbb{R}^2}\left[\left(N'B^\varepsilon + N''\right) f^\varepsilon\right](t, \boldsymbol{x}, \boldsymbol{u})\mathrm{d}\boldsymbol{u} - N'(\mathcal{V}^\varepsilon)\left(B^\varepsilon(t, \boldsymbol{x}, \mathcal{U}^\varepsilon) + \mathcal{E}(f^\varepsilon)\right). \end{cases}$$

The main difficulty here consists in estimating the stiffer term $\mathcal{A}$: this is what we start with. According to the definition of $\mathcal{V}^\varepsilon$, we have

$$\mathcal{A} = -\frac{1}{\varepsilon}\int_{\mathbb{R}^2}\left(N'(v) - N'\left(\mathcal{V}^\varepsilon\right)\right)\left(v - \mathcal{V}^\varepsilon\right) f^\varepsilon(t, \boldsymbol{x}, \boldsymbol{u})\mathrm{d}\boldsymbol{u}.$$

To bound the term inside the latter integral, we distinguish two cases. On the one hand, since $N \in \mathscr{C}^2(\mathbb{R})$ and since $\mathcal{V}^\varepsilon$ is uniformly bounded with respect to $(t, \boldsymbol{x}) \in \mathbb{R}^+ \times K$ and $\varepsilon > 0$ (see item (2) in Theorem 2.4), it holds

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

where constant C is the uniform norm of N'' over the ball of radius $2\sup_{\varepsilon>0}\|\mathcal{V}^\varepsilon\|_{L^\infty}$. On the other hand, using assumption (1.14) and since $\mathcal{V}^\varepsilon$ is uniformly bounded, it holds

$$\left(N'(v) - N'\left(\mathcal{V}^\varepsilon\right)\right)\left(v - \mathcal{V}^\varepsilon\right)\mathbb{1}_{|v| > 2\|\mathcal{V}^\varepsilon\|_{L^\infty}} \leq C\left|v - \mathcal{V}^\varepsilon\right|^{p'+1}.$$

Taking the sum between the last two estimates, we obtain the following bound for $\mathcal{A}$

$$\mathcal{A} \leq \frac{C}{\varepsilon}\left(D_{p'+1}[f^\varepsilon] + D_2[f^\varepsilon]\right),$$

where $D_{p'+1}$ and D_2 are given by (2.2). Hence, we apply item (3) of Theorem 2.4 and deduce

$$\mathcal{A} \leq C\left[\left(D_2 + D_{p'+1}\right)[f^\varepsilon](0,\boldsymbol{x})\varepsilon^{-1}\exp\left(-2m_*\frac{t}{\varepsilon}\right) + 1\right],$$

for all $(t,\boldsymbol{x}) \in \mathbb{R}^+ \times K$ and $\varepsilon > 0$. To conclude this step, we use assumption (1.22b) which ensures that: $\left(D_2 + D_{p'+1}\right)[f^\varepsilon](0,\boldsymbol{x}) = O(\varepsilon)$ as $\varepsilon \to 0$. Therefore, we deduce $\mathcal{A} \leq C$, for some constant C independent of $(t,\boldsymbol{x}) \in \mathbb{R}^+ \times K$ and ε.

Let us now estimate $\mathcal{B}$. According to assumptions (1.16) and (1.18) on Ψ and ρ_0^ε, $\Psi * \rho_0^\varepsilon$ is uniformly bounded with respect to both $\boldsymbol{x} \in K$ and $\varepsilon > 0$. Furthermore, N, N' and N'' have polynomial growth according to (1.13)-(1.14). Hence, we may bound $\mathcal{B}$ by moments of f^ε up to order $2(p+p')$, which are themselves uniformly bounded with respect to both $(t,\boldsymbol{x}) \in \mathbb{R}^+ \times K$ and $\varepsilon > 0$, according to item (2) of Theorem 2.4. Therefore, there exists a constant C such that

$$\mathcal{B} \leq C,$$

for all $(t,\boldsymbol{x}) \in \mathbb{R}^+ \times K$ and $\varepsilon > 0$.

We obtain the expected result gathering the estimates obtained on $\mathcal{A}$ and $\mathcal{B}$ and taking the absolute value in (2.3). □

3. Proof of Theorem 1.5

In this section, we derive uniform L^∞ convergence estimates for the solution ϕ^ε to equation (1.11). To do so, our strategy consists in performing a Hilbert expansion of ϕ^ε with respect to ε and to prove that the higher order terms are uniformly bounded with respect to ε. Denote by ϕ_1^ε the correction of order 1 in the expansion of ϕ^ε with respect to ε

$$\phi^\varepsilon = -\frac{1}{2}\rho_0^\varepsilon\left|v - \mathcal{V}^\varepsilon\right|^2 + \varepsilon\phi_1^\varepsilon. \tag{3.1}$$

Plugging this *ansatz* in (1.11), we find that ϕ_1^ε solves the following equation

$$\partial_t\phi_1^\varepsilon + \nabla_{\boldsymbol{u}}\phi_1^\varepsilon \cdot \mathbf{b}^\varepsilon + \operatorname{div}_{\boldsymbol{u}}\left[\mathbf{b}^\varepsilon\right] - \partial_v^2\phi_1^\varepsilon - \left|\partial_v\phi_1^\varepsilon\right|^2 + \frac{1}{\varepsilon}\rho_0^\varepsilon\left(v - \mathcal{V}^\varepsilon\right)\partial_v\left(\phi_1^\varepsilon - \overline{\phi_1^\varepsilon}\right) = 0, \tag{3.2}$$

where $\overline{\phi_1^\varepsilon}$ is given by

$$\overline{\phi_1^\varepsilon}(t,\boldsymbol{x},\boldsymbol{u}) = n(v) - n(\mathcal{V}^\varepsilon) - (v - \mathcal{V}^\varepsilon)\left(N(\mathcal{V}^\varepsilon) + (w - \mathcal{W}^\varepsilon) + \mathcal{E}(f^\varepsilon) + \frac{1}{2}\Psi * \rho_0^\varepsilon(v - \mathcal{V}^\varepsilon)\right), \tag{3.3}$$

and where n is the primitive of N defined in Theorem 1.5. Keeping the leading order, we expect that ϕ_1^ε will look like $\overline{\phi_1^\varepsilon}$ as $\varepsilon \to 0$. Therefore, we look for sub and super-solutions to equation (3.2) with the form $\overline{\phi_1^\varepsilon} + \psi$, where ψ needs to be determined. This is done in the following lemma, which constitutes the keystone of our analysis.

Lemma 3.1. *Consider some positive constant α_0 and define ψ as follows*

$$\psi(t,\boldsymbol{x},\boldsymbol{u}) = \frac{\alpha_0}{2}\left|v - \mathcal{V}^\varepsilon(t,\boldsymbol{x})\right|^2 + \frac{\alpha(t)}{2}\left|w - \mathcal{W}^\varepsilon(t,\boldsymbol{x})\right|^2,$$

where α is given by

$$\alpha(t) = \alpha_0 e^{2(|a|+b)t} + \frac{1}{|a|+b}\left(e^{2(|a|+b)t} - 1\right),$$

where (a,b) are the coefficients in the definition (1.5) of A. The functions

$$\chi_+ = \overline{\phi_1^\varepsilon} + \psi + m \quad \text{and} \quad \chi_- = \overline{\phi_1^\varepsilon} - \psi - m$$

*are respectively **super** and **sub**-solutions to equation (3.2), where $(t \mapsto m(t))$ is given by*

$$m(t) = m_0 + C\exp\left(6(a+b)t\right),$$

for all $m_0 \in \mathbb{R}$ and where the constant C only depends on α_0, the constants in (1.18)-(1.20) and (1.22b), and the data of the problem N, A and Ψ.

Proof. In this proof, we fix some $(t, \boldsymbol{x}) \in \mathbb{R}^+ \times K$ and denote by C a generic constant depending only on α_0, the constants in (1.18)-(1.20) and (1.22b), and the data of the problem N, A and Ψ. Furthermore we write $\mathcal{U}^\varepsilon$ instead of $\mathcal{U}^\varepsilon(t, \boldsymbol{x})$ for convenience.

