Les Cahiers du GERAD ISSN: 0711–2440

Transmission Neural Networks: Inhibitory and excitatory connections

S. Gao, P. E. Caines

G–2026–62

October 2026

La collection *Les Cahiers du GERAD* est constituée des travaux de recherche menés par nos membres. La plupart de ces documents de travail a été soumis à des revues avec comité de révision. Lorsqu'un document est accepté et publié, le pdf original est retiré si c'est nécessaire et un lien vers l'article publié est ajouté.

Citation suggérée : S. Gao, P. E. Caines (Octobre 2026). Transmission Neural Networks: Inhibitory and excitatory connections, Rapport technique, Les Cahiers du GERAD G– 2026–62, GERAD, HEC Montréal, Canada.

Avant de citer ce rapport technique, veuillez visiter notre site Web (https://www.gerad.ca/fr/papers/G-2026-62) afin de mettre à jour vos données de référence, s'il a été publié dans une revue scientifique.

The series *Les Cahiers du GERAD* consists of working papers carried out by our members. Most of these pre-prints have been submitted to peer-reviewed journals. When accepted and published, if necessary, the original pdf is removed and a link to the published article is added.

Suggested citation: S. Gao, P. E. Caines (October 2026). Transmission Neural Networks: Inhibitory and excitatory connections, Technical report, Les Cahiers du GERAD G–2026–62, GERAD, HEC Montréal, Canada.

Before citing this technical report, please visit our website (https://www.gerad.ca/en/papers/G-2026-62) to update your reference data, if it has been published in a scientific journal.

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2026
– Bibliothèque et Archives Canada, 2026

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2026
– Library and Archives Canada, 2026

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Transmission Neural Networks: Inhibitory and excitatory connections

Shuang Gao [a, c, d]

Peter E. Caines [b, c]

[a] *Department of Electrical Engineering, Polytechnique Montréal, Montréal (Qc), Canada, H3T 1J4*

[b] *Department of Electrical and Computer Engineering, McGill University, Montréal (Qc), Canada, H3T 2A7*

[c] *GERAD, Montréal (Qc), Canada, H3T 1J4*

[d] *UNIQUE (Unifying Neuroscience and Artificial Intelligence – Quebec), Montréal (Qc), Canada, H3T 1P1*

shuang.gao@polymtl.ca
peterc@cim.mcgill.ca

October 2026
Les Cahiers du GERAD
G–2026–62

Copyright © 2026 Gao, Caines

Les textes publiés dans la série des rapports de recherche *Les Cahiers du GERAD* n'engagent que la responsabilité de leurs auteurs. Les auteurs conservent leur droit d'auteur et leurs droits moraux sur leurs publications et les utilisateurs s'engagent à reconnaître et respecter les exigences légales associées à ces droits. Ainsi, les utilisateurs:

- Peuvent télécharger et imprimer une copie de toute publication du portail public aux fins d'étude ou de recherche privée;
- Ne peuvent pas distribuer le matériel ou l'utiliser pour une activité à but lucratif ou pour un gain commercial;
- Peuvent distribuer gratuitement l'URL identifiant la publication.

Si vous pensez que ce document enfreint le droit d'auteur, contactez-nous en fournissant des détails. Nous supprimerons immédiatement l'accès au travail et enquêterons sur votre demande.

The authors are exclusively responsible for the content of their research papers published in the series *Les Cahiers du GERAD*. Copyright and moral rights for the publications are retained by the authors and the users must commit themselves to recognize and abide the legal requirements associated with these rights. Thus, users:

- May download and print one copy of any publication from the public portal for the purpose of private study or research;
- May not further distribute the material or use it for any profit-making activity or commercial gain;
- May freely distribute the URL identifying the publication.

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Abstract : This paper extends the Transmission Neural Network model proposed by Gao and Caines in [9–11] to incorporate inhibitory connections and neurotransmitter populations. The extended network model contains binary neuronal states, transmission dynamics, and inhibitory and excitatory connections. Under technical assumptions, we establish the characterization of the firing probabilities of neurons, and show that such a characterization in the presence of inhibition can be equivalently represented by a neural network where each neuron has a continuous state of dimension 2. In addition, we incorporate neurotransmitter populations into the modeling and establish the limit network model when the number of neurotransmitters at all synaptic connections goes to infinity. Finally, sufficient conditions for stability of the proposed network models are established.

Acknowledgements: This work is supported by NSERC (Canada) Grant RGPIN–2024–06612. SG gratefully acknowledges Roland P. Malhamé and Evelyn Hubbard for their helpful feedback on this work.

