

# A structure preserving numerical scheme for Fokker-Planck equations of structured neural networks with learning rules

Qing He\*, Jingwei Hu† and Zhenman Zhou‡

June 28, 2022

## Abstract

In this work, we are concerned with a Fokker-Planck equation related to the nonlinear noisy leaky integrate-and-fire model for biological neural networks which are structured by the synaptic weights and equipped with the Hebbian learning rule. The equation contains a small parameter  $\varepsilon$  separating the time scales of learning and reacting behavior of the neural system, and an asymptotic limit model can be derived by letting  $\varepsilon \rightarrow 0$ , where the microscopic quasi-static states and the macroscopic evolution equation are coupled through the total firing rate. To handle the endowed flux-shift structure and the multi-scale dynamics in a unified framework, we propose a numerical scheme for this equation that is mass conservative, unconditionally positivity preserving, and asymptotic preserving. We provide extensive numerical tests to verify the schemes' properties and carry out a set of numerical experiments to investigate the model's learning ability, and explore the solution's behavior when the neural network is excitatory.

**Key words:** Fokker-Planck equation, Nonlinear Noisy Leaky Integrate-and-Fire model, neuron network, structure preserving scheme, asymptotic preserving, multiscale model.

**AMS subject classifications:** 35M13, 65M08, 92B20.

## 1 Introduction

In biological neural networks, neurons transmit information in two ways: one is conveying rapid and transient electrical signals on their membranes, and the other is sending neurotransmitters at a synapse. The learning and memory of the network is based on the connection weights of the synapses among the neurons [1]. How a neural network with a given structure reacts to the environment and how the neural network's infrastructure is modified by the environment in the long run are two of the essential topics in neural science.

Mathematically, there has been plenty of models describing the neural networks' behavior from all scales and sizes. The most successful single-neuron model is undoubtedly the Hodgkin-Huxley model [2] and its simplification, the Fitzhugh-Nagumo model [3]. However, when modeling large-scale biological neural networks, these models are far too complicated and impractical. Therefore, scientists and mathematicians usually adopt the integrate-and-fire model [4] when

---

\*Department of Mathematics, Southern Methodist University, Dallas, TX 75275, USA (andrewho@smu.edu).

†Department of Applied Mathematics, University of Washington, Seattle, WA 98195 (hujw@uw.edu).

‡Beijing International Center for Mathematical Research, Peking University, Beijing, China (zhenman@bicmr.pku.edu.cn).

modeling the system that consists of a large number of neurons [5, 6, 7, 8, 9, 10]. One of the most widely used version of the integrate-and-fire model is the nonlinear noisy leaky integrate and fire model (NNLIF). In this model, when the firing event does not occur, the membrane potential  $V(t)$  of a neuron is influenced by a relaxation to a resting potential  $V_I$ , an incoming synaptic current  $\mathcal{I}(t)$  from other neurons and the background. The current  $\mathcal{I}(t)$  is approximately decomposed into a drift term and a term of Brownian motion. The SDE, in the simplest form, is given by

$$dV = -(V - V_I) dt + \mu dt + \sigma dB_t, \quad (1.1)$$

where the parameters  $\mu$  and  $\sigma$  are determined by the synaptic current  $\mathcal{I}(t)$ . The distinguished feature of this model is the incorporation of the firing event: when a neuron reaches a firing potential  $V_F$ , it discharges itself and the potential jumps to a resetting potential  $V_R$  immediately [11, 12]. Here, we assume,  $V_R < V_F$ , and the firing event is expressed by

$$V(t^-) = V_F, \quad V(t^+) = V_R. \quad (1.2)$$

(1.1) and (1.2) constitute the stochastic process for the NNLIF model, which is a SDE coupled with a renewal process. By Itô's formula, one can derive the time evolution of the density function  $p(v, t)$  [13, 14], which represents the probability of finding a neuron at the voltage  $v$  and given time  $t$ :

$$\begin{cases} \frac{\partial p}{\partial t} + \frac{\partial}{\partial v}(hp) - a \frac{\partial^2 p}{\partial v^2} = N(t)\delta(v - V_R), & t \in (0, \infty), v \in (-\infty, V_F), \\ p(v, 0) = p_0(v), \quad p(-\infty, t) = p(V_F, t) = 0, \end{cases} \quad (1.3)$$

where

$$N(t) = -a \frac{\partial p}{\partial v}(V_F, t) \quad (1.4)$$

stands for the firing rate of the network,  $a = \sigma^2/2$  stands for the amplitude of the noise,  $\delta(\cdot)$  stands for the Dirac delta function, and

$$h = h(v, N(t)) = -v + I + wN(t) \quad (1.5)$$

stands for the rate of the rising of the neurons' potential, consisting of a decay term  $-v$ , an external input function  $I$  and a group reaction term  $wN(t)$ . The coefficient  $w$  in (1.5) can be either positive or negative. In the simplest scenario,  $w$  is a constant. We say that the neural network is excitatory when  $w > 0$ , and inhibitory when  $w < 0$ .

When it comes to the learning behavior of the biological neural network, the most seminal work is the Hebbian rule [15] stated as a motto "neurons that fire together wire together". A very simple mathematical setup of this rule is as follows: assuming that the strength of weights  $w_{ij}$  between two neurons  $i$  and  $j$  increases when the two neurons have high activity simultaneously. For  $M$  neurons in interactions, the classical Hebbian rule relates the weights to the activity  $N_i$  of the neuron  $i$  as

$$\frac{d}{dt}w_{ij} = k_{ij}N_iN_j - w_{ij}, \quad 1 \leq i, j \leq M.$$

In [16], the authors designed a model that integrates the NNLIF model and the Hebbian rule stated above. They considered a heterogeneous population of homogeneous neural networks structured by their synaptic weights  $w \in (-\infty, +\infty)$ , negative sign standing for inhibitory neurons and positive sign standing for excitatory neurons. They introduced a learning rule in order to modulate the distribution of synaptic weights  $H$  and allow the network to recognize some

given input signals  $I$  by choosing an appropriate heterogeneous synaptic weight distribution  $H$  adapted to the signal  $I$ . They assumed that the subnetworks interact only via the total firing rate  $N$ , with synaptic weights described with a single variable  $w$ . And they gave the following interpretation: all the subnetworks parametrized by  $w$  may modulate their intrinsic synaptic weight  $w$  with respect to a function  $\Phi$  which depends on the intrinsic activity  $N(w)$  of the network parametrized by  $w$  and of the total activity of the network  $\bar{N}$ . Then, the proposed generalization of the Hebbian rule consists in choosing

$$\Phi(N(w), \bar{N}) = \bar{N}N(w)K(w), \quad \bar{N} = \int_{-\infty}^{\infty} N(w)dw, \quad (1.6)$$

where  $K(\cdot)$  represents the learning strength of the subnetwork with synaptic weight  $w$ . Adding the above choice of the learning rule, they obtained the following equation

$$\frac{\partial p}{\partial t} + \frac{\partial}{\partial v} [(-v + I(w) + w\sigma(\bar{N}(t)))p] + \varepsilon \frac{\partial}{\partial w} [(\Phi - w)p] - a \frac{\partial^2 p}{\partial v^2} = N(w, t)\delta(v - V_R), \quad (1.7)$$

where  $\varepsilon$  stands for a time scale ratio which takes into account that learning is slower than the typical voltage activity of the network.

If we want to study the learning behavior of the model, we may further perform a time rescaling  $t \rightarrow t/\varepsilon$  for (1.7) to arrive at

$$\left\{ \begin{array}{l} \frac{\partial p}{\partial t} + \frac{\partial}{\partial w} [(\bar{N}(t)N(w, t)K(w) - w)p] \\ \quad = \frac{1}{\varepsilon} \left\{ a \frac{\partial^2 p}{\partial v^2} - \frac{\partial}{\partial v} [(-v + I(w) + w\sigma(\bar{N}(t)))p] + N(w, t)\delta(v - V_R) \right\} \\ \quad \text{for } (v, w, t) \in (-\infty, V_F) \times (-\infty, +\infty) \times (0, T), \\ N(w, t) = -a \frac{\partial p}{\partial v}(V_F - 0, w, t) \geq 0, \quad \bar{N}(t) = \int_{-\infty}^{\infty} N(w, t)dw, \end{array} \right. \quad (1.8)$$

where  $\frac{\partial p}{\partial v}(V_F - 0, w, t)$  means the left derivative in the  $v$  direction and the corresponding initial and boundary conditions are

$$p(V_F, w, t) = 0, \quad p(-\infty, w, t) = 0, \quad p(v, \pm\infty, t) = 0, \quad p(v, w, 0) = p^0(v, w). \quad (1.9)$$

Here  $p(v, w, t)$  at a fixed time  $t$  describes the probability density of finding a neuron with synaptic weight  $w$  and voltage  $v$ , henceforth the probability density of finding a single neuron with synaptic weight  $w$  is defined as

$$H(w, t) = \int_{-\infty}^{V_F} p(v, w, t)dv, \quad (1.10)$$

and the normalization condition of probability ensures that  $\int_{-\infty}^{+\infty} H(w, t)dw = 1$  holds for all  $t \geq 0$ .  $N(w, t)$  describes the firing rate of the neurons with synaptic weight  $w$ , at time  $t$  and  $\bar{N}(t)$  describes the entire network's firing rate at time  $t$  summing over all the synaptic weights.  $a$  is a positive noise coefficient. The model is a multi-scale transport equation, where the  $v$ -wise transport represents the integrate-and-fire dynamics of the neural network, and the  $w$ -wise transport represents the network's learning behavior.

Systematic understanding of the model (1.8) is largely lacking, although some of its properties were proved in [16]. Particularly, the numerical studies are far from sufficient for understanding the possible uses and limitations of this model. In this work, we further explore this model from a numerical perspective. Specifically, we aim to design a structure-preserving numerical scheme

for (1.8), and further explore this model with extensive numerical experiments to investigate possible solution structures and interpretations.

In our work, we propose a numerical scheme with the positivity preserving and asymptotic preserving properties. There are two key ingredients that make our scheme meet the desired properties. The first one is the Scharfetter-Gummel symmetric reconstruction [17] of the  $v$ -wise convection and diffusion terms in (1.8). The second one is that we treat the flux shift term  $N\delta(v - V_R)$  in (1.8) implicitly. With the symmetric reconstruction and this implicit treatment, we can prove that implementing our scheme means inverting a so-called  $M$ -matrix at each time step, and the positivity preserving properties of the scheme is independent of the respective ratio between  $\Delta t$  and  $\Delta v$  or  $\varepsilon$ . The only grid ratio requirement is that  $\Delta t/\Delta w$  has to satisfy the  $w$ -wise CFL condition, which is a natural requirement. Furthermore, when letting  $\varepsilon \rightarrow 0$  without adjusting  $\Delta t$ , we obtain a numerical scheme consistent with the asymptotic behavior of the model, consisting of a  $w$ -wise convection equation for  $H(w, t)$  defined in (1.10) and a  $v$ -wise quasi-steady equation for  $p(v, w, t)$  with given  $H(w, t)$ .

Numerical methods for Fokker-Planck type equations are abundant in the literature. In recent years many works have been proposed to preserve the structure of the solution such as mass conservation, positivity, entropy decay, and steady state preserving. This often relies on delicate choice of numerical fluxes and suitable time discretization. Without being exhaustive, we mention a few relevant works. Regarding spatial discretization, there are mainly two kinds of methods: one is based on formulating the equation into the transport form and using upwind fluxes [18]. This is later extended to general nonlinear aggregation and diffusion equations [19]. The other is based on formulating the equation into the diffusion form (i.e., Scharfetter-Gummel form). Depending on how to choose the fluxes, it bifurcates to several variants, for example, the Chang-Cooper scheme [20] and its generalization [21]; the symmetrized finite difference scheme [17], and more recently the high order finite difference scheme [22]. Regarding time discretization, high order explicit or semi-implicit schemes (c.f. [23]) can be used. However, we mention that to get positivity preserving implicit schemes for Fokker-Planck type equations, one is basically limited to first order scheme [24, 25, 26, 27]. The only exception we are aware of is the second order exponential method in [28] which is applicable to the linear Fokker-Planck equation and hard to generalize to more general cases. For the above reason, we limit ourself to first order method in this paper.