The first step of the proof consists in proving that the following term, obtained when evaluating equation (3.2) in $\overline{\phi_1^\varepsilon}$, is of order 0 with respect to ε

$$\mathcal{A} = \partial_t \overline{\phi_1^\varepsilon} + \nabla_{\boldsymbol{u}} \overline{\phi_1^\varepsilon} \cdot \mathbf{b}^\varepsilon + \mathrm{div}_{\boldsymbol{u}} [\mathbf{b}^\varepsilon] - \partial_v^2 \overline{\phi_1^\varepsilon} - \left| \partial_v \overline{\phi_1^\varepsilon} \right|^2 . \tag{3.4}$$

To this aim, we compute the derivatives of $\overline{\phi_1^\varepsilon}$. To simplify the computations, we rewrite $\overline{\phi_1^\varepsilon}$ in a more convenient way: we consider some $(\tilde{v}, w) \in \mathbb{R}^2$ and take the difference between $B^\varepsilon(t, \boldsymbol{x}, \tilde{v}, w)$ given in (1.5) and the first line of the system (1.7)

$$\left(B^\varepsilon - \partial_t \mathcal{V}^\varepsilon\right)(t, \boldsymbol{x}, \tilde{v}, w) = \left[N(\tilde{v}) - N(\mathcal{V}^\varepsilon) - (w - \mathcal{W}^\varepsilon) - (\tilde{v} - \mathcal{V}^\varepsilon)\ \Psi * \rho_0^\varepsilon - \mathcal{E}(f^\varepsilon)\right](t, \boldsymbol{x}) .$$

Integrating the latter relation between $\mathcal{V}^\varepsilon$ and v with respect to $\tilde{v}$, we deduce that $\overline{\phi_1^\varepsilon}$ verifies

$$\overline{\phi_1^\varepsilon}(t, \boldsymbol{x}, \boldsymbol{u}) = \int_{\mathcal{V}^\varepsilon}^{v} \left(B^\varepsilon - \partial_t \mathcal{V}^\varepsilon\right)(t, \boldsymbol{x}, \tilde{v}, w)\ \mathrm{d}\tilde{v} .$$

Using the last two relations, we deduce that for all $\boldsymbol{u} \in \mathbb{R}^2$, it holds

$$\begin{cases} \nabla_{\boldsymbol{u}} \overline{\phi_1^\varepsilon}(t, \boldsymbol{x}, \boldsymbol{u}) &= \begin{pmatrix} B^\varepsilon(t, \boldsymbol{x}, \boldsymbol{u}) - \partial_t \mathcal{V}^\varepsilon \\ -(v - \mathcal{V}^\varepsilon) \end{pmatrix} , \\ \partial_v^2 \overline{\phi_1^\varepsilon}(t, \boldsymbol{x}, \boldsymbol{u}) &= \mathrm{div}_{\boldsymbol{u}} [\mathbf{b}^\varepsilon](\boldsymbol{x}, \boldsymbol{u}) - \partial_w A(\boldsymbol{u}) , \\ \partial_t \overline{\phi_1^\varepsilon}(t, \boldsymbol{x}, \boldsymbol{u}) &= \partial_t \mathcal{V}^\varepsilon \left(\partial_t \mathcal{V}^\varepsilon - B^\varepsilon(t, \boldsymbol{x}, \mathcal{V}^\varepsilon, w)\right) - (v - \mathcal{V}^\varepsilon) \partial_t \left(\partial_t \mathcal{V}^\varepsilon - \Psi * (\rho_0^\varepsilon\, \mathcal{V}^\varepsilon)\right) . \end{cases}$$

Relying on the latter expressions for the derivatives of $\overline{\phi_1^\varepsilon}$ we deduce

$$\mathcal{A} = \partial_t \mathcal{V}^\varepsilon \left(B^\varepsilon(t, \boldsymbol{x}, \boldsymbol{u}) - B^\varepsilon(t, \boldsymbol{x}, (\mathcal{V}^\varepsilon, w))\right) - (v - \mathcal{V}^\varepsilon) \partial_t \left(\partial_t \mathcal{V}^\varepsilon - \Psi * (\rho_0^\varepsilon\, \mathcal{V}^\varepsilon)\right) + \partial_w A(\boldsymbol{u}) - (v - \mathcal{V}^\varepsilon) A(\boldsymbol{u}) .$$

On the one hand, since B^ε is given by (1.5), it holds

$$B^\varepsilon(t, \boldsymbol{x}, \boldsymbol{u}) - B^\varepsilon(t, \boldsymbol{x}, (\mathcal{V}^\varepsilon, w)) = N(v) - N(\mathcal{V}^\varepsilon) - \Psi * \rho_0^\varepsilon (v - \mathcal{V}^\varepsilon) .$$

On the other hand, taking the time derivative in the first line of equation (1.7) we obtain

$$\partial_t \left(\partial_t \mathcal{V}^\varepsilon - \Psi * (\rho_0^\varepsilon\, \mathcal{V}^\varepsilon)\right) = \partial_t\, \mathcal{V}^\varepsilon \left(N'(\mathcal{V}^\varepsilon) - \Psi * \rho_0^\varepsilon\right) - A(\mathcal{U}^\varepsilon) + \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}(f^\varepsilon) .$$

Therefore, $\mathcal{A}$ may be expressed as follows

$$\mathcal{A} = \partial_t \mathcal{V}^\varepsilon \left(N(v) - N(\mathcal{V}^\varepsilon)\right) - (v - \mathcal{V}^\varepsilon) \left(\partial_t \mathcal{V}^\varepsilon N'(\mathcal{V}^\varepsilon) + (A(\boldsymbol{u}) - A(\mathcal{U}^\varepsilon)) + \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}(f^\varepsilon)\right) + \partial_w A(\boldsymbol{u}) .$$

On the one hand, $\partial_t \mathcal{V}^\varepsilon$ and $\mathcal{V}^\varepsilon$ are uniformly bounded according to item (2) in Theorem 2.4. On the other hand, we apply Corollary 2.5 which ensures that under the smallness assumption (1.22b), the time derivative of $\mathcal{E}(f^\varepsilon)$ is uniformly bounded as well. Consequently, for all positive ε, it holds

$$\left(|\partial_t \mathcal{V}^\varepsilon| + |\mathcal{U}^\varepsilon| + \left|\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}(f^\varepsilon)\right|\right)(s, \boldsymbol{y}) \leq C , \quad \forall (s, \boldsymbol{y}) \in \mathbb{R}^+ \times K ,$$

for some constant C depending only on the constants in assumptions (1.18)-(1.20), (1.22b) and on the data of the problem N, A and Ψ. Therefore, we deduce the following bound for $\mathcal{A}$

$$|\mathcal{A}| \leq C \left(1 + |v - \mathcal{V}^\varepsilon| + |v - \mathcal{V}^\varepsilon|^2 + |N(v) - N(\mathcal{V}^\varepsilon)|\right) + b\ |(v - \mathcal{V}^\varepsilon)(w - \mathcal{W}^\varepsilon)| .$$

Then we apply Young's inequality to bound the crossed term between v and w and use assumption (1.13) to bound $N(v)$. In the end, it yields

$$|\mathcal{A}| \leq C\,(1+|v-\mathcal{V}^\varepsilon|^p) + \frac{1}{2}\,|w-\mathcal{W}^\varepsilon|^2\,. \tag{3.5}$$

Building on this estimate, we can now pass to the heart of the proof and show that χ_+ and χ_- are respectively super- and sub-solutions to (3.2). We evaluate equation (3.2) in χ_+ and χ_- and obtain

$$\partial_t\chi_\pm + \nabla_{\boldsymbol{u}}\chi_\pm\cdot\mathbf{b}^\varepsilon + \mathrm{div}_{\boldsymbol{u}}\left[\mathbf{b}^\varepsilon\right] - \partial_v^2\chi_\pm - |\partial_v\chi_\pm|^2 + \frac{1}{\varepsilon}\,\rho_0^\varepsilon\,(v-\mathcal{V}^\varepsilon)\,\partial_v\chi_\pm = \mathcal{A}\pm\mathcal{B}\pm m'(t) - |\partial_v\psi|^2\,,$$

where $\mathcal{A}$ is given by (3.4) and $\mathcal{B}$ is given by

$$\mathcal{B} \,=\, \partial_t\psi + \nabla_{\boldsymbol{u}}\psi\cdot\mathbf{b}^\varepsilon - \partial_v^2\psi - 2\,\partial_v\psi\,\partial_v\overline{\phi_1^\varepsilon} + \frac{1}{\varepsilon}\,\rho_0^\varepsilon\,(v-\mathcal{V}^\varepsilon)\,\partial_v\psi\,. \tag{3.6}$$