1 Introduction

Modelling neuronal systems is important to understand intelligence and to analyze and control such systems. Networks of neurons can learn input-output relations in the context of artificial neural networks [3, 13, 23, 24]. Moreover, recent advances show that combining detailed brain networks, such as the Drosophila connectome, with relatively simple neuronal dynamics can predict neural activities associated with specific sensorimotor processing [26]. Neuronal models with different levels of abstraction have been proposed to characterize the behaviors of neuronal systems [4,12], ranging from the detailed descriptions of the dynamics of individual neurons by Hodgkin and Huxley [17] to network-level models characterizing interactions [12, 18, 19].

Neural network models that adopt a binary-state representation of neuronal systems offer certain advantages: (a) one can focus on network level properties as in the work of Hopfield [18] and those on Boltzmann Machines [1, 16] and (b) continuous-valued neural networks can be binarized to provide more efficient algorithms and learning models [6, 20]. In addition, the binary state for each neuron can be naturally linked with a continuous value by taking the probability of neurons being activated as a neuronal state [11, 19], for which control-theoretic properties including stability can be established (see e.g. [4, 11]).

The Transmission Neural Network (TransNN) model proposed in [9–11] has established a natural connection between neural networks and virus spread models, where the connection of the nodes resembles the process of synaptic transmission. The work [9, 10] investigated how TransNN approximates stochastic neural networks with binary nodal states, and proposed TransNN-based approximate control algorithms for controlling these stochastic networks.

Inhibition that suppresses the activity of neurons is essential for neuronal systems [7, 15, 22]. Inhibitory properties have been considered in models of neurons with different formulations (e.g. the Wilson-Cowan model for neuronal populations [28, 29] and models with continuous-time stochastic dynamics [5, 27]).

The main focus of the present paper is to generalize TransNN models [9, 10] to account for inhibitory connections. The central challenges are to identify a biologically motivated and mathematically tractable inhibition mechanism and to derive approximate models that capture its effect.

Contribution: This work extends the TransNN models with binary nodal states in [9,10] to include inhibition, and identifies the corresponding model for the probability of excitation. A key consequence of introducing inhibition is that the scalar-state simplification in [9, 10] no longer applies directly. We show that such networks with inhibition can be equivalently represented by neural networks where a two-dimensional nodal state is associated with each neuron and the Tuneable Log-Sigmoid activation function in [11] with each synaptic connection. Moreover, we incorporate neurotransmitter populations into an extended TransNN model and establish its limit model by letting the number of neurotransmitters at all synaptic connections go to infinity. Finally, sufficient conditions for stability of the proposed network models are established.

Notation: R denotes the set of real numbers. Let $\bar{\mathrm{R}} \triangleq \mathrm{R} \cup \{+\infty\}$ and $[n] \triangleq \{1, 2, ..., n\}$. We use $W \triangleq [W_{ij}] \in \mathrm{R}^{n\times n}$ to denote the matrix with its ij^{th} element specified by W_{ij} for all $i, j \in [n]$. For a vector $v \in \bar{\mathrm{R}}^n$, we use both v_i and $[v]_i$ to denote its i^{th} element. For a matrix $W \in \mathrm{R}^{n\times n}$, $\|W\|_1 \triangleq \max_{j\in[n]} \sum_{i=1}^n |W_{ij}|$ and $\|W\|_\infty \triangleq \max_{i\in[n]} \sum_{j=1}^n |W_{ij}|$. For a vector $v \in \mathrm{R}^n$, $\|v\|_1 \triangleq \sum_{i=1}^n |v_i|$ and $\|v\|_\infty \triangleq \max_{i\in[n]} |v_i|$. Let $\odot$ denote Hadamard product.