In the numerical experiment part, we verify the scheme's order of accuracy (first order for  $w, t$ , second order for  $v$ ) and asymptotic preserving properties, test the model's recognition properties presented in [16] and further explore the model's behavior when the solution  $p(v, w, t)$  is supported in the region  $w > 0$ . The numerical results suggest that this model shows great potential in distinguishing a general input by producing distinct and unsymmetrical responses to orthogonal basis functions. When  $w > 0$ , the solution  $p(v, w, t)$  may either develop to a steady state supported in a region  $w \in [0, A]$  for some positive  $A$ , or expand its support rapidly towards  $w \rightarrow \infty$ .

The rest of the paper is organized as follows. In Section 2, we present our schemes in detail, including the equations and their implementations; in Section 3, we give a detailed analysis of our schemes' properties, including conservation, positivity preserving properties and asymptotic preserving properties; and in Section 4, we present the numerical results including the verification of the schemes' accuracy, asymptotic preserving properties, numerical exploration of the model's learning and recognition ability and the behavior of the model with positive synaptic weights.

## 2 The Schemes

Our work aims to provide a numerical scheme for (1.8). In this section, we will give a detailed description of the schemes we designed.

This section is organized as follows: in Section 2.1 and Section 2.2, we introduce the grid point settings and the numerical scheme for (1.8), including a  $v$ -wise semi-implicit ( $v$ -SI) implementation and a  $v$ -wise fully-implicit ( $v$ -FI) implementation; and in Section 2.3, we introduce the matrix form of the  $v$ -SI and  $v$ -FI schemes, from which we introduce a useful matrix  $\mathbf{M}_j^{\bar{N}}$  to formulate further analysis.

### 2.1 Grid point settings and discretizations

Our scheme is a grid-based method, and this subsection gives the basic grid point settings of our scheme.

We assume that  $p(v, w, t)$  vanishes fast enough as  $v \rightarrow -\infty$  or  $w \rightarrow \pm\infty$  such that we truncate the domain into the bounded region

$$(v, w) \in [V_{\min}, V_F] \times [W_{\min}, W_{\max}] \times [0, T_{\max}],$$

and suppose that  $p(v, w, t)$  is negligible out of this region. Then we divide the intervals  $[V_{\min}, V_F]$ ,  $[W_{\min}, W_{\max}]$ ,  $[0, T_{\max}]$  into  $n_v, n_w, n_t$  equal sub-intervals with size

$$\Delta v = \frac{V_F - V_{\min}}{n_v}, \quad \Delta w = \frac{W_{\max} - W_{\min}}{n_w}, \quad \Delta t = \frac{T_{\max}}{n_t}, \quad (2.1)$$

respectively, so that the grid points are assigned as follows

$$\begin{cases} v_i = V_{\min} + i\Delta v & i = 0, 1, 2, \dots, n_v \\ w_j = W_{\min} + j\Delta w & j = 0, 1, 2, \dots, n_w \\ t^m = m\Delta t & m = 0, 1, 2, \dots, n_t, \end{cases} \quad (2.2)$$

then  $p_{i,j}^m$  represents the numerical approximation of  $p(v_i, w_j, t^m)$ . We always assume that there is an index  $r$  satisfying

$$V_R = v_r, \quad (2.3)$$

which naturally treats the derivative's discontinuity at  $v = V_R$ . Actually it can also be seen from (2.1) and (2.2) that  $v_{n_v} = V_F$ , so the two points related to flux shift are both aligned with the grid point.

For the discretization of  $N = -a\partial p/\partial v|_{v=V_F}$ , we approximate the derivative with the first-order finite difference

$$N_j^m = -a \frac{p_{n_v,j}^m - p_{n_v-1,j}^m}{\Delta v} = a \frac{p_{n_v-1,j}^m}{\Delta v} \quad (2.4)$$

(here we have applied the boundary condition  $p_{n_v,j}^m = 0$ ); for the discretization of  $\bar{N}(t) = \int_{-\infty}^{\infty} N(w, t)dw$ , we apply the simplest rectangular rule of numerical integration

$$\bar{N}^m = \Delta w \sum_{j=0}^{n_w} N_j^m, \quad (2.5)$$

which can also be explained as a trapezoidal rule of numerical integration since the boundary values  $N_0^m$  and  $N_{n_w}^m$  are supposed to be negligible (zero); furthermore, we also denote the discrete

representation of  $H(w, t) = \int_{-\infty}^{V_F} p(v, w, t) dv$  as

$$H_j^m = \Delta v \sum_{i=1}^{n_v} p_{i,j}^m. \quad (2.6)$$

## 2.2 The numerical scheme for the Fokker-Planck equation (1.8)

We design a scheme with a finite volume construction: for each  $i = 0, 1, \dots, n_v - 1$ ,  $j = 0, 1, \dots, n_w$ ,  $m = 0, 1, \dots, n_t - 1$ , we have the difference equation

$$\frac{p_{i,j}^{m+1} - p_{i,j}^m}{\Delta t} + \frac{\Phi_{i,j+1/2}^m - \Phi_{i,j-1/2}^m}{\Delta w} = \frac{1}{\varepsilon} \frac{F_{i+1/2,j}^m - F_{i-1/2,j}^m}{\Delta v}, \quad (2.7)$$

or, in an equivalent splitting form

$$\frac{p_{i,j}^{*m} - p_{i,j}^m}{\Delta t} + \frac{\Phi_{i,j+1/2}^m - \Phi_{i,j-1/2}^m}{\Delta w} = 0, \quad (2.8)$$

$$\frac{p_{i,j}^{m+1} - p_{i,j}^{*m}}{\Delta t} = \frac{1}{\varepsilon} \frac{F_{i+1/2,j}^m - F_{i-1/2,j}^m}{\Delta v}. \quad (2.9)$$

This splitting form is introduced only for convenience when discussing the scheme's properties in section 3.3. Our scheme is not based on the idea of operator splitting.

For the  $w$ -wise transport, we take the following explicit flux construction adapted from Godunov's Method (see, for example, (13.24) in [29]):

$$\Phi_{i,j+1/2}^m = \begin{cases} \begin{cases} \min \{ \Phi_{i,j}^m, \Phi_{i,j+1}^m \} & p_{i,j}^m \leq p_{i,j+1}^m \\ \max \{ \Phi_{i,j}^m, \Phi_{i,j+1}^m \} & p_{i,j}^m > p_{i,j+1}^m \end{cases} & j = 0, \dots, n_w - 1 \\ 0 & j = -1, n_w \end{cases} \quad (2.10)$$

where

$$\Phi_{i,j}^m = (\bar{N}^m N_j^m K(w_j) - w_j) p_{i,j}^m \quad \text{for } j = 0, \dots, n_w. \quad (2.11)$$

For the  $v$ -wise transport, we need to treat the operator partially implicitly in an appropriate way such that the scheme's stability is not limited by the grid ratio  $\Delta t / (\Delta v)^2$  and the stiffness introduced by the smallness of  $\varepsilon$ . We inherit the idea from [30] which implements a Scharfetter-Gummel symmetric reconstruction for the convection-diffusion operator and imposes a flux-shift from the boundary  $V_F$  back to  $V_R$ , but we treat the flux shift operator  $N\delta(V - V_R)$  implicitly:

$$F_{i+1/2,j}^m = \begin{cases} a \frac{M_{i+1/2,j}^{\bar{N}^\alpha}}{\Delta v} \left( \frac{p_{i+1,j}^{m+1}}{M_{i+1,j}^{\bar{N}^\alpha}} - \frac{p_{i,j}^{m+1}}{M_{i,j}^{\bar{N}^\alpha}} \right) + N_j^{m+1} \eta(v_{i+1/2} - V_R) & i = 0, 1, \dots, n_v - 2, \\ 0 & i = -1, n_v - 1, \end{cases} \quad (2.12)$$

where  $\eta(\cdot)$  is the Heaviside function ( $\eta(x) = 1$  when  $x \geq 0$  and  $\eta(x) = 0$  when  $x < 0$ ), and

$$M_{i,j}^{\bar{N}^\alpha} = e^{-\frac{[-v_i + I(w_j) + \sigma(\bar{N}^\alpha)]^2}{2\alpha}}, \quad M_{i+1/2,j}^{\bar{N}^\alpha} = \frac{2M_{i,j}^{\bar{N}^\alpha} M_{i+1,j}^{\bar{N}^\alpha}}{M_{i,j}^{\bar{N}^\alpha} + M_{i+1,j}^{\bar{N}^\alpha}}. \quad (2.13)$$

Now to complete the scheme, the only thing that remains is to specify the expression of  $\bar{N}^\alpha$ . As we shall see, the choice of  $\bar{N}^\alpha$  will highly affect the implementations and properties

of the scheme. Here we propose two choices of  $\overline{N}^\alpha$ : either  $\overline{N}^m$  or  $\overline{N}^{m+1}$ . We call the scheme (2.7)~(2.13) with

$$\overline{N}^\alpha = \overline{N}^m$$

the  $v$ -wise semi-implicit ( $v$ -SI) scheme, and the scheme (2.7)~(2.13) with

$$\overline{N}^\alpha = \overline{N}^{m+1}$$

the  $v$ -wise fully-implicit ( $v$ -FI) scheme. The advantage of the  $v$ -SI scheme is that its implementation is free of nonlinear solvers, and thus is the main focus of this work. The  $v$ -FI scheme is mainly introduced for comparison.

To avoid ambiguity, when discussing the  $v$ -SI scheme, we add a superscript "SI" for all the variables, which will be put at the first place and separated with other superscripts by a semicolon:

$$p_{i,j}^m, N_j^m, \overline{N}^m, H_j^m \quad \text{are written as} \quad p_{i,j}^{\text{SI};m}, N_j^{\text{SI};m}, \overline{N}^{\text{SI};m}, H_j^{\text{SI};m} \quad \text{for the } v\text{-SI scheme;}$$

and a similar set of notations will be introduced for the  $v$ -FI scheme:

$$p_{i,j}^m, N_j^m, \overline{N}^m, H_j^m \quad \text{are written as} \quad p_{i,j}^{\text{FI};m}, N_j^{\text{FI};m}, \overline{N}^{\text{FI};m}, H_j^{\text{FI};m} \quad \text{for the } v\text{-FI scheme.}$$

We still use  $p_{i,j}^m, N_j^m, \overline{N}^m, H_j^m$  when discussing the properties that the  $v$ -SI scheme and the  $v$ -FI scheme have in common (like the conservation and positivity preserving property).

It is worth noting that, in (2.12), the flux-shift part is  $\overline{N}^{m+1}$  for both the  $v$ -SI and  $v$ -FI construction of the scheme. This is an apparently small but important difference from the scheme proposed in [30]. This difference will be proved to make the scheme (unconditionally) positivity preserving for all  $\varepsilon$  and  $\Delta v$ .