In order to conclude that χ_+ and χ_- are respectively sub- and super-solutions to (3.2), it is sufficient to prove

$$\mathcal{B} + m'(t) - |\partial_v\psi|^2 - |\mathcal{A}| \geq 0\,.$$

Therefore, we focus on proving the latter inequality. To begin with, we have

$$\begin{aligned}
&\mathcal{B} - |\partial_v\psi|^2\\
=&-\alpha_0\,(v-\mathcal{V}^\varepsilon)\,\partial_t\mathcal{V}^\varepsilon + \frac{\alpha'(t)}{2}\,|w-\mathcal{W}^\varepsilon|^2 - \alpha(t)\,(w-\mathcal{W}^\varepsilon)\,\partial_t\mathcal{W}^\varepsilon\\
&+\alpha_0\,(v-\mathcal{V}^\varepsilon)\,B^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) + \alpha(t)\,(w-\mathcal{W}^\varepsilon)\,A(\boldsymbol{u}) - \alpha_0 - 2\alpha_0\,(v-\mathcal{V}^\varepsilon)\,(B^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) - \partial_t\mathcal{V}^\varepsilon)\\
&+\frac{\rho_0^\varepsilon}{\varepsilon}\alpha_0\,|v-\mathcal{V}^\varepsilon|^2 - \alpha_0^2\,|v-\mathcal{V}^\varepsilon|^2\\
=&\,\frac{\alpha'(t)}{2}\,|w-\mathcal{W}^\varepsilon|^2 - \alpha_0\,(v-\mathcal{V}^\varepsilon)\,(B^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) - \partial_t\,\mathcal{V}^\varepsilon)\\
&+\alpha(t)\,(w-\mathcal{W}^\varepsilon)\,(A(\boldsymbol{u}) - A(\mathcal{U}^\varepsilon)) - \alpha_0 + \left(\frac{\rho_0^\varepsilon}{\varepsilon}\alpha_0 - \alpha_0^2\right)|v-\mathcal{V}^\varepsilon|^2\\
=&\frac{\alpha'(t)}{2}\,|w-\mathcal{W}^\varepsilon|^2 - \alpha_0\,(v-\mathcal{V}^\varepsilon)\,(B^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) - B^\varepsilon(t,\boldsymbol{x},\mathcal{U}^\varepsilon) - \mathcal{E}(f^\varepsilon))\\
&+\alpha(t)\,(w-\mathcal{W}^\varepsilon)\,(a\,(v-\mathcal{V}^\varepsilon) - b\,(w-\mathcal{W}^\varepsilon)) - \alpha_0 + \left(\frac{\rho_0^\varepsilon}{\varepsilon}\alpha_0 - \alpha_0^2\right)|v-\mathcal{V}^\varepsilon|^2\\
\geq&\left(\frac{\alpha'(t)}{2} - \alpha(t)(|a|+b)\right)|w-\mathcal{W}^\varepsilon|^2 - \alpha_0\,(v-\mathcal{V}^\varepsilon)\,(B^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) - B^\varepsilon(t,\boldsymbol{x},\mathcal{U}^\varepsilon) - \mathcal{E}(f^\varepsilon))\\
&-\alpha_0 + \left(\frac{\rho_0^\varepsilon}{\varepsilon}\alpha_0 - \alpha_0^2 - \frac{|a|}{4}\alpha(t)\right)|v-\mathcal{V}^\varepsilon|^2\,,
\end{aligned}$$

where we have used the Young inequality at the last line. Gathering the latter estimate and (3.5) we obtain

$$\begin{aligned}
&\mathcal{B} + m'(t) - |\partial_v\psi|^2 - |\mathcal{A}|\\
\geq& m'(t) + \left(\frac{\alpha'(t)}{2} - \alpha(t)(|a|+b)\right)|w-\mathcal{W}^\varepsilon|^2 - \alpha_0\,(v-\mathcal{V}^\varepsilon)\,(B^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) - B^\varepsilon(t,\boldsymbol{x},\mathcal{U}^\varepsilon) - \mathcal{E}(f^\varepsilon))\\
&-\alpha_0 + \left(\frac{\rho_0^\varepsilon}{\varepsilon}\alpha_0 - \alpha_0^2 - \frac{|a|}{4}\alpha(t)\right)|v-\mathcal{V}^\varepsilon|^2 - C\,(1+|v-\mathcal{V}^\varepsilon|^p) - |w-\mathcal{W}^\varepsilon|^2\,.\\
\geq& m'(t) + \left(\frac{\alpha'(t)}{2} - \alpha(t)(|a|+b) - 1\right)|w-\mathcal{W}^\varepsilon|^2 - \alpha_0 - \alpha_0\,(v-\mathcal{V}^\varepsilon)\,(N(v) - N(\mathcal{V}^\varepsilon))\\
&+\left(\alpha_0\Psi * \rho_0^\varepsilon(\boldsymbol{x}) + \frac{\rho_0^\varepsilon}{\varepsilon}\alpha_0 - \frac{3}{2}\,\alpha_0^2 - \frac{|a|}{4}\alpha(t) - \left(|a| + \frac{b}{2} + \frac{1}{2}\right) - \frac{1}{2}\,\alpha_0\right)|v-\mathcal{V}^\varepsilon|^2\\
&+\alpha_0\,(v-\mathcal{V}^\varepsilon)\,\mathcal{E}(f^\varepsilon) - C\,(1+|v-\mathcal{V}^\varepsilon|^p)\,,
\end{aligned}$$

where we have used Young inequality and the following relation

$$B^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) - B^\varepsilon(t,\boldsymbol{x},\mathcal{U}^\varepsilon) = N(v) - N(\mathcal{V}^\varepsilon) - (w - \mathcal{W}^\varepsilon) - \Psi * \rho_0^\varepsilon(\boldsymbol{x})(v - \mathcal{V}^\varepsilon)\,.$$

To control the contribution of the term $|w - \mathcal{W}^\varepsilon|^2$ in the latter expression, we choose $\alpha(t)$ such that $\frac{\alpha'(t)}{2} - \alpha(t)(|a| + b) - 1 = 0$, that is

$$\alpha(t) = \alpha_0 \exp\left(2(|a| + b)t\right) + \frac{1}{|a| + b}\left(\exp\left(2(|a| + b)t\right) - 1\right).$$

Furthermore, since $\mathcal{E}(f^\varepsilon)$, ρ_0^ε and $\Psi * \rho_0^\varepsilon$ are uniformly bounded according to (respectively) item (2) in Theorem 2.4, assumptions (1.18) and (1.16), we deduce

$$\begin{aligned}
&\mathcal{B} + m'(t) - |\partial_v \psi|^2 - |\mathcal{A}| \\
&\geq m'(t) - \alpha_0\,(v - \mathcal{V}^\varepsilon)\,(N(v) - N(\mathcal{V}^\varepsilon)) - C\,\Big(1 + \exp\left(2\,(|a| + b)\,t\right)|v - \mathcal{V}^\varepsilon|^2 + |v - \mathcal{V}^\varepsilon|^p\Big)\,,
\end{aligned}$$

for some constant $C > 0$. To control the terms of order $|v - \mathcal{V}^\varepsilon|^2$ and $|v - \mathcal{V}^\varepsilon|^p$ in the latter expression, we rely on the confining property (1.13) of N, which ensures

$$(v - \mathcal{V}^\varepsilon)\,(N(v) - N(\mathcal{V}^\varepsilon)) \leq C - \frac{1}{C}\,|v - \mathcal{V}^\varepsilon|^{p+1}\,,$$

for some constant C great enough. Hence, we obtain

$$\begin{aligned}
&\mathcal{B} + m'(t) - |\partial_v \psi|^2 - |\mathcal{A}| \\
&\geq m'(t) + \frac{1}{C}\,|v - \mathcal{V}^\varepsilon|^{p+1} - C\,\Big(1 + \exp\left(2\,(|a| + b)\,t\right)|v - \mathcal{V}^\varepsilon|^2 + |v - \mathcal{V}^\varepsilon|^p\Big)\,.
\end{aligned}$$

Then we find that

$$\frac{1}{C}\,|v - \mathcal{V}^\varepsilon|^{p+1} - C\,\Big(1 + \exp\left(2\,(|a| + b)\,t\right)|v - \mathcal{V}^\varepsilon|^2 + |v - \mathcal{V}^\varepsilon|^p\Big) \geq -\tilde{C}\,\exp\left(6\,(|a| + b)\,t\right).$$

Hence choosing the function $(t \mapsto m(t))$ such that

$$m'(t) = \tilde{C}\,\exp\left(6\,(|a| + b)\,t\right),$$

we obtain

$$\mathcal{B} + m'(t) - |\partial_v \psi|^2 - |\mathcal{A}| \geq 0,$$

which concludes the proof. □

We are now able to proceed to the proof of Theorem 1.5. Indeed, relying on Lemma 3.1 and applying a comparison principle to equation (3.2), we deduce convergence estimates for the Hopf-Cole transform ϕ^ε of f^ε.