2 Transmission dynamics with inhibitory and excitatory connections

Consider a network of n neurons with interconnections through chemical synapses [21]. The synaptic connection structure at time or layer[1] $k \geq 0$ is represented by a directed graph $\mathcal{G}^k = ([n], \mathcal{E}^k)$ with the node set $[n] \triangleq \{1, 2, ..., n\}$ and the edge set $\mathcal{E}^k \subset [n] \times [n]$, which may include self-loops. A directed connection from neuron j to neuron i, denoted by the node pair $(i, j) \subset \mathcal{E}^k$, exists if the axon terminal of neuron j has at least one synapse onto neuron i at step k. Such synaptic connections could be either excitatory or inhibitory [21]. Let $\mathcal{G}_h^k = ([n], \mathcal{E}_h^k)$ denote the subgraph of $\mathcal{G}^k$ with all inhibitory connections, and $\mathcal{G}_e^k = ([n], \mathcal{E}_e^k)$ the subgraph of $\mathcal{G}^k$ with all excitatory connections. Furthermore, $\mathcal{E}_h^k \cup \mathcal{E}_e^k = \mathcal{E}^k$ and $\mathcal{E}_h^k \cap \mathcal{E}_e^k = \varnothing$, that is, a connection between two neurons will be either inhibitory or excitatory. We allow self-loops to model autapses (i.e. synapses from a neuron onto itself) in neuronal systems [2]. Furthermore, since an autapse is either inhibitory or excitatory [2], the self-loops considered are either inhibitory or excitatory, that is, for each node $i \in [n]$, its self-loop (i.e. the edge pair (i, i)) may appear in either $\mathcal{E}_h^k$ or $\mathcal{E}_e^k$, but not in both. Let B_E^k denote the binary-valued adjacency matrix of the subgraph $\mathcal{G}_e^k = ([n], \mathcal{E}_e^k)$ with all excitatory connections at step k, whose ij^{th} element is 1 if $(i, j) \in \mathcal{E}_e^k$, and 0 otherwise. Similarly, let B_I^k denote the binary-valued adjacency matrix of the subgraph $\mathcal{G}_h^k = ([n], \mathcal{E}_h^k)$ with all inhibitory connections at step k.

2.1 Transmission dynamics

The state of a neuron $i \in [n]$ at step k is denoted by a binary variable $X_i(k)$ that takes 1 if the neuron fires and 0 otherwise. We introduce the transmission dynamics with both excitatory and inhibitory connections as follows:

$$\begin{aligned} X_i(k+1) = &\Big(1 - \prod_{j \in E_i^{\circ k}} (1 - W_{ij}^k X_j(k))\Big) \\ &\times \prod_{j \in I_i^{\circ k}} (1 - W_{ij}^k X_j(k)) \end{aligned} \tag{1}$$

where $E_i^{\circ k} \triangleq \{j : (i, j) \in \mathcal{E}_e^k\}$ denotes the set of incoming neighbouring nodes of i with excitatory connection that potentially includes node i at step k, $I_i^{\circ k} \triangleq \{j : (i, j) \in \mathcal{E}_h^k\} = \{j : (i, j) \in \mathcal{E}^k - \mathcal{E}_e^k\}$ denotes the set of incoming neighbouring nodes of i with inhibitory connections that potentially includes node i at step k, and $W_{ij}^k \in \{0, 1\}$ is the binary variable that represents the successful transmission when taking value 1, otherwise 0.

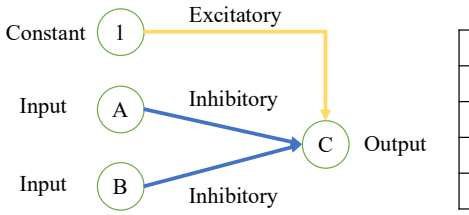
Remark 1. The network model with binary states in (1) generalizes that in [9,10] to include inhibitory neuronal connections. In the dynamics (1), a single effective inhibitory connection to a neuron suppresses its excitation, whereas in the absence of inhibition, a single effective excitatory connection is sufficient to activate it. This rule provides a coarse-grained representation of inhibitory suppression and is qualitatively reminiscent of the functional effect of shunting inhibition (also known as the short-circuiting effect) [21, p. 290], in the sense that inhibitory input can suppress the effect of excitatory input.

Remark 2 (Functional Completeness). The existence of a constant input is possible for neuronal systems (e.g. persistent firing for working memory [8]). With a constant input 1 as well as excitatory and inhibitory connections, the interaction rules in (1) can form an NOR gate with a simple network structure as illustrated in Fig. 1. Since any Boolean function can be implemented by combining NOR gates (i.e. the NOR gate is functionally complete), the interaction rule in (1) with the constant signal 1 can reproduce any Boolean function by appropriate network designs.

2.2 Stochasticity in the transmission dynamics

Now we consider the case where the state $X_i(k)$ and the transmission W_{ij}^k for $i \in [n]$ and for $k \geq 0$ are stochastic. Let $X(k) \triangleq [X_1(k), \cdots, X_n(k)]^\intercal$.

[1] We note that, throughout the paper, time step k can also be interpreted as layer k in the context of neural networks with multiple layers.



A	B	C
0	0	1
0	1	0
1	0	0
1	1	0

Figure 1: NOR gate with inputs A and B, and output C, created with inhibitory connections and a constant signal 1.

We introduce the assumptions on the independence of the transmissions and that of the states.