### 2.3 The matrix problem involved in the schemes and the iterative methods for the $v$ -FI scheme

In this subsection, we introduce the matrix equation one will face when implementing the proposed schemes. To begin with, we define the class of matrices which will frequently appear throughout our work:

**Definition 1** For a given  $\overline{N} \in \mathbb{R}$  and  $j = 0, 1, \dots, n_w$ , we define the matrix  $\mathbf{M}_j^{\overline{N}}$  as an  $n_v \times n_v$  matrix such that

$$\left(\mathbf{M}_j^{\overline{N}}\right)_{kl} = \begin{cases} -M_{k-3/2,j}^{\overline{N}}/M_{k-2,j}^{\overline{N}} & l = k - 1 \\ -M_{k-1/2,j}^{\overline{N}}/M_{k,j}^{\overline{N}} & l = k + 1 \\ \left(M_{k-3/2,j}^{\overline{N}} + M_{k-1/2,j}^{\overline{N}}\right)/M_{k-1,j}^{\overline{N}} & l = k \text{ and } l \neq 1, n_v \\ M_{n-3/2,j}^{\overline{N}}/M_{n-1,j}^{\overline{N}} + 1 & l = k = n_v \\ M_{1/2,j}^{\overline{N}}/M_{0,j}^{\overline{N}} & l = k = 1 \\ -1 & k = r + 1 \text{ and } l = n_v \\ 0 & \text{otherwise} \end{cases} \quad (2.14)$$

where

$$M_{i,j}^{\overline{N}} = e^{-\frac{[-v_i + I(w_j) + \sigma(\overline{N})]^2}{2\alpha}} \quad \text{and} \quad M_{i+1/2,j}^{\overline{N}} = \frac{2M_{i,j}^{\overline{N}}M_{i+1,j}^{\overline{N}}}{M_{i,j}^{\overline{N}} + M_{i+1,j}^{\overline{N}}}. \quad (2.15)$$

$\mathbf{M}_j^{\bar{N}}$  is tridiagonal except for one entry at  $k = r + 1$ ,  $l = n_v$ , and its properties will be further discussed in section 3.2. For now, having defined  $\mathbf{M}_j^{\bar{N}}$ , we can write the matrix form of the proposed schemes.

For the  $v$ -SI scheme, assume that we have already solved  $p_{i,j}^{\text{SI};*m}$  from the explicit convection step (2.8), then one easily checks<sup>1</sup> that solving the  $v$ -wise transport step (2.9) means solving the system

$$\left( \varepsilon \mathbf{I} + \frac{a\Delta t}{(\Delta v)^2} \mathbf{M}_j^{\bar{N}^{\text{SI};m}} \right) \mathbf{p}_{:,j}^{\text{SI};m+1} = \varepsilon \mathbf{p}_{:,j}^{\text{SI};*m} \quad (2.16)$$

for each  $j = 0, 1, \dots, n_w$ , where

$$\begin{aligned} \mathbf{p}_{:,j}^{\text{SI};m+1} &= \left( p_{0,j}^{\text{SI};m+1}, p_{1,j}^{\text{SI};m+1}, \dots, p_{n_v-1,j}^{\text{SI};m+1} \right)^\top, \\ \mathbf{p}_{:,j}^{\text{SI};*m} &= \left( p_{0,j}^{\text{SI};*m}, p_{1,j}^{\text{SI};*m}, \dots, p_{n_v-1,j}^{\text{SI};*m} \right)^\top, \end{aligned}$$

$\mathbf{I}$  is the  $n \times n$  identity matrix and  $\mathbf{M}_j^{\bar{N}^{\text{SI};m}}$  follows the definition given in (2.14). (2.16) is a linear equation for  $\mathbf{p}_{:,j}^{\text{SI};m+1}$ .

For the  $v$ -FI scheme, solving the  $v$ -wise transport equation means solving the system

$$\left( \varepsilon \mathbf{I} + \frac{a\Delta t}{(\Delta v)^2} \mathbf{M}_j^{\bar{N}^{\text{FI};m+1}} \right) \mathbf{p}_{:,j}^{\text{FI};m+1} = \varepsilon \mathbf{p}_{:,j}^{\text{FI};*m} \quad (2.17)$$

for each  $j = 0, 1, \dots, n_w$ , where

$$\begin{aligned} \mathbf{p}_{:,j}^{\text{FI};m+1} &= \left( p_{0,j}^{\text{FI};m+1}, p_{1,j}^{\text{FI};m+1}, \dots, p_{n_v-1,j}^{\text{FI};m+1} \right)^\top, \\ \mathbf{p}_{:,j}^{\text{FI};*m} &= \left( p_{0,j}^{\text{FI};*m}, p_{1,j}^{\text{FI};*m}, \dots, p_{n_v-1,j}^{\text{FI};*m} \right)^\top. \end{aligned}$$

(2.17) is a nonlinear system since the matrix  $\mathbf{M}_j^{\bar{N}^{\text{FI};m+1}}$  depends directly on the unsolved variable  $\bar{N}^{\text{FI};m+1}$ . Therefore, we have to solve such a system with a nonlinear solver.

We propose a fix point iteration method as follows. We set the initial guess as  $p_{i,j}^{(0)} = p_{i,j}^{\text{FI};*m}$ , hence  $\bar{N}^{(0)} = \bar{N}^m$ , then for  $k = 0, 1, 2, \dots$ , let

$$\begin{cases} \left( \varepsilon \mathbf{I} + \frac{a\Delta t}{(\Delta v)^2} \mathbf{M}_j^{\bar{N}^{(k)}} \right) \mathbf{p}_{:,j}^{(k+1)} = \varepsilon \mathbf{p}_{:,j}^{\text{FI};*m} \text{ for } j = 0, 1, \dots, n_w \\ \bar{N}^{(k+1)} = \Delta w \sum_{j=0}^{n_v} N^{(k+1)} = \Delta w \sum_{j=0}^{n_v} \frac{p_{n_v-1,j}^{(k+1)}}{\Delta v}, \end{cases} \quad (2.18)$$

where

$$\mathbf{p}_{:,j}^{(k+1)} = \left( p_{0,j}^{(k+1)}, p_{1,j}^{(k+1)}, \dots, p_{n_v-1,j}^{(k+1)} \right)^\top.$$

We iterate  $K$  rounds to satisfy the stopping criterion  $|\bar{N}^{(K+1)} - \bar{N}^{(K)}| \leq \epsilon$  for some small enough  $\epsilon$ , and then take

$$\mathbf{p}_{:,j}^{\text{FI};m+1} = \mathbf{p}_{:,j}^{(K)}, \quad N_j^{\text{FI};m+1} = N_j^{(K)}, \quad \bar{N}^{\text{FI};m+1} = \bar{N}^{(K)}.$$

<sup>1</sup>The only notable detail here is that  $N_j^{m+1}$  in (2.12) should be replaced with  $ap_{n_v-1,j}^{m+1}/\Delta v$  according to (2.4).

### 3 The Properties of the Scheme

Having proposed the scheme in Section 2.2, we discuss how it preserves the properties of the original model.

We show that both the  $v$ -SI and the  $v$ -FI schemes proposed in Section 2.2 are mass-conservative and positivity-preserving as long as the  $w$ -wise CFL condition is satisfied (no matter how small  $\varepsilon$  and  $\Delta v$  is). As for the AP property, the  $v$ -FI scheme and the  $v$ -SI scheme preserve the asymptotic limit in slightly different forms.

#### 3.1 Conservative properties

Being total mass invariant is one of the most basic properties of our model (1.8), and our scheme satisfies the conservation property naturally thanks to the finite volume construction.

**Theorem 1** *The schemes (2.7)~(2.13) satisfies*

$$\sum_{i=0}^{n_v} \sum_{j=0}^{n_w} p_{i,j}^{m+1} = \sum_{i=0}^{n_v} \sum_{j=0}^{n_w} p_{i,j}^m, \quad (3.1)$$

Proof: Summing (2.7) over  $i$  from 0 to  $n_v - 1$ , over  $j$  from 0 to  $n_w$  and using (2.10) (2.12), we have

$$\begin{aligned} & \sum_{i=0}^{n_v-1} \sum_{j=0}^{n_w} p_{i,j}^{m+1} \\ &= \sum_{i=0}^{n_v-1} \sum_{j=0}^{n_w} p_{i,j}^m + \frac{\Delta t}{\varepsilon \Delta v} \sum_{i=0}^{n_v-1} \sum_{j=0}^{n_w} (F_{i+1/2,j}^m - F_{i-1/2,j}^m) - \frac{\Delta t}{\Delta w} \sum_{i=0}^{n_v-1} \sum_{j=0}^{n_w} (\Psi_{i,j+1/2}^m - \Psi_{i,j-1/2}^m) \\ &= \sum_{i=0}^{n_v-1} \sum_{j=0}^{n_w} p_{i,j}^m + \frac{\Delta t}{\varepsilon \Delta v} \sum_{j=0}^{n_w} (F_{n_v-1/2,j}^m - F_{-1/2,j}^m) - \frac{\Delta t}{\Delta w} \sum_{i=0}^{n_v-1} (\Psi_{i,n_w+1/2}^m - \Psi_{i,-1/2}^m) \\ &= \sum_{i=0}^{n_v-1} \sum_{j=0}^{n_w} p_{i,j}^m + \frac{\Delta t}{\varepsilon \Delta v} \sum_{j=0}^{n_w} (0 - 0) - \frac{\Delta t}{\Delta w} \sum_{i=0}^{n_v-1} (0 - 0) = \sum_{i=0}^{n_v-1} \sum_{j=0}^{n_w} p_{i,j}^m, \end{aligned} \quad (3.2)$$

and because of the boundary condition  $p_{n_v,j}^m = 0$ , we can change “ $\sum_{i=0}^{n_v-1}$ ” on both ends of (3.2) into “ $\sum_{i=0}^{n_v}$ ”, which gives us (3.1).

#### 3.2 Properties of the matrix $M_j^{\bar{N}}$

Before discussing the properties of our schemes, we firstly investigate the matrix  $M_j^{\bar{N}}$  defined in (2.14) and (2.15). This matrix has a one dimensional kernel spanned by a strictly positive vector. This property will help us to show the positivity of the schemes.

The main result in this section is

**Lemma 1** *For all  $\bar{N} \geq 0$ ,  $M_j^{\bar{N}}$  defined in (2.14) and (2.15) has a unique (up to a positive multiple) right eigenvector  $\mathbf{q} = (q_0, \dots, q_{n_v-1})^\top$  for eigenvalue 0 which is strictly positive.*

Proof: We want to study the solution of

$$M_j^{\bar{N}} \mathbf{q} = \mathbf{0}. \quad (3.3)$$

We left-multiply both sides of (3.3) by the  $n \times n$  bi-diagonal matrix

$$\mathbf{L} = \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & 1 \end{bmatrix}, \text{ or entry-by-entry : } L_{kl} = \begin{cases} 1 & k = l + 1 \text{ or } k = l \\ 0 & \text{otherwise} \end{cases}$$

obtaining  $(\mathbf{L}\mathbf{M}_j^{\overline{N}})\mathbf{q} = \mathbf{0}$  (note that  $\mathbf{L}$  is non-singular, so  $\mathbf{L}\mathbf{M}_j^{\overline{N}}$  has the same kernel with  $\mathbf{M}_j^{\overline{N}}$ ), which can be rearranged into the following system:

$$\begin{cases} q_i = \frac{M_{i,j}^{\overline{N}}}{M_{i,j+1}^{\overline{N}}}q_{i+1} & i = r-1, r-2, \dots, 0 \\ q_i = \frac{M_{i,j}^{\overline{N}}}{M_{i,j+1}^{\overline{N}}}q_{i+1} + \frac{M_{i,j}^{\overline{N}}}{M_{i,j+1/2}^{\overline{N}}}q_{n-1} & i = n-3, n-4, \dots, r \\ q_i = \left( \frac{M_{i,n-2}^{\overline{N}}}{M_{i,n-1}^{\overline{N}}} + \frac{M_{i,n-2}^{\overline{N}}}{M_{i,n-3/2}^{\overline{N}}} \right) q_{n-1} & i = n-2 \end{cases} \quad (3.4)$$

One can derive from (3.4) recurrently (in a reversed order, from  $i = n-2$  to  $i = 0$ ) that  $q_{n-2}, q_{n-3}, \dots, q_0$  can all be written as a positive multiple of  $q_{n-1}$  whose expressions only contains entries of  $\mathbf{M}_j^{\overline{N}}$ . This implies strict-positiveness and uniqueness (up to a positive number multiplication) of  $\mathbf{q}$ .  $\square$

### 3.3 Positivity preserving properties

Solutions being non-negative is another basic property of model (1.8), hence we also expect our scheme to be positivity preserving. In Theorem 2.2 [30], it is proved that the scheme's positivity preserving property relies on the grid ratio  $\Delta t/(\Delta v)^2$ . In this section, we see that the positivity preserving property of our scheme does not rely on  $\Delta t/(\Delta v)^2$  or  $\varepsilon$ , and this improvement comes from our implicit treatment to the flux-shift operator that is stated in Section 2.