Proof of Theorem 1.5. All along this proof, we consider some positive constants α_0 (to be determined later on) and we work with the associated quantities ψ, χ_+ and χ_- defined in Proposition 3.1. We proceed in three steps

(1) we prove that under our set of assumptions, it holds uniformly in ε

$$\chi_-(0,\boldsymbol{x},\boldsymbol{u}) \leq \phi_1^\varepsilon(0,\boldsymbol{x},\boldsymbol{u}) \leq \chi_+(0,\boldsymbol{x},\boldsymbol{u}), \quad \forall (\boldsymbol{x},\boldsymbol{u}) \in K \times \mathbb{R}^2,$$

where χ_+ and χ_- are defined in Lemma 3.1,

(2) we apply Lemma 3.1 and prove a comparison principle to deduce that the latter inequality holds for all positive time, that is

$$\chi_-(t,\boldsymbol{x},\boldsymbol{u}) \leq \phi_1^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) \leq \chi_+(t,\boldsymbol{x},\boldsymbol{u}), \quad \forall (t,\boldsymbol{x}) \in \mathbb{R}^+ \times K, \text{ a.e. in } \boldsymbol{u} \in \mathbb{R}^2,$$

(3) we conclude that ϕ^ε converges locally uniformly to $-\frac{1}{2}\rho_0\,|v - \mathcal{V}|^2$.

We start with step (1). Dividing (1.22a) by ε and replacing ϕ^ε with ϕ_1^ε according to (3.1) we obtain the following bound for all positive ε

$$n(v) - C\left(1+|\boldsymbol{u}|^2\right) \leq \left(\phi_1^\varepsilon + \frac{1}{\varepsilon}\left(-\frac{1}{2}\rho_0^\varepsilon |v-\mathcal{V}^\varepsilon|^2 + \frac{1}{2}\rho_0 |v-\mathcal{V}|^2\right)\right)(0,\boldsymbol{x},\boldsymbol{u}) \leq n(v) + C\left(1+|\boldsymbol{u}|^2\right),$$

for all $(\boldsymbol{x},\boldsymbol{u}) \in K \times \mathbb{R}^2$. On the one hand, according to assumptions (1.18), (1.19) and (1.21), it holds

$$\frac{1}{\varepsilon}\left|-\frac{1}{2}\rho_0^\varepsilon |v-\mathcal{V}_0^\varepsilon(\boldsymbol{x})|^2 + \frac{1}{2}\rho_0 |v-\mathcal{V}_0(\boldsymbol{x})|^2\right| \leq C\left(1+|\boldsymbol{u}-\mathcal{U}_0^\varepsilon(\boldsymbol{x})|^2\right),$$

for all $(\boldsymbol{x},\boldsymbol{u}) \in K\times\mathbb{R}^2$, for some constant C depending only on the initial condition f_0^ε (only through the constants appearing in assumptions (1.18), (1.19) and (1.21)). On the other hand, according to assumptions (1.16) and (1.18), $\Psi * \rho_0^\varepsilon$ is uniformly bounded with respect to both $\boldsymbol{x} \in K$ and $\varepsilon > 0$. On top of that, $\mathcal{U}_0^\varepsilon$ and $\mathcal{E}\left(f_0^\varepsilon\right)$ are also uniformly bounded with respect to both $\boldsymbol{x} \in K$ and $\varepsilon > 0$ according to assumptions (1.13) and (1.19). Therefore, according to the definition (3.3) of $\overline{\phi_1^\varepsilon}$, it holds

$$\left|n - \overline{\phi_1^\varepsilon}\right|(0,\boldsymbol{x},\boldsymbol{u}) \leq C\left(1+|\boldsymbol{u}-\mathcal{U}_0^\varepsilon(\boldsymbol{x})|^2\right),$$

for all $(\boldsymbol{x},\boldsymbol{u}) \in K \times \mathbb{R}^2$, for some constant C depending on the initial condition f_0^ε (only through the constants appearing in assumptions (1.18)-(1.19)) and N. Gathering these considerations and writing

$$\phi_1^\varepsilon - \overline{\phi_1^\varepsilon} = \frac{1}{\varepsilon}\left(\phi^\varepsilon + \frac{1}{2}\rho_0 |v-\mathcal{V}|^2 - \varepsilon n\right) + \frac{1}{\varepsilon}\left(-\frac{1}{2}\rho_0 |v-\mathcal{V}|^2 + \frac{1}{2}\rho_0^\varepsilon |v-\mathcal{V}^\varepsilon|^2\right) + n - \overline{\phi_1^\varepsilon},$$

we deduce that according to assumption (1.22a), for all positive ε, it holds

$$\overline{\phi_1^\varepsilon}(0,\boldsymbol{x},\boldsymbol{u}) - C\left(1+|\boldsymbol{u}-\mathcal{U}_0^\varepsilon(\boldsymbol{x})|^2\right) \leq \phi_1^\varepsilon(0,\boldsymbol{x},\boldsymbol{u}) \leq \overline{\phi_1^\varepsilon}(0,\boldsymbol{x},\boldsymbol{u}) + C\left(1+|\boldsymbol{u}-\mathcal{U}_0^\varepsilon(\boldsymbol{x})|^2\right),$$

for all $(\boldsymbol{x},\boldsymbol{u}) \in K \times \mathbb{R}^2$. Therefore, taking $\alpha_0/2$ and $m(0)$ greater than C, we conclude step (1), indeed for all positive ε it holds

$$\chi_-(0,\boldsymbol{x},\boldsymbol{u}) \leq \phi_1^\varepsilon(0,\boldsymbol{x},\boldsymbol{u}) \leq \chi_+(0,\boldsymbol{x},\boldsymbol{u}), \quad \forall (\boldsymbol{x},\boldsymbol{u}) \in K \times \mathbb{R}^2,$$

where χ_+ and χ_- are given in Proposition 3.1.

Let us now turn to step (2), which consists in proving that the latter estimate holds true for all positive time by applying a comparison principle. For technical reasons detailed in Appendix A, we apply the comparison principle for the following linearized version of the kinetic equation (1.4), instead of working directly on equation (3.2)

$$\partial_t f + \operatorname{div}_{\boldsymbol{u}}\left[\mathbf{b}^\varepsilon f\right] - \partial_v^2 f = \frac{1}{\varepsilon}\rho_0^\varepsilon \partial_v\left[(v-\mathcal{V}^\varepsilon)f\right]. \tag{3.7}$$

Therefore, we define f_+ and f_- for all $(t,\boldsymbol{x},\boldsymbol{u}) \in \mathbb{R}^+ \times K \times \mathbb{R}^2$ as follows

$$f_\pm(t,\boldsymbol{x},\boldsymbol{u}) = \sqrt{\frac{\rho_0(\boldsymbol{x})}{2\pi\varepsilon}} \exp\left(-\frac{1}{2\,\varepsilon}\rho_0^\varepsilon |v-\mathcal{V}^\varepsilon(t,\boldsymbol{x})|^2 + \chi_\pm(t,\boldsymbol{x},\boldsymbol{u})\right).$$

We prove that these quantities are classical super- and sub-solutions of (3.7) applying jointly Lemma 3.1, to ensure that χ_- and χ_+ are respectively sub and super-solutions to equation (3.2) and Lemma A.1, which ensures that under the regularity condition $f_\pm, \partial_t f_\pm, \nabla_{\boldsymbol{u}}^2 f_\pm \in \mathscr{C}^0\left(\mathbb{R}^+ \times K \times \mathbb{R}^2\right)$, f_- and f_+ are respectively sub and super-solutions to (3.7) if and only if χ_- and χ_+ are respectively sub and super-solutions to (3.2). To verify the regularity assumption, we apply Lemma 2.1 which ensures that $\mathcal{U}^\varepsilon$ and $\mathcal{E}(f^\varepsilon)$ are continuous and have continuous time derivative over $\mathbb{R}^+ \times K$. Therefore, $f_\pm$, $\partial_t f_\pm$ and $\nabla_{\boldsymbol{u}}^2 f_\pm$ lie in $\mathscr{C}^0\left(\mathbb{R}^+ \times K \times \mathbb{R}^2\right)$. To conclude, we rely on the previous step, which ensures

$$f_-(0,\boldsymbol{x},\boldsymbol{u}) \leq f_0^\varepsilon(\boldsymbol{x},\boldsymbol{u}) \leq f_+(0,\boldsymbol{x},\boldsymbol{u}), \quad \forall (\boldsymbol{x},\boldsymbol{u}) \in K \times \mathbb{R}^2.$$

Therefore, relying on the comparison principle proved in Lemma A.2, we deduce

$$f_-(t,\boldsymbol{x},\boldsymbol{u}) \leq f^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) \leq f_+(t,\boldsymbol{x},\boldsymbol{u}), \quad \forall (t,\boldsymbol{x}) \in \mathbb{R}^+ \times K, \text{a.e. in } \boldsymbol{u} \in \mathbb{R}^2.$$

Taking the logarithm of the latter relation, we deduce that the bound obtained in step (1), propagates through time, that is, for all positive ε, it holds

$$\chi_-(t,\boldsymbol{x},\boldsymbol{u}) \leq \phi_1^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) \leq \chi_+(t,\boldsymbol{x},\boldsymbol{u}), \quad \forall (t,\boldsymbol{x}) \in \mathbb{R}^+ \times K, \text{a.e. in } \boldsymbol{u} \in \mathbb{R}^2.$$

We can now turn to the last step and prove our main result. According to the definition of ϕ_1^ε and the result of step (2), it holds

$$-\frac{1}{2}\rho_0^\varepsilon |v-\mathcal{V}^\varepsilon|^2 + \frac{1}{2}\rho_0 |v-\mathcal{V}|^2 + \varepsilon\chi_- \leq \phi^\varepsilon + \frac{1}{2}\rho_0 |v-\mathcal{V}|^2 \leq -\frac{1}{2}\rho_0^\varepsilon |v-\mathcal{V}^\varepsilon|^2 + \frac{1}{2}\rho_0 |v-\mathcal{V}|^2 + \varepsilon\chi_+,$$

for all $\forall (t,\boldsymbol{x}) \in \mathbb{R}^+ \times K$, a.e. in $\boldsymbol{u} \in \mathbb{R}^2$. On the one hand, relying on item (1) in Theorem 2.4 and since the initial condition f_0^ε meets the compatibility assumption (1.21), there exists two positive constants C and ε_0 such that for all $\varepsilon \leq \varepsilon_0$, it holds

$$\left| -\frac{1}{2}\rho_0^\varepsilon |v-\mathcal{V}^\varepsilon|^2 + \frac{1}{2}\rho_0 |v-\mathcal{V}|^2 \right| (t,\boldsymbol{x},\boldsymbol{u}) \leq \varepsilon C e^{Ct} \left(1+|\boldsymbol{u}|^2\right),$$

for all $(t,\boldsymbol{x},\boldsymbol{u}) \in \mathbb{R}^+ \times K \times \mathbb{R}^2$, where constants C and ε_0 only depends on the data of our problem: f_0^ε (only through the constants appearing in assumptions (1.18)-(1.20) and (1.21)), N, A and Ψ. On the other hand, according to assumption (1.18) and (1.16), $\Psi * \rho_0^\varepsilon$ is uniformly bounded with respect to both $\boldsymbol{x} \in K$ and $\varepsilon > 0$. On top of that, $\mathcal{U}^\varepsilon$ and $\mathcal{E}(f^\varepsilon)$ are also uniformly bounded with respect to both $(t,\boldsymbol{x}) \in \mathbb{R}^+ \times K$ and $\varepsilon > 0$ according item (2) in Theorem 2.4. Therefore, according to the definitions of χ_- and χ_+ (see Lemma 3.1) and since N is continuous, it holds

$$|\chi_\pm - n|(t,\boldsymbol{x},\boldsymbol{u}) \leq C e^{Ct}\left(1+|\boldsymbol{u}|^2\right),$$

for all $(t,\boldsymbol{x},\boldsymbol{u}) \in K \times \mathbb{R}^2$, where the constant C only depends on the data of our problem: f_0^ε (only through the constants appearing in assumptions (1.18)-(1.20)), N, A and Ψ. Hence, we deduce the result: for all $\varepsilon \leq \varepsilon_0$ it holds

$$\varepsilon\left(n(v) - Ce^{Ct}\left(1+|\boldsymbol{u}|^2\right)\right) \leq \left(\phi^\varepsilon + \frac{1}{2}\rho_0 |v-\mathcal{V}|^2\right)(t,\boldsymbol{x},\boldsymbol{u}) \leq \varepsilon\left(n(v) + Ce^{Ct}\left(1+|\boldsymbol{u}|^2\right)\right),$$

for all $\forall (t,\boldsymbol{x}) \in \mathbb{R}^+ \times K$, a.e. in $\boldsymbol{u} \in \mathbb{R}^2$, where constants C and ε_0 only depends on the data of our problem: f_0^ε (only through the constants appearing in assumptions (1.18)-(1.20) and (1.21)-(1.22b)), N, A and Ψ. □

Acknowledgment

The authors thank warmly Francis Filbet all the discussions without which it would not have been possible to achieve this work, Philippe Laurençot for his enlightening suggestions and comments and Sepideh Mirrahimi for her precious explanations regarding Hamilton-Jacobi equations. A.B. gratefully acknowledge the support of ANITI (Artificial and Natural Intelligence Toulouse Institute). This project has received support from ANR ChaMaNe No: ANR-19-CE40-0024.

Appendix A. Comparison principles

The object of this section is to prove a comparison principle for equation (3.2) in order to complete step (2) in the proof of Theorem 1.5. More precisely, we prove that if the quantity ϕ_1^ε defined by (3.1) verifies

$$\chi_-(0,\boldsymbol{x},\boldsymbol{u}) \leq \phi_1^\varepsilon(0,\boldsymbol{x},\boldsymbol{u}) \leq \chi_+(0,\boldsymbol{x},\boldsymbol{u}), \quad \forall (\boldsymbol{x},\boldsymbol{u}) \in K \times \mathbb{R}^2,$$

where χ_- an χ_+ are respectively sub and super-solutions to (3.2), then the latter estimate propagates through time, that is

$$\chi_-(t,\boldsymbol{x},\boldsymbol{u}) \leq \phi_1^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) \leq \chi_+(t,\boldsymbol{x},\boldsymbol{u}), \quad \forall (t,\boldsymbol{x},\boldsymbol{u}) \in \mathbb{R}^+ \times K \times \mathbb{R}^2.$$

Instead of working directly on equation (3.2), our strategy consists in proving a comparison principle for the linearized version (3.7) of the kinetic equation (1.4). Indeed, it is more convenient to work on equation (3.7) since we can rely on the decaying properties of solutions to (1.4) provided by Theorem 1.4. From the comparison principle on equation (3.7), we will easily deduce the expected

result. This approach is made possible since, according to the following lemma, there is a direct link between sub- and super-solutions to equations (3.7) and (3.2)

Lemma A.1. *Consider some fixed $\varepsilon > 0$. Under the assumptions of Theorem 1.4, consider the solution f^ε to equation (1.4) and its associated macroscopic quantities $(\mathcal{V}^\varepsilon, \mathcal{W}^\varepsilon)$ provided by Theorem 1.4. Furthermore, consider a strictly positive function f such that f, $\partial_t f$ and $\nabla^2_{\boldsymbol{u}} f$ lie in $\mathscr{C}^0\left(\mathbb{R}^+ \times K \times \mathbb{R}^2\right)$ and define χ as follows*

$$f(t,\boldsymbol{x},\boldsymbol{u}) = \sqrt{\frac{\rho_0(\boldsymbol{x})}{2\pi\varepsilon}} \exp\left(\left(-\frac{1}{2\varepsilon}\rho_0^\varepsilon \left|v - \mathcal{V}^\varepsilon\right|^2 + \chi\right)(t,\boldsymbol{x},\boldsymbol{u})\right), \quad \forall (t,\boldsymbol{x},\boldsymbol{u}) \in \mathbb{R}^+ \times K \times \mathbb{R}^2.$$

Then the following statements are equivalent

(1) f is a super-solution (resp. sub-solution) to (3.7).
(2) χ is a super-solution (resp. sub-solution) to (3.2).