We start with the general matrix form of our problem (both for the  $v$ -SI scheme and for the  $v$ -FI scheme):

$$\left( \varepsilon \mathbf{I} + \frac{a\Delta t}{(\Delta v)^2} \mathbf{M}_j^{\overline{N}\alpha} \right) \mathbf{p}_{:,j}^{m+1} = \varepsilon \mathbf{p}_{:,j}^{*m}, \quad (3.5)$$

where

$$\begin{aligned} \mathbf{p}_{:,j}^{m+1} &= (p_{0,j}^{m+1}, p_{1,j}^{m+1}, \dots, p_{n_v-1,j}^{m+1})^\top, \\ \mathbf{p}_{:,j}^{*m} &= (p_{0,j}^{*m}, p_{1,j}^{*m}, \dots, p_{n_v-1,j}^{*m})^\top. \end{aligned}$$

A crucial property of system (3.5) is that the to-be-inverted matrix  $\varepsilon \mathbf{I} + \frac{a\Delta t}{(\Delta v)^2} \mathbf{M}_j^{\overline{N}\alpha}$  is the so-called M-matrix, which is an important topic in the area of positive operators. In [31], the authors provided forty equivalent statements that ensures a Z-matrix (a matrix with all diagonal entries non-negative and all non-diagonal entries non-positive) to be an M-matrix. We pick out the statement F16 and K33 in this article, forming the following lemma:

**Lemma 2 (Equivalence of statements F16 and K33 of Theorem 1 in [31])** *A semi-positive Z-matrix is an M-matrix, and henceforth monotone. Specifically, if a  $n \times n$  matrix  $\mathbf{M}$  satisfies*

$$\begin{cases} \mathbf{M}_{i,j} \geq 0 & j = i \\ \mathbf{M}_{i,j} \leq 0 & j \neq i \end{cases}$$

*and there exists a strictly positive (every entry is positive)  $n \times 1$  vector  $\mathbf{v}^* > 0$  such that  $\mathbf{M}\mathbf{v}^* > 0$ , then  $\mathbf{M}$  is a non-singular M-matrix, and henceforth  $\mathbf{M}\mathbf{v} \geq 0$  implies  $\mathbf{v} \geq 0$ . Here*

- $\mathbf{v} > 0$  means that all the entries of  $\mathbf{v}$  are positive, and
- $\mathbf{v} \geq 0$  means that all the entries of  $\mathbf{v}$  are non-negative.

With Lemma 2, we can further prove the following Lemma:

**Lemma 3** *For (3.5),  $\mathbf{p}_{:,j}^{*m} \geq 0$  indicates  $\mathbf{p}_{:,j}^{m+1} \geq 0$ .*

*Proof:* Firstly, it is not hard to check that  $\varepsilon \mathbf{I} + \frac{a\Delta t}{(\Delta v)^2} \mathbf{M}_j^{\overline{N}^\alpha}$  is a Z-matrix according to the definitions. Secondly, we notice that

$$\left( \varepsilon \mathbf{I} + \frac{a\Delta t}{(\Delta v)^2} \mathbf{M}_j^{\overline{N}^\alpha} \right) \mathbf{Q}_{:,j}^{\overline{N}^\alpha} = \varepsilon \mathbf{Q}_{:,j}^{\overline{N}^\alpha} > 0$$

since  $\mathbf{Q}_{i,j}^{\overline{N}^\alpha}$  is in the kernel of  $\mathbf{M}_j^{\overline{N}^\alpha}$ , according to Lemma 2 (taking  $\mathbf{v}^*$  as  $\mathbf{Q}_j^{\overline{N}^\alpha}$ ), the matrix  $\varepsilon \mathbf{I} + \frac{a\Delta t}{(\Delta v)^2} \mathbf{M}_j^{\overline{N}^\alpha}$  is a non-singular M-Matrix and henceforth  $\mathbf{p}_j^{*m} \geq 0$  indicates  $\mathbf{p}_j^{m+1} \geq 0$ .  $\square$

Applying Lemma 3 and (2.8), we can give out the following sufficient conditions of the positivity preserving property of the proposed schemes

**Theorem 2 (Positivity Preserving Properties)**  *$p_{i,j}^{m+1} \geq 0$ , holds for all  $i = 0, \dots, n_v$  and  $j = 0, \dots, n_w$  as long as*

$$p_{i,j}^m - \frac{\Delta t}{\Delta w} \left( \Phi_{i,j+1/2}^m - \Phi_{i,j-1/2}^m \right) \geq 0 \quad (3.6)$$

*holds for all  $i = 0, \dots, n_v, j = 0, \dots, n_w$ .*

(3.6) tells us that as long as the grid ratio  $\Delta t/\Delta w$  is not too big, (which is the most basic requirement for almost all numerical schemes for hyperbolic conservation law, and is also known as the CFL condition), the numerical solution will be non-negative. Such a property is unrelated to the grid ratio  $\Delta t/\Delta v^2$  or the scaling parameter  $\varepsilon$ . This means that we can take arbitrarily small  $\Delta v$  and  $\varepsilon$  without worrying about violations of the positivity of the solutions.

Especially, as we can take  $\varepsilon$  arbitrarily small for a fixed  $\Delta t$ , it becomes possible for us to further analyze the asymptotic behavior of the scheme when  $\varepsilon \rightarrow 0$ .

### 3.4 Asymptotic preserving properties

We first discuss the asymptotic properties of the original continuous model. (1.8)

Integrating the first equation of (1.8) over  $v$  from  $-\infty$  to  $V_F$  and applying the boundary condition, we get

$$\frac{\partial H}{\partial t} + \frac{\partial}{\partial w} [(\overline{N}(t)N(w,t)K(w) - w)H] = 0, \quad (3.7)$$

where  $H = H(w,t)$  is defined in (1.10). When taking  $\varepsilon \rightarrow 0$  in (1.8), we have

$$\begin{cases} a \frac{\partial^2 p}{\partial v^2} - \frac{\partial}{\partial v} [(-v + I(w) + w\sigma(\overline{N}(t)))p] + N(w,t)\delta(v - V_R) = 0, \\ N(w,t) := -a \frac{\partial p}{\partial v}(V_F - 0, w, t), \quad \overline{N}(t) = \int_{-\infty}^{\infty} N(w,t)dw. \end{cases} \quad (3.8)$$

According to Theorem 3.1 in [16], the solution of equation (3.8) is guaranteed to exist and be unique when  $H(w, t)$  is given. Equations (1.10), (3.7) and (3.8) gives the asymptotic limit equation of (1.8) as  $\varepsilon \rightarrow 0$ , where (3.7) can be viewed as the macroscopic equation governing the slow dynamics while (3.8) gives the local microscopic equilibrium of the fast dynamics. We consider our scheme with only time  $t$  discretized

$$\left\{ \begin{aligned} & \frac{p^{n+1} - p^n}{\Delta t} + \frac{\partial}{\partial w} [(\bar{N}^n N^n(w)K(w) - w)p^n] \\ & = \frac{1}{\varepsilon} \left\{ a \frac{\partial^2 p^{n+1}}{\partial v^2} - \frac{\partial}{\partial v} [(-v + I(w) + w\sigma(\bar{N}^\alpha))p^{n+1}] + N^{n+1}(w)\delta(v - V_R) \right\} \\ & \quad \text{for } (v, w, t) \in (-\infty, V_F) \times (-\infty, +\infty) \times (0, T) \\ & N^{n+1}(w) := -a \frac{\partial p}{\partial v}(V_F - 0, w, t), \quad \bar{N}^{n+1} = \int_{-\infty}^{\infty} N^{n+1}(w)dw; \end{aligned} \right. \quad (3.9)$$

where  $\alpha$  is  $n$  and  $n + 1$  for the  $v$ -SI scheme and  $v$ -FI scheme, respectively.

For the  $v$ -FI scheme, integrating the first equation in (3.9) and applying the boundary conditions, we have

$$\frac{H^{n+1} - H^n}{\Delta t} + \frac{\partial}{\partial w} [(\bar{N}^n N^n(w)K(w) - w)H^n] = 0 \quad (3.10)$$

which is a discretization of (3.7); and when  $\varepsilon \rightarrow 0$ , (3.9) becomes

$$\left\{ \begin{aligned} & a \frac{\partial^2 p^{n+1}}{\partial v^2} - \frac{\partial}{\partial v} [(-v + I(w) + w\sigma(\bar{N}^{n+1}))p^{n+1}] + N^{n+1}(w)\delta(v - V_R) = 0 \\ & N^{n+1}(w) := -a \frac{\partial p}{\partial v}(V_F - 0, w, t), \quad \bar{N}^{n+1} = \int_{-\infty}^{\infty} N^{n+1}(w)dw, \end{aligned} \right. \quad (3.11)$$

which is directly the same as (3.8).

For the  $v$ -SI scheme, Integrating the first equation in (3.9) and applying the boundary conditions, we obtain (3.10) again, while taking  $\varepsilon \rightarrow 0$  in (3.9), we have

$$\left\{ \begin{aligned} & a \frac{\partial^2 p^{n+1}}{\partial v^2} - \frac{\partial}{\partial v} [(-v + I(w) + w\sigma(\bar{N}^n))p^{n+1}] + N^{n+1}(w)\delta(v - V_R) = 0 \\ & N^{n+1}(w) := -a \frac{\partial p}{\partial v}(V_F - 0, w, t), \quad \bar{N}^{n+1} = \int_{-\infty}^{\infty} N^{n+1}(w)dw. \end{aligned} \right. \quad (3.12)$$

We remark that the local equilibrium (3.12) can be viewed as a linearization of (3.11) such that the nonlinear coefficient  $\sigma(\bar{N}^{n+1})$  is replaced by  $\sigma(\bar{N}^n)$ ,

The schemes (3.10) and (3.12) can be viewed as the limiting scheme for the following time rescaled maybe “delayed” equation as  $\varepsilon \rightarrow 0$

$$\left\{ \begin{aligned} & \frac{\partial p}{\partial t} + \frac{\partial}{\partial w} [(\bar{N}(t)N(w, t)K(w) - w)p] \\ & = \frac{1}{\varepsilon} \left\{ a \frac{\partial^2 p}{\partial v^2} - \frac{\partial}{\partial v} [(-v + I(w) + w\sigma(\bar{N}(t - \Delta t)))p] + N(w, t)\delta(v - V_R) \right\} \\ & \quad \text{for } (v, w, t) \in (-\infty, V_F) \times (-\infty, +\infty) \times (0, T) \\ & N(w) := -a \frac{\partial p}{\partial v}(V_F - 0, w, t), \quad \bar{N} = \int_{-\infty}^{\infty} N(w)dw. \end{aligned} \right. \quad (3.13)$$

Therefore, the  $v$ -SI scheme is asymptotic preserving up to a  $\Delta t$  time delay. In other words, when  $\varepsilon \rightarrow 0$ , it is still a consistent first-order in time discretization to the limit of the original model.

## 4 Numerical Results

This section gives the numerical results obtained from our schemes, involving both the tests on the schemes' performance and the explorations of the model by our scheme. In Section 4.1, we numerically test the order of accuracy of our schemes; in Section 4.2, we numerically validate the asymptotic preserving properties of our schemes; in Section 4.3, we test the model's learning and discrimination abilities; in Section 4.4, we focus on exploring the model's behavior when the neural system is excitatory.