Proof. We consider f and χ as in Lemma A.1 and proceed in two steps. On the one hand, plugging f in equation (3.7), one has the following relation

$$\partial_t f + \mathrm{div}_{\boldsymbol{u}}\left[\mathbf{b}^\varepsilon f\right] - \partial_v^2 f - \frac{1}{\varepsilon}\rho_0^\varepsilon \partial_v\left[(v - \mathcal{V}^\varepsilon) f\right] = \frac{1}{\varepsilon} f\,\mathcal{A}\,,$$

for all $(t,\boldsymbol{x},\boldsymbol{u}) \in \mathbb{R}^+ \times K \times \mathbb{R}^2$ where $\mathcal{A}$ gathers the terms obtained plugging $\phi := -\frac{1}{2}\rho_0^\varepsilon \left|v - \mathcal{V}^\varepsilon\right|^2 + \varepsilon\chi$ into equation (1.11), that is

$$\mathcal{A} = \partial_t \phi + \nabla_{\boldsymbol{u}}\phi \cdot \mathbf{b}^\varepsilon + \varepsilon \mathrm{div}_{\boldsymbol{u}}\left[\mathbf{b}\right] - \partial_v^2 \phi - \rho_0^\varepsilon - \frac{1}{\varepsilon}\left(\partial_v\left(\frac{1}{2}\rho_0^\varepsilon\left|v - \mathcal{V}^\varepsilon\right|^2 + \phi\right)\partial_v \phi\right).$$

On the other hand, according to computations already detailed at the beginning of Section 3, $\mathcal{A}$ also corresponds to the terms obtained plugging χ into equation (3.2), that is

$$\mathcal{A} = \varepsilon\left(\partial_t \chi + \nabla_{\boldsymbol{u}}\chi \cdot \mathbf{b}^\varepsilon + \mathrm{div}_{\boldsymbol{u}}\left[\mathbf{b}^\varepsilon\right] - \partial_v^2 \chi - \left|\partial_v \chi\right|^2 + \frac{1}{\varepsilon}\rho_0^\varepsilon\left(v - \mathcal{V}^\varepsilon\right)\partial_v\left(\chi - \overline{\phi_1^\varepsilon}\right)\right),$$

for all $(t,\boldsymbol{x},\boldsymbol{u}) \in \mathbb{R}^+ \times K \times \mathbb{R}^2$. Therefore, we deduce

$$\partial_t f + \mathrm{div}_{\boldsymbol{u}}\left[\mathbf{b}^\varepsilon f\right] - \partial_v^2 f - \frac{1}{\varepsilon}\rho_0^\varepsilon \partial_v\left[(v - \mathcal{V}^\varepsilon) f\right] =$$

$$f\left(\partial_t \chi + \nabla_{\boldsymbol{u}}\chi \cdot \mathbf{b}^\varepsilon + \mathrm{div}_{\boldsymbol{u}}\left[\mathbf{b}^\varepsilon\right] - \partial_v^2 \chi - \left|\partial_v \chi\right|^2 + \frac{1}{\varepsilon}\rho_0^\varepsilon\left(v - \mathcal{V}^\varepsilon\right)\partial_v\left(\chi - \overline{\phi_1^\varepsilon}\right)\right),$$

which yields the result since f is strictly positive. □

It is now left to prove that a comparison principle holds for equation (3.7). It is the object of the following Lemma

Lemma A.2. *Consider some fixed $\varepsilon > 0$. Under the assumptions of Theorem 1.4 and Proposition 2.2, consider the solution f^ε to equation (1.4) and its associated macroscopic quantities $(\mathcal{V}^\varepsilon, \mathcal{W}^\varepsilon)$ provided by Theorem 1.4. Furthermore, consider a strictly positive function f such that f, $\partial_t f$ and $\nabla^2_{\boldsymbol{u}} f$ lie in $\mathscr{C}^0\left(\mathbb{R}^+ \times K \times \mathbb{R}^2\right)$. Suppose that at initial time, it holds*

$$f_0^\varepsilon(\boldsymbol{x},\boldsymbol{u}) \leq f(0,\boldsymbol{x},\boldsymbol{u}), \quad \forall (\boldsymbol{x},\boldsymbol{u}) \in K \times \mathbb{R}^2,$$

and that f is super-solution to equation (3.7), that is

$$0 \leq \partial_t f + \mathrm{div}_{\boldsymbol{u}}\left[\mathbf{b}^\varepsilon f\right] - \partial_v^2 f - \frac{1}{\varepsilon}\rho_0^\varepsilon \partial_v\left[(v - \mathcal{V}^\varepsilon) f\right],$$

for all $(t,\boldsymbol{x},\boldsymbol{u}) \in \mathbb{R}^+ \times K \times \mathbb{R}^2$. Then it holds

$$f^\varepsilon(t,\boldsymbol{x},\boldsymbol{u}) \leq f(t,\boldsymbol{x},\boldsymbol{u}), \quad \forall (t,\boldsymbol{x}) \in \mathbb{R}^+ \times K, \text{a.e. in } \boldsymbol{u} \in \mathbb{R}^2.$$

Furthermore, the latter statement is also true if we replace "super-solution" by "sub-solution" and the symbol "$\geq$" by "$\leq$".

Proof. Since the proof of the comparison principle for super and sub-solutions is the same, we only detail it for super-solutions. According to Proposition 2.2, f^ε is a classical solution to (1.4), hence it holds

$$\partial_t(f^\varepsilon - f) + \mathrm{div}_{\boldsymbol{u}}\left[\mathbf{b}^\varepsilon(f^\varepsilon - f)\right] - \partial_v^2(f^\varepsilon - f) - \frac{1}{\varepsilon}\rho_0^\varepsilon\partial_v\left[(v - \mathcal{V}^\varepsilon)(f^\varepsilon - f)\right] \leq 0$$

a.e. with respect to $(t, \boldsymbol{x}, \boldsymbol{u}) \in \mathbb{R}^+ \times K \times \mathbb{R}^2$. Therefore, multiplying the latter relation by $\mathbb{1}_{f^\varepsilon \geq f}$, we deduce that in the weak sense, it holds

$$\partial_t(f^\varepsilon - f)_+ + \mathrm{div}_{\boldsymbol{u}}\left[\mathbf{b}^\varepsilon(f^\varepsilon - f)_+\right] - \partial_v^2(f^\varepsilon - f)_+ - \frac{1}{\varepsilon}\rho_0^\varepsilon\partial_v\left[(v - \mathcal{V}^\varepsilon)(f^\varepsilon - f)_+\right] \leq 0,$$

where $(\cdot)_+$ stands for the positive part and is defined by $(\cdot)_+ = (|\cdot| + id_{\mathbb{R}})/2$. In order to derive the latter relation, we follow a classical procedure which we do not detail here and which consists in regularizing the positive part $(\cdot)_+$. Then we integrate the latter relation with respect to time and $\boldsymbol{u}$ and obtain

$$\int_{\mathbb{R}^2}(f^\varepsilon - f)_+(t, \boldsymbol{x}, \boldsymbol{u})\mathrm{d}\boldsymbol{u} \leq \int_{\mathbb{R}^2}(f^\varepsilon - f)_+(0, \boldsymbol{x}, \boldsymbol{u})\mathrm{d}\boldsymbol{u}.$$

The latter computations are justified since $0 \leq (f^\varepsilon - f)_+ \leq f^\varepsilon$ and since, according to Theorem 1.4, f^ε has moments up to any order with respect to $\boldsymbol{u}$. Since $f^\varepsilon \leq f$ at time $t = 0$, we deduce the result. □

Appendix B. Regularity estimates

In this section, we derive regularity estimates for the solution f^ε to equation (1.4) and therefore prove Proposition 2.2. The main difficulty here consists in dealing with the contribution due to the non-linear drift N. We bypass this difficulty by estimating the norm of f^ε and its derivatives in the following weighted L^1 spaces

$$L^1(\omega_q) = \left\{f : \mathbb{R}^2 \mapsto \mathbb{R}, \int_{\mathbb{R}^2}|f|\,\omega_q(\boldsymbol{u})\mathrm{d}\boldsymbol{u} < +\infty\right\},$$

where $\omega_q(\boldsymbol{u}) = 1 + |\boldsymbol{u}|^q$, for $q \geq 2$. The first step consists in estimating the norm of f^ε in $L^1(\omega_q)$. This step relies on previous results obtained in [7]. Then we adapt these computations to evaluate the norm of the derivatives of f^ε in $L^1(\omega_q)$. Indeed, the derivatives of f^ε solve equation (1.4) with additional source terms whose contribution can be controlled thanks to the confining properties of the non-linear drift N. Let us outline the strategy in the case of the first order derivatives: equation (1.4) on f^ε reads as follows

$$\partial_t f^\varepsilon = \mathscr{A}^\varepsilon f^\varepsilon,$$

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

$$\mathscr{A}^\varepsilon f = \partial_v^2 f + \frac{1}{\varepsilon}\rho_0^\varepsilon\partial_v\left[(v - \mathcal{V}^\varepsilon)f\right] - \mathrm{div}_{\boldsymbol{u}}\left[\mathbf{b}^\varepsilon f\right].$$

Relying the arguments developed in the proof of [7, Proposition 3.1] it holds

Lemma B.1. *Consider some fixed ε and some $q \geq 2$. Under the assumptions of Theorem 1.4, consider the operator $\mathscr{A}^\varepsilon$ associated to the solution f^ε to equation (1.4). There exists a positive constant C such that for all $\boldsymbol{x} \in K$, it holds*

$$\int_{\mathbb{R}^2}\mathrm{sgn}\,(f)\mathscr{A}^\varepsilon(f)\,\omega_q(\boldsymbol{u})\mathrm{d}\boldsymbol{u} \leq C\|f\|_{L^1(\omega_q)} + q\int_{\mathbb{R}^2}\mathbb{1}_{|v|\geq 1}\frac{N(v)}{v}|v|^q f(\boldsymbol{u})\mathrm{d}\boldsymbol{u},$$

for all function f lying in $W^{2,1}(\omega_{p+q})$.

As a direct consequence of Lemma B.1, we deduce that for any smooth initial data f_0^ε, the solution f^ε to equation (1.4) provided by Theorem 1.4 lies in $L^\infty_{loc}\left(\mathbb{R}^+ \times K, L^1\left(\omega_q\right)\right)$ for all exponent $q \geq 2$. Indeed, since N meets the confining assumption (1.13) we have: $\mathbb{1}_{|v|\geq 1}N(v)/v \leq C$ for some

constant C. Therefore, multiplying equation (1.4) by $\operatorname{sgn}(f^\varepsilon)\omega_q(\boldsymbol{u})$, integrating with respect to $\boldsymbol{u}$ and applying Lemma B.1, we obtain that for all exponent q, it holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\|f^\varepsilon\|_{L^1(\omega_q)} \leq C\|f^\varepsilon\|_{L^1(\omega_q)}.$$

Hence, applying Gronwall's lemma and taking the supremum over all $\boldsymbol{x} \in K$, we deduce

$$\|f^\varepsilon(t)\|_{L^\infty(K,L^1(\omega_q))} \leq e^{Ct}\|f_0^\varepsilon\|_{L^\infty(K,L^1(\omega_q))}, \tag{B.1}$$

for all $t \in \mathbb{R}^+$. We follow the same strategy for derivatives of f^ε: writing $(g,h) = (\partial_v f^\varepsilon, \partial_w f^\varepsilon)$ and differentiating equation (1.4) with respect to v and w we obtain

$$\begin{cases} \partial_t g = \mathscr{A}^\varepsilon g - \left(N' - \Psi * \rho_0^\varepsilon - \dfrac{1}{\varepsilon}\rho_0^\varepsilon\right) g - N'' f^\varepsilon - ah, \\ \partial_t h = \mathscr{A}^\varepsilon h + g + bh. \end{cases} \tag{B.2}$$

Therefore, g and h solve the same equation as f^ε with additional an high order term due to the non-linear drift N. We control this additional term thanks to the confining properties of N.

Proof of Proposition 2.2. Let us consider some initial data f_0^ε lying in $\mathscr{C}^0\left(K, \mathscr{C}_c^\infty\left(\mathbb{R}^2\right)\right)$ and the associated solution f^ε to (1.4) provided by Theorem 1.4. We start by proving that f^ε lies in $L^\infty_{loc}\left(\mathbb{R}^+ \times K, W^{1,1}\left(\mathbb{R}^2\right)\right)$. We fix some $\boldsymbol{x} \in K$, some exponent $q \geq 2$ and integrate with respect to $\boldsymbol{u}$ the sum between the first equation in (B.2) multiplied by $\operatorname{sgn}(g)\omega_q(\boldsymbol{u})$ and the second multiplied by $\operatorname{sgn}(h)\omega_q(\boldsymbol{u})$. According to Lemma B.1, assumptions (1.16) on Ψ, (1.18) on ρ_0^ε and (1.14) on N, we obtain

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left(\|g\|_{L^1(\omega_q)} + \|h\|_{L^1(\omega_q)}\right) \leq & C\left(\|g\|_{L^1(\omega_q)} + \|h\|_{L^1(\omega_q)} + \|f^\varepsilon\|_{L^1\left(\omega_{p'}\right)}\right) \\ & + \int_{\mathbb{R}^2}\left(q\mathbb{1}_{|v|\geq 1}\frac{N(v)}{v} - N'(v)\right)|v|^q g(\boldsymbol{u})\mathrm{d}\boldsymbol{u}, \end{aligned}$$

where p' is given in assumption (1.14). Since N meets assumptions (1.13) and (2.1), we deduce that for q great enough, it holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\|g\|_{L^1(\omega_q)} + \|h\|_{L^1(\omega_q)}\right) \leq C\left(\|g\|_{L^1(\omega_q)} + \|h\|_{L^1(\omega_q)} + \|f^\varepsilon\|_{L^1\left(\omega_{p'}\right)}\right).$$

Therefore, applying Gronwall's lemma, taking the supremum over all $\boldsymbol{x} \in K$ and replacing $\|f^\varepsilon\|_{L^1\left(\omega_{p'}\right)}$ according to estimate (B.1), we deduce

$$\begin{aligned} & \|g(t)\|_{L^\infty(K,L^1(\omega_q))} + \|h(t)\|_{L^\infty(K,L^1(\omega_q))} \leq \\ & \qquad e^{Ct}\left(\|g_0\|_{L^\infty(K,L^1(\omega_q))} + \|h_0\|_{L^\infty(K,L^1(\omega_q))} + \|f_0^\varepsilon\|_{L^\infty(K,L^1(\omega_q))}\right), \end{aligned}$$

for all time $t \in \mathbb{R}^+$. As a straightforward consequence, we obtain the expected result: $f^\varepsilon \in L^\infty_{loc}\left(\mathbb{R}^+ \times K, W^{1,1}\left(\mathbb{R}^2\right)\right)$.

We obtain that $f^\varepsilon \in L^\infty_{loc}\left(\mathbb{R}^+ \times K, W^{2,1}\left(\mathbb{R}^2\right)\right)$ iterating the same procedure as before but this time on the derivatives of g and h and using assumption (2.1), which ensures that N''' has polynomial growth.

To end with, we obtain $\partial_t f^\varepsilon \in L^\infty_{loc}\left(\mathbb{R}^+ \times K, L^1\left(\mathbb{R}^2\right)\right)$ noticing that according to the definition of $\mathscr{A}^\varepsilon$, it holds

$$\|\mathscr{A}^\varepsilon f^\varepsilon\|_{L^\infty(K,L^1(\mathbb{R}^2))} \leq C\|f^\varepsilon\|_{L^\infty(K,W^{2,1}(\omega_q))},$$

for q great enough. Then we apply the previous estimates to the relation

$$\partial_t f^\varepsilon = \mathscr{A}^\varepsilon f^\varepsilon.$$

□

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. *J. Math. Neurosci.*, 10, 2012.

[2] G. Barles. A weak Bernstein method for fully nonlinear elliptic equations. *Differential Integral Equations*, 4(2):241–262, 1991.

[3] G. Barles, S. Mirrahimi, and B. Perthame. Concentration in Lotka-Volterra parabolic or integral equations: a general convergence result. *Methods Appl. Anal.*, 16(3):321–340, 2009.

[4] G. Barles and B. Perthame. Discontinuous solutions of deterministic optimal stopping time problems. *RAIRO Modél. Math. Anal. Numér.*, 21(4):557–579, 1987.

[5] G. Barles and B. Perthame. Dirac concentrations in Lotka-Volterra parabolic PDEs. *Indiana Univ. Math. J.*, 57:3275–3302, 2008.

[6] A. Blaustein. Large coupling in a FitzHug-Nagumo neural network: quantitative and strong convergence results. *arXiv:2210.10396v2 [math.AP]*.

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

[8] F. Bolley, J. A. Cañizo, and J. A. Carrillo. Stochastic mean-field limit: non-Lipschitz forces and swarming. *Math. Models Methods Appl. Sci.*, 21(11):2179–2210, 2011.

[9] 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'. *J. Math. Neurosci.*, 5, 2015.

[10] E. Bouin and S. Mirrahimi. A Hamilton-Jacobi approach for a model of population structured by space and trait. *Commun. Math. Sci.*, 13(6):1431–1452, 2015.