### 4.1 Order of Convergence

In this part, we test the order of accuracy of the  $v$ -SI schemes. Since the exact solution is unavailable, we estimate the order of the error by

$$O_{h,L^p} = \log_2 \frac{\|\omega_h - \omega_{\frac{h}{2}}\|_p}{\|\omega_{\frac{h}{2}} - \omega_{\frac{h}{4}}\|_p},$$

where  $\omega_h$  is the numerical solution with step length  $h$ . The term  $O_h$  above is an approximation for the accuracy order. Errors in both  $L^1$  and  $L^2$  norms are examined.

We choose  $V_F = 2, V_R = 1, V_{\min} = -4, a = 1, \varepsilon = 0.5, W_{\min} = -1.1, W_{\max} = 0.1, \sigma(\bar{N}) = \bar{N}, I(w) \equiv 0$  and

$$p_{\text{init}} = \begin{cases} \sin^2(\pi v) \sin^2(\pi w) & -1 < w < 0 \text{ and } -1 < v < 1 \\ 0 & \text{otherwise} \end{cases}$$

throughout this section, and we test the accuracy for both  $T_{\max} = 0.1$  and  $T_{\max} = 1$ . We test the numerical results on  $v, w, t$  directions by fixing two of  $\Delta t, \Delta w$  and  $\Delta v$  while adjusting the third one. The results are shown in Table 1, Table 2 and Table 3, respectively.

$\Delta v$	$\ p_{\Delta v} - p_{\Delta v/2}\ _1$	$O_{\Delta v, L^1}$	$\ p_{\Delta v} - p_{\Delta v/2}\ _2$	$O_{\Delta v, L^2}$
2.0e-1	1.1337e-03	2.0818	3.5479e-03	2.0675
1.0e-1	2.6781e-04	2.0122	8.4643e-04	2.0080
5.0e-2	6.6390e-05	1.9340	2.1044e-04	1.8739
2.5e-2	1.7374e-05	-	5.7417e-05	-

$\Delta v$	$\ p_{\Delta v} - p_{\Delta v/2}\ _1$	$O_{\Delta v, L^1}$	$\ p_{\Delta v} - p_{\Delta v/2}\ _2$	$O_{\Delta v, L^2}$
2.0e-1	1.4571e-01	1.8905	2.8929e-01	1.7984
1.0e-1	3.9299e-02	1.7441	8.3168e-02	1.7309
5.0e-2	1.1731e-02	1.7872	2.5056e-02	1.8471
2.5e-2	3.3990e-03	-	6.9641e-03	-

Table 1:  $v$ -wise order of accuracy,  $\Delta t = 10^{-3}$ ,  $\Delta w = 10^{-2}$  are fixed. Upper table:  $T_{\max} = 0.1$ ; Lower table:  $T_{\max} = 2.5$ .

It can be seen that when  $T_{\max} = 0.1$ , the scheme shows the expected second-order convergence rate in the  $v$  direction, and the first-order convergence in  $w, t$  directions. However, when  $T_{\max} = 2.5$ , the  $v$ - and  $w$ - wise orders of accuracy decrease a bit. This may be caused by the  $w$ -wise nonlinearity of the equation (1.8). Indeed, when  $T = 2.5$  the solution has already developed into a discontinuous pattern (Figure 1).

$\Delta w$	$\ p_{\Delta w} - p_{\Delta w/2}\ _1$	$O_{\Delta w, L^1}$	$\ p_{\Delta w} - p_{\Delta w/2}\ _2$	$O_{\Delta w, L^2}$
4.0e-2	4.2858e-03	0.9550	9.9587e-03	0.9543
2.0e-2	2.2108e-03	1.0038	5.1394e-03	1.0030
1.0e-2	1.1025e-03	0.9849	2.5643e-03	0.9801
0.5e-2	5.5703e-04	-	1.3000e-03	-

$\Delta w$	$\ p_{\Delta w} - p_{\Delta w/2}\ _1$	$O_{\Delta w, L^1}$	$\ p_{\Delta w} - p_{\Delta w/2}\ _2$	$O_{\Delta w, L^2}$
4.0e-2	4.5348e-01	1.0102	3.9149e-01	0.7353
2.0e-2	2.2514e-01	1.0095	2.3516e-01	0.7557
1.0e-2	1.1183e-01	0.9717	1.3928e-01	0.6145
0.5e-2	5.7022e-02	-	9.0971e-02	-

Table 2:  $w$ -wise order of accuracy,  $\Delta t = 10^{-3}$ ,  $\Delta v = 10^{-1}$  are fixed. Upper table:  $T_{\max} = 0.1$ ; Lower table:  $T_{\max} = 2.5$ .

$\Delta t$	$\ p_{\Delta t} - p_{\Delta t/2}\ _1$	$O_{\Delta t, L^1}$	$\ p_{\Delta t} - p_{\Delta t/2}\ _2$	$O_{\Delta t, L^2}$
2.0e-3	3.8523e-04	0.9730	1.2219e-03	0.9647
1.0e-3	1.9625e-04	0.9686	6.2608e-04	0.9626
5.0e-4	1.0028e-04	1.0093	3.2125e-04	1.0089
2.5e-4	4.9819e-05	-	1.5963e-04	-

$\Delta t$	$\ p_{\Delta t} - p_{\Delta t/2}\ _1$	$O_{\Delta t, L^1}$	$\ p_{\Delta t} - p_{\Delta t/2}\ _2$	$O_{\Delta t, L^2}$
2.0e-3	4.9612e-03	0.9473	4.5521e-03	0.9676
1.0e-3	2.5729e-03	0.9884	2.3278e-03	0.9954
5.0e-4	1.2969e-03	0.9781	1.1677e-03	0.9895
2.5e-4	6.5836e-04	-	5.8810e-04	-

Table 3:  $t$ -wise order of accuracy,  $\Delta v = 10^{-1}$ ,  $\Delta w = 10^{-2}$  are fixed. Upper table:  $T_{\max} = 0.1$ ; Lower table:  $T_{\max} = 2.5$ .

## 4.2 Asymptotic behaviors of the $v$ -SI and $v$ -FI schemes

In this subsection, we numerically validate the asymptotic preserving properties of both the  $v$ -SI scheme and the  $v$ -FI scheme by validating that with fixed grid size, the numerical solution of the scheme converge to the  $v$ -wise quasi-steady state at the discretized level as  $\varepsilon \rightarrow 0$ .

Firstly we introduce a discretized version of the time-independent quasi-steady state equations (1.10) and (3.8) for a given  $H = H(w)$ . This will serve as a reference solution. We denote the solution of the continuous time-independent quasi-steady state equations (1.10) and (3.8) for a given  $H = H(w)$  as  $P^H(v, w)$ . Since (3.8) means that  $P^H(v, w)$  should nullify the stiff operator in (1.8), our discretized version can also be directly defined to nullify the stiff terms (order  $1/\varepsilon$  terms) in the main equation (2.7) in our scheme. For a given discretized grid function  $H_j$  such that  $\Delta w \sum_{j=0}^{n_w} H_j = 1$  the  $v$ -wise quasi-steady state in the discretized level  $P_{i,j}^H$  and corresponding  $N_j^H, \bar{N}^H$  is defined to satisfy the following equations

$$\Delta v \sum_{i=0}^{n_v-1} P_{i,j}^H = H_j \quad \text{for } j = 0, 1, \dots, n_w, \quad (4.1)$$

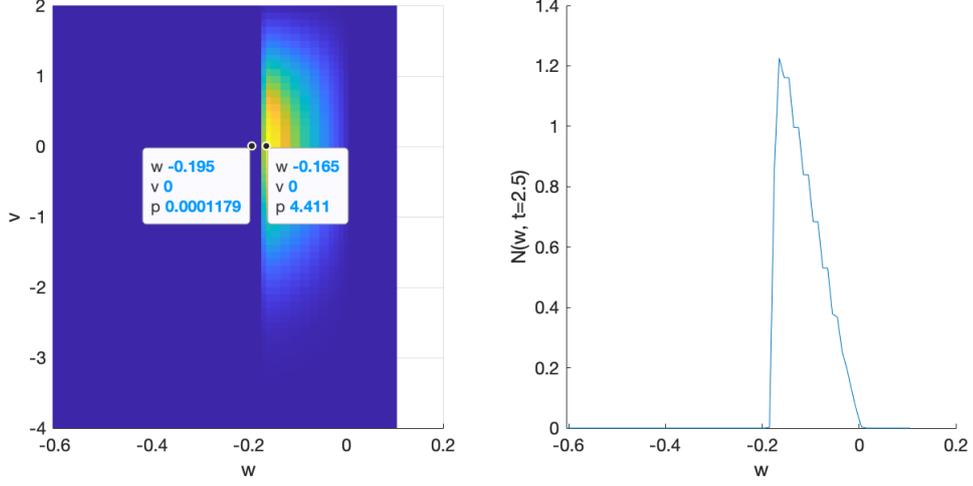


Figure 1: When  $t = 2.5$ , there is a discontinuity in the numerical solution. Left:  $p(v, w, t = 2.5)$ , with the marked data showing that small variance along the  $w$  direction causes a big variance of  $p$ ; Right:  $N(w, t = 2.5)$ .

$$F_{i+1/2,j}^H - F_{i-1/2,j}^H = 0 \quad (4.2)$$

$$F_{i+1/2,j}^H = \begin{cases} a \frac{M_{i+1/2,j}^{\bar{N}^H}}{\Delta v} \left( \frac{P_{i+1,j}^H}{M_{i+1,j}^{\bar{N}^H}} - \frac{P_{i,j}^H}{M_{i,j}^{\bar{N}^H}} \right) + N_j^H \eta(v_{i+1/2} - V_R) & i = 0, 1, \dots, n_v - 2 \\ 0 & i = -1, n_v - 1 \end{cases} \quad (4.3)$$

$$N_j^H = a P_{n_v-1,j}^H / \Delta t, \quad \bar{N}^H = \Delta w \sum_{j=0}^{n_w} N_j^H \quad (4.4)$$

To generate this discretized  $v$ -wise quasi-steady state for a given  $H_j$  with  $\Delta v \sum_{j=0}^{n_w} H_j = 1$ , one may solve the system

$$\begin{cases} M_j^{\bar{N}^{(k)}} \mathbf{P}_{:,j}^{(k+1)} = \mathbf{0} \text{ and } \Delta t \sum_{i=0}^{n_v-1} P_{i,j}^{(k+1)} = H_j \text{ for } j = 0, 1, \dots, n_w \\ \bar{N}^{(k+1)} = \Delta w \sum_{j=0}^{n_w} N^{(k+1)} = \Delta w \sum_{j=0}^{n_w} \frac{P_{n_v-1,j}^{(k+1)}}{\Delta v} \end{cases} \quad (4.5)$$

iteratively from some initial guess of  $\bar{N}^{(0)}$ , and take  $P_{i,j}^H = P_{i,j}^{(K)}$ ,  $N_j^H = N_j^{(K)}$ ,  $\bar{N}^H = \bar{N}^{(K)}$  for sufficiently large  $K$  that satisfies some stopping criterion like  $|N^{(K+1)} - N^{(K)}| \leq \epsilon$  for some small enough  $\epsilon$ .

The numerical experiments are conducted as follows:

1. Fix all the scheme parameters except  $\epsilon$  (these parameters include all the grid lengths, initial/boundary values and coefficients, especially,  $\Delta t$  is fixed), and solve the scheme. The solution for a fixed  $\epsilon$  is denoted as  $p_\epsilon$ .