[11] P. C. Bressloff. Spatially periodic modulation of cortical patterns by long-range horizontal connections. *Physica D: Nonlinear Phenomena*, 185(3):131–157, 2003.

[12] P. C. Bressloff. Waves in Neural Media: From single Neurons to Neural Fields. *Springer*, 2014.

[13] M. J. Càceres, J. A. Carrillo, and B. Perthame. Analysis of nonlinear noisy integrate & fire neuron models: blow-up and steady states. *J. Math. Neurosci.*, 1(7):33, 2011.

[14] M. J. Càceres, P. Roux, D. Salort, and R. Schneider. Global-in-time solutions and qualitative properties for the NNLIF neuron model with synaptic delay. *Comm. Partial Differential Equations*, 44(12):1358–1386, 2019.

[15] G. A. Carpenter. A geometric approach to singular perturbation problems with applications to nerve impulse equations. *Journal of Differential Equations*, 1977.

[16] J. Carrillo, X. Dou, and Z. Zhou. A simplified voltage-conductance kinetic model for interacting neurons and its asymptotic limit. *arXiv:2203.02746 [math.AP]*, 2022.

[17] 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(9):3365–3388, 2015.

[18] N. Champagnat and P.-E. Jabin. The evolutionary limit for models of populations interacting competitively via several resources. *Journal of Differential Equations*, 251(1):176–195, 2011.

[19] C.-N. Chen, S.-I. Ei, Y.-P. Lin, and S.-Y. Kung. Standing waves joining with Turing patterns in FitzHugh-Nagumo type systems. *Comm. Partial Differential Equations*, 36(6):998–1015, 2011.

[20] J. Chevallier. Mean-field limit of generalized Hawkes processes. *Stochastic Process. Appl.*, 127(12):3870–3912, 2017.

[21] J. Chevallier, M. J. Càceres, M. Doumic, and P. Reynaud-Bouret. Microscopic approach of a time elapsed neural model. *Math. Models Methods Appl. Sci.*, 25(14):2669—2719, 2015.

[22] J. Chevallier, A. Duarte, E. Locherbach, and G. Ost. Mean-field limits for nonlinear spatially extended Hawkes processes with exponential memory kernels. *Stochastic Process. Appl.*, 25(129):1—-27, 2019.

[23] D. Chowdhury and M. C. Cross. Synchronization of oscillators with long-range power law interactions. *Phys. Rev. E*, 82:016205, Jul 2010.

[24] J. Crevat. Mean-field limit of a spatially-extended Fitzhugh-Nagumo neural network. *Kinet. Relat. Models*, 12(6):1329–1358, 2019.

[25] J. Crevat, G. Faye, and F. Filbet. Rigorous derivation of the nonlocal reaction-diffusion Fitzhugh-Nagumo system. *SIAM J. Math. Anal.*, 51(1):346–373, 2019.

[26] J. Crevat and F. Filbet. Asymptotic preserving schemes for the FitzHugh-Nagumo transport equation with strong local interactions. *BIT*, 61(3):771–804, 2021.

[27] P. Dayan and L. F. Abbott. *Theoretical neuroscience. Computational and mathematical modeling of neural systems*. 2001.

[28] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré. Particle systems with a singular mean-field self-excitation. application to neuronal networks. *Stochastic Processes and their Applications*, 125(6):2451–2492, 2015.

[29] O. Diekmann, P.-E. Jabin, S. Mischler, and B. Perthame. The dynamics of adaptation: An illuminating example and a Hamilton-Jacobi approach. *Theoretical Population Biology*, 67(4):257–271, 2005.

[30] R. FitzHugh. Impulses and Physiological States in Theoretical Models of Nerve Membrane. *Biophysical Journal*, 1(6):445–466, 1961.

[31] S. Gupta, M. Potters, and S. Ruffo. One-dimensional lattice of oscillators coupled through power-law interactions: Continuum limit and dynamics of spatial fourier modes. *Phys. Rev. E*, 85:066201, Jun 2012.
[32] S. P. Hastings. On the existence of homoclinic and periodic orbits for the FitzHugh-Nagumo equations. *The Quarterly Journal of Mathematics*, 27(1):123–134, 1976.
[33] A. Hodgkin and A. Huxley. Action Potentials Recorded from Inside a Nerve Fibre. *Nature*, 144:710–711, 1939.
[34] A. Hodgkin and A. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol.*, 117(4):500–544, 1952.
[35] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160, 1962.
[36] P.-E. Jabin, D. Poyato, and J. Soler. Mean-field limit of non-exchangeable systems. *arXiv:2112.15406*, 2021.
[37] C. K. R. T. Jones. Stability of the Travelling Wave Solution of the Fitzhugh-Nagumo System. *Transactions of the American Mathematical Society*, 286(2):431–469, 1984.
[38] E. Kandel, T. Jessell, J. Schwartz, S. Siegelbaum, and A. Hudspeth. *Principles of Neural Science, Fifth Edition*. McGraw-Hill's AccessMedicine. McGraw-Hill Education, 2013.
[39] H. Leman, S. Méléard, and S. Mirrahimi. Influence of a spatial structure on the long time behavior of a competitive Lotka-Volterra type system. *Discrete and Continuous Dynamical Systems - B*, 20(2):469–493, 2015.
[40] E. Luçon and C. Poquet. Emergence of oscillatory behaviors for excitable systems with noise and mean-field interaction: a slow-fast dynamics approach. *Comm. Math. Phys.*, 373(3):907–969, 2020.
[41] E. Luçon and W. Stannat. Mean field limit for disordered diffusions with singular interactions. *Ann. Appl. Probab.*, 24(5):1946–1993, 2014.
[42] J. S. Lund, A. Angelucci, and P. Bressloff. Anatomical substrates for functional columns in macaque monkey primary visual cortex. *Cereb Cortex*, 12:15–24, 2003.
[43] S. Méléard and S. Mirrahimi. Singular limits for reaction-diffusion equations with fractional Laplacian and local or nonlocal nonlinearity. *Comm. Partial Differential Equations*, 40(5):957–993, 2015.
[44] S. Mirrahimi and J.-M. Roquejoffre. Uniqueness in a class of Hamilton-Jacobi equations with constraints. *C. R. Math. Acad. Sci. Paris*, 353(6):489–494, 2015.
[45] S. Mischler, C. Quiñinao, and J. Touboul. On a kinetic Fitzhugh-Nagumo model of neuronal network. *Comm. Math. Phys.*, 342(3):1001–1042, 2016.
[46] S. Mischler and Q. Weng. Relaxation in Time Elapsed Neuron Network Models in the Weak Connectivity Regime. *Acta Applicandae Mathematicae*, 157:45–74, 2018.
[47] V. B. Mountcastle. Modality and topographic properties of single neurons of cat's somatic sensory cortex. *J Neurophysiol*, 20(4):408–434, 1957.
[48] V. B. Mountcastle. *11 Dynamic Neural Operations in Somatic Sensibility*. Harvard University Press, Cambridge, MA and London, England, 2005.
[49] J. Nagumo, S. Arimoto, and S. Yoshizawa. An Active Pulse Transmission Line Simulating Nerve Axon. *Proceedings of the IRE*, 50(10):2061–2070, 1962.
[50] K. Oelschläger. A law of large numbers for moderately interacting diffusion processes. *Z. Wahrscheinlichkeitstheorie verw Gebiete*, 69:279–322, 1985.
[51] I. Omelchenko, Y. Maistrenko, P. Hövel, and E. Schöll. Loss of coherence in dynamical networks: Spatial chaos and chimera states. *Phys. Rev. Lett.*, 106:234102, Jun 2011.
[52] I. Omelchenko, B. Riemenschneider, P. Hövel, Y. Maistrenko, and E. Schöll. Transition from spatial coherence to incoherence in coupled chaotic systems. *Phys. Rev. E*, 85:026212, Feb 2012.
[53] K. Pakdaman, B. Perthame, and D. Salort. Dynamics of a structured neuron population. *Nonlinearity*, 23(1):55–75, 2010.
[54] B. Perthame and D. Salort. On a voltage-conductance kinetic system for integrate and fire neural networks. *Kinet. Relat. Models*, 6(4):841–864, 2013.
[55] C. Quiñinao and J. Touboul. Limits and dynamics of randomly connected neuronal networks. *Acta Appl. Math.*, 136:167–192, 2015.
[56] C. Quiñinao and J. D. Touboul. Clamping and synchronization in the strongly coupled FitzHugh-Nagumo model. *SIAM J. Appl. Dyn. Syst.*, 19(2):788–827, 2020.