- Having solved for  $p_\varepsilon$ , find  $H_\varepsilon$  by (2.6), and thereafter the corresponding discretized  $v$ -wise quasi-steady state  $P^{H_\varepsilon}$  by (4.1)~(4.4), which can be numerically obtained by implementing the iterations described in (4.5);
- Plot  $\|p_\varepsilon^m - P^{H_\varepsilon^m}\|$  versus  $t^m$  for each  $\varepsilon$  (the norm is taken at a single time layer, chosen as  $L^1$  norm in our numerical test). AP property means that after a short period of time (the length of which decreases with  $\varepsilon$ ), the difference between the solution  $p_\varepsilon^m$  and the corresponding discretized  $v$ -wise quasi-steady state  $P^{H_\varepsilon^m}$  should drop to a small value which vanishes as  $\varepsilon \rightarrow 0$ , i.e.

$$\|p_\varepsilon^m - P^{H_\varepsilon^m}\| \xrightarrow{\varepsilon \rightarrow 0} 0. \quad (4.6)$$

In our simulations, we choose  $V_F = 2, V_R = 1, V_{\min} = -4, a = 1, W_{\min} = -1.1, W_{\max} = 0.1, T_{\max} = 0.3, \sigma(\bar{N}) = \bar{N}, \Delta t = 5 \times 10^{-4}, \Delta v = 0.1, \Delta w = 0.01$ ,

$$p_{\text{init}} = \begin{cases} \sin^2(\pi v) \sin^2(\pi w) & -1 < w < 0 \text{ and } -1 < v < 1 \\ 0 & \text{otherwise} \end{cases} \quad (4.7)$$

and

$$I(w) = \frac{1}{2} e^{-(10w+5)^2} \quad (4.8)$$

We perform the aforementioned AP test for both the  $v$ -FI scheme and the  $v$ -SI scheme, choosing  $\varepsilon$  ranging from  $10^{-1}$  to  $10^{-7}$ , and the results are shown in Figure 2 and Figure 3, respectively.

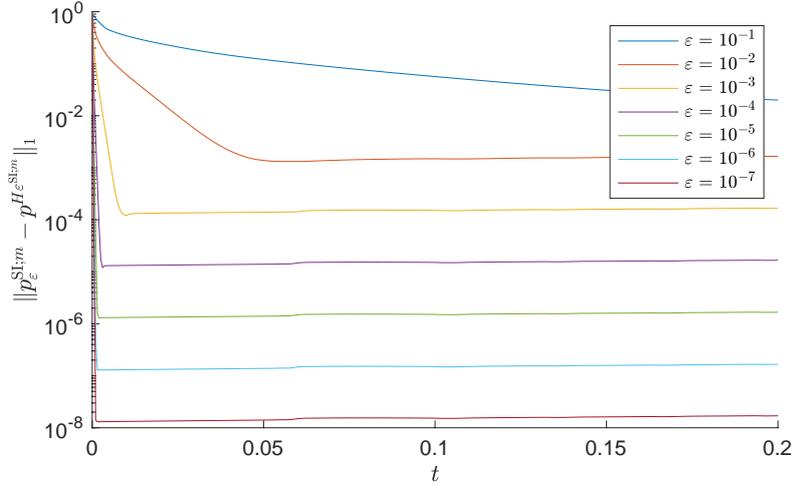


Figure 2: AP property of the  $v$ -FI scheme.  $\|p_\varepsilon^{\text{FI};m} - P^{H_\varepsilon^{\text{FI};m}}\|$  is plotted over  $t \in [0, T]$ , discretized with the time step  $\Delta t = 5 \times 10^{-4}$ . The discretized quasi-steady state  $P^{H_\varepsilon^{\text{FI};m}}$  is obtained by the iterations described in (4.5). It can be roughly seen that when  $\varepsilon$  is small,  $\|p_\varepsilon^{\text{FI};m} - P^{H_\varepsilon^{\text{FI};m}}\| = o(\varepsilon)$  after a few time steps.

From Figure 2, we can see that for the  $v$ -FI scheme, the differences between  $p_\varepsilon^m$  and  $P^{H_\varepsilon^m}$  is decreasing when  $\varepsilon \rightarrow 0$ , we can even assert  $\|p_\varepsilon^m - P^{H_\varepsilon^m}\| = o(\varepsilon)$  from the figure.

From Figure 3, we can see that for the  $v$ -SI scheme, the differences between  $p_\varepsilon^m$  and  $P^{H_\varepsilon^m}$  is also decreasing when  $\varepsilon \rightarrow 0$ . However, this decreasing trend is halting when  $\varepsilon = 10^{-5}$  for

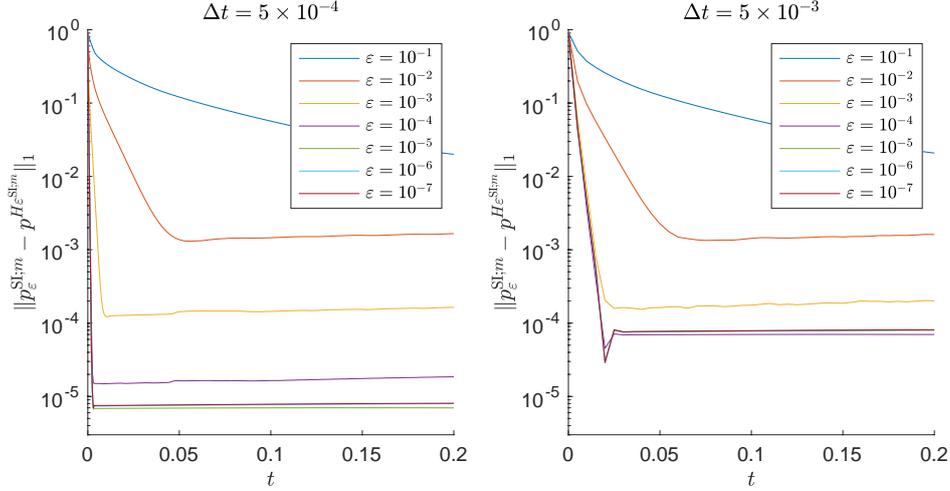


Figure 3: AP property of the  $v$ -SI scheme.  $\|p_\epsilon^{\text{SI};m} - P^{H_\epsilon^{\text{SI};m}}\|$  is plotted over  $t \in [0, T]$ , discretized with the grid size  $\Delta t = 5 \times 10^{-4}$  (left) and  $5 \times 10^{-3}$  (right). The discretized quasi-steady state  $P^{H_\epsilon^{\text{SI};m}}$  is obtained by the iterations described in (4.5). Comparing to the  $v$ -FI scheme, the property  $\|p_\epsilon^{\text{SI};m} - P^{H_\epsilon^{\text{SI};m}}\|$  is of order  $\epsilon$  when  $\epsilon$  is comparable to or greater than  $\Delta t$ . When  $\epsilon \ll \Delta t$  (which is the case when  $\epsilon = 10^{-5}, 10^{-6}, 10^{-7}$  in the figure, represented by the green, cyan and blue curves, respectively),  $\|p_\epsilon^{\text{SI};m} - P^{H_\epsilon^{\text{SI};m}}\|$  becomes roughly of order of  $\Delta t$ , so smaller  $\epsilon$  does not make this difference uniformly smaller as it does for the  $v$ -FI scheme.

$\Delta t = 5 \times 10^{-4}$  and  $\epsilon = 10^{-4}$  for  $\Delta t = 5 \times 10^{-3}$ , and further decrease of  $\epsilon$  does not uniformly draw  $p_\epsilon^m$  and  $P^{H_\epsilon^m}$  closer. This can be explained by the time-delayed effect of the  $v$ -SI scheme, which was presented in section 3.4. This time-delay effect introduces an  $o(\Delta t)$  difference between  $p_\epsilon^m$  and  $P^{H_\epsilon^m}$ , which will not further decrease when  $\epsilon$  becomes a small number comparing to  $\Delta t$ . We may say that  $\|p_\epsilon^m - P^{H_\epsilon^m}\| = o(\Delta t)$  when  $\epsilon \ll \Delta t$  and  $\|p_\epsilon^m - P^{H_\epsilon^m}\| = o(\epsilon)$  when  $\epsilon \gg \Delta t$  for the  $v$ -SI scheme.

### 4.3 Learning and reacting

From this subsection on, we explore model (1.8) with our scheme. We always use the  $v$ -SI scheme in the rest of this section.

In this subsection, we test the model's learning and discrimination abilities. We let the model perform learning-testing tasks that was originally proposed in [16], containing a learning phase and a testing phase:

**Learning Phase** A heterogeneous input  $I(w)$  is presented to the system, while the learning process is active. The learning rule is determined for the inhibitory weights by  $-N(w)\bar{N}$  by taking  $K(w) = -1$  if  $w \leq 0$ . After some time, the synaptic weight distribution  $H(w, t)$  converges to an equilibrium distribution  $H_I^*(w)$ , which depends on  $I$ .

**Testing Phase** A new input  $J(w)$  is presented to the system. This time, the neural system reacts to  $J$  but does not learn it. In this stage, there is only  $v$ -wise transport and no  $w$ -wise convection. To achieve this goal, we implement the  $v$ -wise quasi-steady state solver (4.5) with  $I$  replaced by the testing signal  $J$ , during which we keep  $H_I^*(w)$  we obtained in the Learning Phase.

The pattern recognition property of the system is that it can detect whether the new input  $J(w)$  is actually the same one that has been presented during the learning phase, i.e.  $I(w)$ : indeed, in this case,  $N_{I,J}^*(w) = w\mathbf{1}[-A, 0)$  has a very specific shape that does not depend upon the original input  $I$  that has been learned in the learning phase.

For the learning phase, We implement the  $v$ -SI scheme; for the testing phase, we implement the  $v$ -wise quasi-steady state solver (4.5), replacing the input function  $I(w)$  with the testing input function  $J(w)$ .

In our numerical experiments, the learning and testing input functions are chosen from the set

$$\{\psi_i(10w + 5) + 1 | i = 0, 1, \dots, 4\} \quad (4.9)$$

where  $\psi_i(x)$  stands for the  $i$ -th normalized Hermite function defined recursively as

$$\begin{aligned} \psi_0(y) &\equiv \pi^{-\frac{1}{4}} e^{-\frac{1}{2}y^2}; & \psi_1(y) &\equiv \pi^{-\frac{1}{4}} \sqrt{2} y e^{-\frac{1}{2}y^2} \\ \psi_{n+1}(y) &= \sqrt{\frac{2}{n+1}} y \psi_n(y) - \sqrt{\frac{n}{n+1}} \psi_{n-1}(y). \end{aligned} \quad (4.10)$$

This set of functions are well known for forming a set of bases for the Hilbert space  $L^2(-\infty, +\infty)$ . We take the argument  $10w + 5$  so that the input functions are centralized at  $w = -1/2$  and almost entirely supported in our truncated solving region; and the extra “+1” term widens the support of the solution so that it can react to input signals at wider range of  $w$ . We choose  $V_F = 2, V_R = 1, V_{\min} = -4, a = 1, \varepsilon = 0.1, W_{\min} = -1.1, W_{\max} = 0.1, T_{\max} = 5, \sigma(\bar{N}) = \bar{N}, \Delta t = 0.005, \Delta v = 0.1, \Delta w = 0.01$ , and  $p_{\text{init}}$  is given as (4.7).

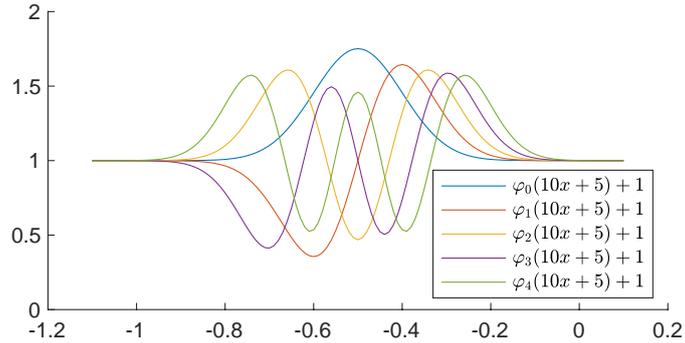


Figure 4: The candidate input functions  $\varphi_i(10w + 5) + 1$  where  $i = 0, 1, 2, 3, 4$  and  $\varphi_i$  stands for the  $i$ -th normalized Hermite functions.

For each  $i, j = 0, 1, \dots, 4$ , we let the model perform the fore-mentioned task with  $I(w) = \psi_i(10w + 5)$  and  $J(w) = \psi_j(10w + 5)$ , and the corresponding network activity  $N(w)$  is shown in the  $(i+1)$ -th row,  $(j+1)$ -th column in Figure 5. It can be seen that, when the learned input  $I(w)$  and the testing input  $J(w)$  are the same (corresponding to the diagonal entries in the figure), the firing pattern is of the nice triangular shape  $N_{I,J}^*(w) = w\mathbf{1}[-A, 0)$ ; but when  $I(w)$  and  $J(w)$  are different, the firing pattern is not in a regular shape. It is worth noting that in [16], the authors provided a primitive numerical result for the same learning-testing task, the networks firing pattern after reacting to the same learned input signal, shown in the article’s Figure 3, turned out to be quite zigzag. In comparison, with our numerical method, an almost-perfect triangular-shaped pattern is produced after the model learn and react to the same input (see the subfigures in the diagonal entries in Figure 5).

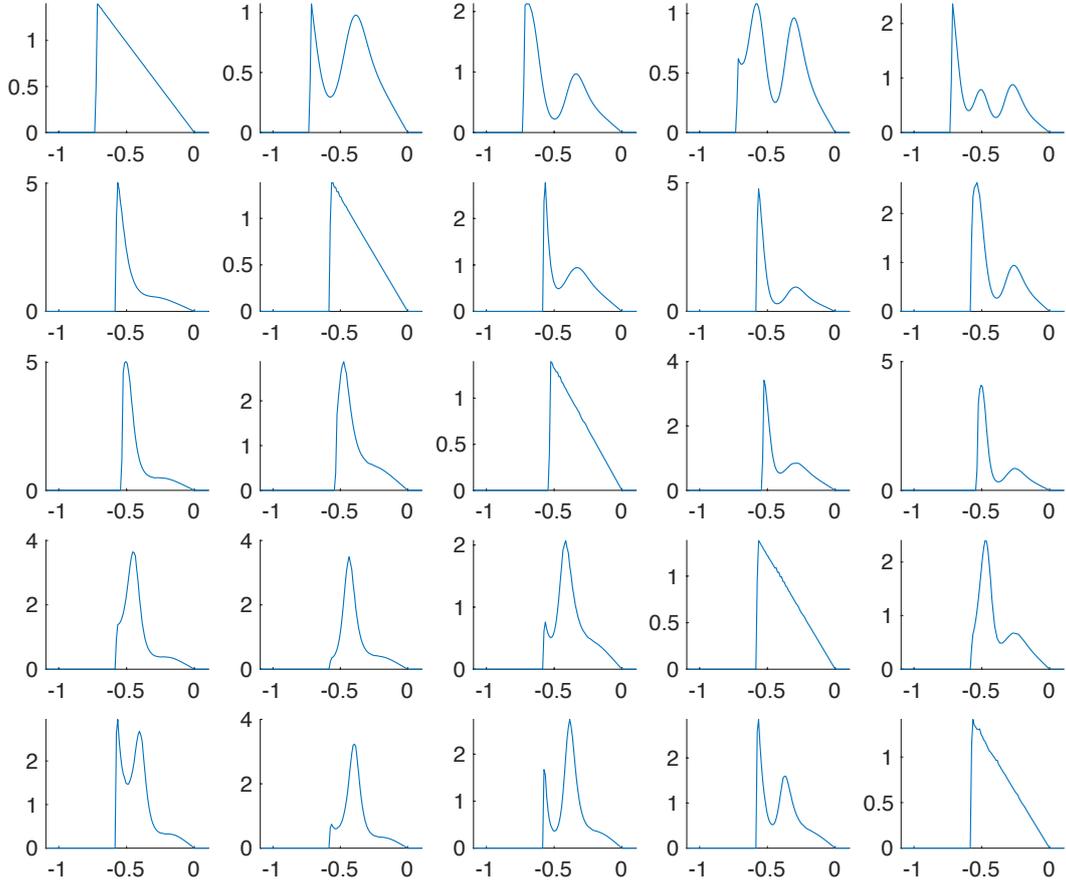


Figure 5: Test on the neural network’s recognition property. The pattern recognition task described at the beginning of subsection 4.3 is done for all possible pairs of input functions given in (4.9). For each  $i, j = 0, 1, \dots, 4$ , the subfigure at the  $(i+1)$ -th row,  $(j+1)$ -th column gives the final firing pattern  $N_{I,J}^*(w)$  versus after learning the signal  $I(w) = \varphi_i(10w + 5) + 1$  and reacting to the signal  $J(w) = \varphi_j(10w + 5) + 1$ . For each subfigure, the horizontal coordinate represents  $w$ , and the vertical axis represents  $N(w)$  at the end of the task. It can be seen that, when the learned input  $I(w)$  and the testing input  $J(w)$  are the same (corresponding to the diagonal entries in the figure), the firing pattern is of the nice triangular shape  $N_{I,J}^*(w) = w\mathbf{1}[-A, 0)$ ; but when  $I(w)$  and  $J(w)$  are different, the firing pattern will not be regular-shaped.

## 4.4 Positive synaptic weights

This subsection focus on exploring model's behavior when the neural system is excitatory, i.e. the solution is supported in the with region with  $w \geq 0$ . In [16], the authors' analytic work focus mainly on inhibitory networks. And for the excitatory neural system, they provided that, under some choice of parameters, the steady-state solution may not exist. We further explore with our numerical scheme the behavior of the solution when the synaptic weights are positive.

Our numerical exploration shows that in the excitatory cases, the network can either develop to a steady state like it did in the inhibitory cases, or exhibit a trend of accelerating expansion unlike it did with inhibitory cases under different parameter settings.

### 4.4.1 Long-time steady solution

We firstly test some scenarios in which the solutions present behavior similar to it would present with negative synaptic weights  $w$ .

We choose  $I \equiv 1$ ,  $a = 1$ ,  $\varepsilon = 0.2$ ,  $\sigma(\bar{N}) = \bar{N}/(1 + \bar{N})$ ,  $\Delta t = 5 \times 10^{-3}$ ,  $\Delta v = 0.1$ ,  $\Delta w = 0.01$ ,  $T_{\max} = 5$ ,  $K(w) \equiv 1$  and initial condition given as (4.7). The solution finally develops to a steady-state with  $N^*(w)$  forming a right triangular pattern supported in a single interval  $w \in [0, A]$  for some  $A > 0$ , just like it did with negative synaptic weights. This is the case that corresponds to the correct biological picture.

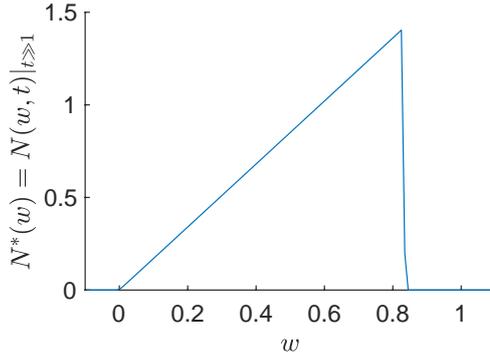


Figure 6: Solution with positive synaptic weight. With  $I = 0$ ,  $a = 1$ ,  $\varepsilon = 0.2$ , initial condition given as (4.7),  $\sigma(\bar{N}) = \bar{N}/(1 + \bar{N})$ ,  $\Delta t = 5 \times 10^{-3}$ ,  $\Delta v = 0.1$ ,  $\Delta w = 0.01$ ,  $T_{\max} = 5$ ,  $K(w) \equiv 1$ . The solution finally develops to a steady-state with  $N^*(w)$  forming the right triangular pattern just like it did with negative synaptic weights.

Indeed, as long as it can develop to a steady state, the excitatory neural network also has the recognition ability shown in subsection 4.3. To show this, we perform a similar learning-testing task as we did in the last subsection. The parameter settings are the same as they were in the last subsection except for that the initial condition  $p_{\text{init}}$  is supported in the region with  $w \geq 0$ ,  $K(w) \equiv 1$ ,  $\sigma(\bar{N}) = \bar{N}/(1 + \bar{N})$  and the candidate input functions become

$$\psi_i(10w - 5) + 1 \quad i = 0, 1, \dots, 4 \quad (4.11)$$

so that the input are centered at  $+1/2$  instead of  $-1/2$  for the inhibitory cases. The results are given in Figure 7. The arrangement of the figures are the same as that of Figure 5. It can be

seen that  $w - N(w)$  subfigures at the diagonal entries are in perfect right triangular shape, which is the same as it is in Figure 7.

#### 4.4.2 Unsteady solution

Now we turn our attention to a different case, where the solution is not converging to a steady solution. In [16], it was presented that even if we choose the firing function  $\sigma$  to be bounded

$$\sigma(\bar{N}) = k \frac{\bar{N}}{1 + \bar{N}}, \quad (4.12)$$

we may still face some scenarios that a steady state does not theoretically exist. And we further explore the solution's behavior in this case.

We choose  $I(w) \equiv 1$ ,  $k = 3$ , and the rest parameters were same as it was in the first numerical experiment in sub-subsection 4.4.1. The solution turns out to propagate to the region with greater  $w$ . The  $w$ -wise right boundary of the support of  $N$  grows increasingly fast (Figure 8 right), and henceforth  $\bar{N}$  grows increasingly big (Figure 8 left). It should be noticed that our result shown in Figure 8 does not necessarily imply that this is a blow up solution, since we are not sure whether  $\bar{N}(t)$  will grow to  $\infty$  in a finite time.

Even if the solution is not a blow-up one, it still does not practically make sense that the network's synaptic weight grows increasingly fast to  $\infty$ . Therefore, the solution's fast growing behavior is a non-physical effect of our model.

## 5 Conclusion

In this work, we have proposed a numerical scheme for the Fokker-Planck equation of structured neural networks (1.8), which is shown to be conservative, positivity-preserving, and asymptotic-preserving.

The key to success of our scheme is the implicit treatment for flux shift operators in (1.8), based on which the numerical scheme has an improved positivity preserving property: there is no  $v$ -wise grid ratio constraint to guarantee the positivity, which is otherwise a major bottleneck for efficient simulation.

As for numerical exploration, we have conducted the learning-testing numerical experiments and yielded numerical solutions that not only match the theoretical results perfectly, but also fully demonstrate the recognition ability of the model and provide insights for potential uses and future studies. The numerical tests are also extended to the cases of excitatory neural networks. We have also explored with our numerical experiments the behavior of the model when a steady state solution does not exist. These numerical experiments are successful in capturing partial long time trends of the solution but still preliminary. Further numerical experiments can be done to explore, for example, the model's behavior when the input signal  $I$  becomes time-dependent, or alternative learning rules that can prevent  $\bar{N}(t)$  from growing to infinity for the excitatory cases.

## Acknowledgement

J. Hu's research was partially supported by NSF CAREER grant DMS-1654152. Z. Zhou is supported by the National Key R&D Program of China, Project Number 2020YFA0712000 and NSFC grant No. 11801016, No. 12031013.

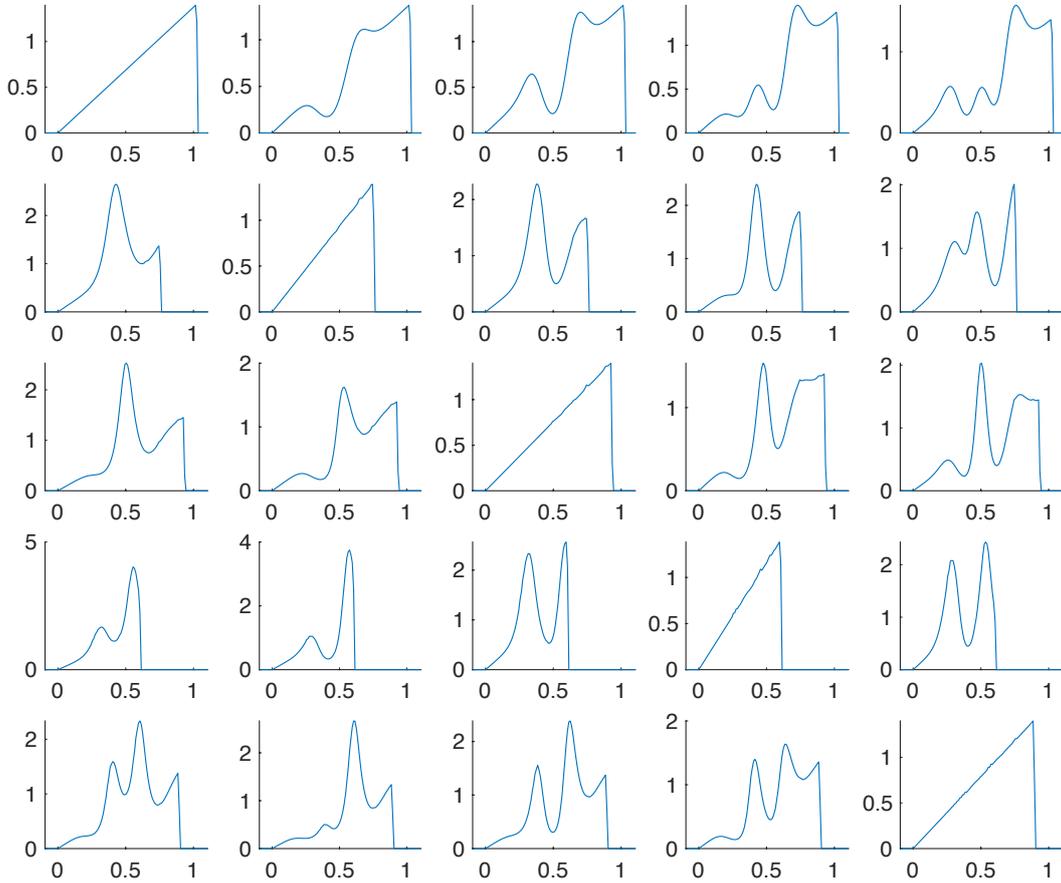


Figure 7: Test of recognition property of excitatory neural networks. The pattern recognition task described at the beginning of subsection 4.3 is done for all possible pairs of input functions given in (4.9). For each  $i, j = 0, 1, \dots, 4$ , the subfigure at the  $(i+1)$ -th row,  $(j+1)$ -th column gives the final firing pattern  $N_{I,J}^*(w)$  versus  $w$  after learning the signal  $I(w) = \varphi_i(10w - 5) + 1$  and reacting to the signal  $J(w) = \varphi_j(10w - 5) + 1$ . For each subfigure, the horizontal coordinate represents  $w$ , and the vertical axis represents  $N(w)$  at the end of the task. It can be seen that, when the learned input  $I(w)$  and the testing input  $J(w)$  are the same (corresponding to the diagonal entries in the figure), the firing pattern is of the nice triangular shape  $N_{I,J}^*(w) = w\mathbf{1}(0, A)$ ; but when  $I(w)$  and  $J(w)$  are different, the firing pattern will not be regular-shaped.

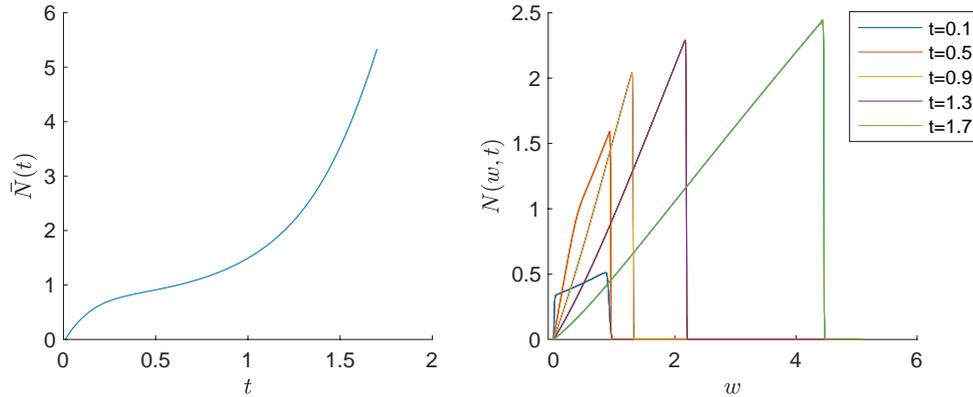


Figure 8: Solution with positive synaptic weight. With  $I(w) \equiv 1$ ,  $a = 1$ ,  $\varepsilon = 0.2$ , initial condition given as (4.7),  $\sigma(\bar{N}) = 3\bar{N}/(1 + \bar{N})$ ,  $\Delta t = 5 \times 10^{-3}$ ,  $\Delta v = 0.1$ ,  $\Delta w = 0.01$ ,  $T_{\max} = 5$ ,  $K(w) \equiv 1$ . Left: in the unsteady scenarios,  $\bar{N}$  grows with accelerate. Right: The solution turns out to propagate to the region with greater  $w$ . The  $w$ -wise right boundary of the support of  $N$  grows increasingly fast.

## References

- [1] Stephen J Martin, Paul D Grimwood, and Richard GM Morris. Synaptic plasticity and memory: an evaluation of the hypothesis. *Annual review of neuroscience*, 23(1):649–711, 2000.
- [2] Alan L Hodgkin and Andrew F Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500, 1952.
- [3] Richard FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1(6):445–466, 1961.
- [4] Louis Lapicque. Recherches quantitatives sur l’excitation électrique des nerfs traitée comme une polarisation. *J. Physiol. Pathol. Gen.*, 9(1):620–635, 1907.
- [5] María J Cáceres, José A Carrillo, and Benoît Perthame. Analysis of nonlinear noisy integrate & fire neuron models: blow-up and steady states. *The Journal of Mathematical Neuroscience*, 1(1):1–33, 2011.
- [6] José A Carrillo, María d M González, Maria P Galdani, and Maria E Schonbek. Classical solutions for a nonlinear fokker-planck equation arising in computational neuroscience. *Communications in Partial Differential Equations*, 38(3):385–409, 2013.
- [7] María J. Cáceres, Pierre Roux, Delphine Salort, and Ricarda Schneider. Global-in-time classical solutions and qualitative properties for the nlif neuron model with synaptic delay. *Communications in Partial Differential Equations*, 2018.
- [8] Nicolas Brunel. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of computational neuroscience*, 8(3):183–208, 2000.

- [9] Grégory Dumont and Pierre Gabriel. The mean-field equation of a leaky integrate-and-fire neural network: measure solutions and steady states. *Nonlinearity*, 33, 10 2017.
- [10] Nicolas Brunel and Vincent Hakim. Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural computation*, 11(7):1621–1671, 1999.
- [11] François Delarue, James Inglis, Sylvain Rubenthaler, and Etienne Tanré. Global solvability of a networked integrate-and-fire model of mckean–vlasov type. *The Annals of Applied Probability*, 25(4):2096–2133, 2015.
- [12] François Delarue, James Inglis, Sylvain Rubenthaler, and Etienne Tanré. Particle systems with a singular mean-field self-excitation. application to neuronal networks. *Stochastic Processes and their Applications*, 125(6):2451–2492, 2015.
- [13] Jian-guo Liu, Ziheng Wang, Yuan Zhang, and Zhennan Zhou. Rigorous justification of the fokker-planck equations of neural networks based on an iteration perspective. *arXiv preprint arXiv:2005.08285*, 2020.
- [14] Jian-Guo Liu, Ziheng Wang, Yantong Xie, Yuan Zhang, and Zhennan Zhou. Investigating the integrate and fire model as the limit of a random discharge model: a stochastic analysis perspective. *Mathematical Neuroscience and Applications*, 1, 2021.
- [15] Donald Olding Hebb. *The organization of behavior: A neuropsychological approach*. John Wiley & Sons, 1949.
- [16] Benoît Perthame, Delphine Salort, and Gilles Wainrib. Distributed synaptic weights in a LIF neural network and learning rules. *Physica D: Nonlinear Phenomena*, 353-354:20–30, Sep 2017.
- [17] Shi Jin and Bokai Yan. A class of asymptotic-preserving schemes for the fokker-planck-landau equation. *J. Comput. Physics*, 230:6420–6437, 07 2011.
- [18] Marianne Bessemoulin-Chatard and Francis Filbet. A finite volume scheme for nonlinear degenerate parabolic equations. *SIAM Journal on Scientific Computing*, 34(5):B559–B583, 2012.
- [19] José A Carrillo, Alina Chertock, and Yanghong Huang. A finite-volume method for nonlinear nonlocal equations with a gradient flow structure. *Communications in Computational Physics*, 17(1):233–258, 2015.
- [20] JS Chang and G Cooper. A practical difference scheme for fokker-planck equations. *Journal of Computational Physics*, 6(1):1–16, 1970.
- [21] Lorenzo Pareschi and Mattia Zanella. Structure preserving schemes for nonlinear fokker–planck equations and applications. *Journal of Scientific Computing*, 74(3):1575–1600, 2018.
- [22] Jingwei Hu and Xiangxiong Zhang. Positivity-preserving and energy-dissipative finite difference schemes for the fokker-planck and keller-segel equations. *arXiv preprint arXiv:2103.16790*, 2021.
- [23] Sebastiano Boscarino, Francis Filbet, and Giovanni Russo. High order semi-implicit schemes for time dependent partial differential equations. *Journal of Scientific Computing*, 68(3):975–1001, 2016.

- [24] Jian-Guo Liu, Li Wang, and Zhennan Zhou. Positivity-preserving and asymptotic preserving method for 2d keller-segal equations. *Mathematics of Computation*, 87(311):1165–1189, 2018.
- [25] Jingwei Hu and Xiaodong Huang. A fully discrete positivity-preserving and energy-dissipative finite difference scheme for poisson–nernst–planck equations. *Numerische Mathematik*, 145(1):77–115, 2020.
- [26] Rafael Bailo, Jose A Carrillo, and Jingwei Hu. Fully discrete positivity-preserving and energy-dissipating schemes for aggregation-diffusion equations with a gradient flow structure. *Commun. Math. Sci.*, 18:1259–1303, 2020.
- [27] Yong Zhang, Yu Zhao, and Zhennan Zhou. A unified structure preserving scheme for a multispecies model with a gradient flow structure and nonlocal interactions via singular kernels. *SIAM Journal on Scientific Computing*, 43(3):B539–B569, 2021.
- [28] Jingwei Hu and Ruiwen Shu. A second-order asymptotic-preserving and positivity-preserving exponential runge–kutta method for a class of stiff kinetic equations. *Multiscale Modeling & Simulation*, 17(4):1123–1146, 2019.
- [29] Randall J. LeVeque. *Numerical methods for conservation laws (2. ed.)*. Lectures in mathematics. Birkhäuser, 1992.
- [30] Jingwei Hu, Jian-Guo Liu, Yantong Xie, and Zhennan Zhou. A structure preserving numerical scheme for fokker-planck equations of neuron networks: Numerical analysis and exploration. *Journal of Computational Physics*, 433:110195, 2021.
- [31] Robert J Plemmons. M-matrix characterizations. i—nonsingular m-matrices. *Linear Algebra and its Applications*, 18(2):175–188, 1977.