(A1) (Memoryless Transmission) For any $k > 0$, $W^k \triangleq [W_{ij}^k]$ is independent of $\{W^t : 0 \le t < k\}$ and $\{X(t) : 0 \le t < k\}$. W^0 is independent of $X(0)$.

(A2) (Transmission Conditional Independence) For any $k \ge 0$, the binary random variables $\{W_{ij}^k : i, j \in [n]\}$ representing the transmissions are jointly conditionally independent given the current states $X(k)$.

Remark 3. The assumption (A1) introduces the independence of the transmission at the current step (or layer) from all the past transmissions and states. (A2) imposes the conditional independence of transmissions across different links given the current state. A special case is as follows: W_{ij}^k is determined by flipping a biased coin independently with probability w_{ij}^k being head corresponding to the value 1, and $X_i(k)$ is determined similarly.

Under the assumption (A1), the transmission dynamics in (1) are Markovian and

$$\mathbb{E}(X_i(k+1)|(X(t))_{t\in[k]}) = \mathbb{E}(X_i(k+1)|X(k)). \tag{2}$$

Furthermore, under (A2), we have

$$\begin{aligned}\mathbb{E}(X_i(k+1)|X(k)) = &\Big(1 - \prod_{j\in E_i^{\circ k}} \mathbb{E}(1 - W_{ij}^k X_j(k)|X(k))\Big) \\ &\times \prod_{j\in I_i^{\circ k}} \mathbb{E}(1 - W_{ij}^k X_j(k)|X(k)).\end{aligned}$$

For a state configuration $q \in \{0,1\}^n$, the conditional probability of reaching q is given by

$$\Pr(X(k+1) = q|X(k)) = \prod_{i=1}^n \Pr(X_i(k+1) = q_i|X(k)) \tag{3}$$

where the equality is due to the conditional independence of the transmissions $\{W_{ij}^k : i, j \in [n]\}$ assumed in (A2).

Proposition 1. *Assume (A1) and (A2) hold. Given a state configuration $x \in \{0,1\}^n$ at step k, the transition probability to a state configuration $q \in \{0,1\}^n$ is given by*

$$\begin{aligned}&\Pr(X(k+1) = q|X(k) = x) \\ &\quad = \prod_{i=1}^n \Big(q_i\rho_i(k+1) + (1-q_i)(1-\rho_i(k+1))\Big)\end{aligned} \tag{4}$$

where

$$\begin{aligned}\rho_i(k+1) = &\Big(1 - \prod_{j\in E_i^{\circ k}} (1 - w_{ij}^k x_j)\Big) \\ &\times \prod_{j\in I_i^{\circ k}} (1 - w_{ij}^k x_j), \quad i \in [n]\end{aligned} \tag{5}$$

and $w_{ij}^k \triangleq \Pr(W_{ij}^k = 1|X_j(k) = 1)$.

Proof. Following (3), we note that

$$\begin{aligned}&\Pr(X(k+1)=q|X(k)=x)\\&=\prod_{i=1}^{n}\Pr(X_i(k+1)=q_i|X(k)=x).\end{aligned}\tag{6}$$

Then, by explicitly evaluating each probability $\Pr(X_i(k+1)=q_i|X(k)=x)$, using the dynamics (1), we obtain the desired result. □

The result above generalizes that of [10, Prop. 1] by including inhibitions in the transmission dynamics.

To further simplify the model, we introduce two other assumptions below.

(A3) (Transmission Independence at Step k) For $k\geq 0$, $\{W_{ij}^k : i,j\in[n]\}$ are independent and for each $i,j\in[n]$, W_{ij}^k is independent of $\{X_q(k) : q\in[n], q\neq j\}$.

(A4) (State Independence among Neurons) Up to some terminal step T, for each $k\leq T$, the states $\{X_i(k) : i\in[n]\}$ are independent.

Remark 4. The assumption (A3) introduces the independence of the transmissions across different links and their independence of each link W_{ij}^k from the states of neurons other than its source neuron j, and (A3) is more restrictive than (A2). (A4) can be satisfied with appropriate tree-like network structures and small time horizon T that avoid state correlations. Such structures as well as the characterization of the error incurred in terms of the second moment when (A4) does not hold will be presented in the journal extension.

Under (A1), (A3) and (A4), taking the expectation on both sides of the Equation (1) yields

$$\begin{aligned}\mathbb{E}X_i(k+1)=&\Big(1-\prod_{j\in E_i^{\circ k}}\mathbb{E}(1-W_{ij}^kX_j(k))\Big)\\&\times\prod_{j\in I_i^{\circ k}}\mathbb{E}(1-W_{ij}^kX_j(k))\end{aligned}\tag{7}$$

for every $k\in\{0,\dots,T-1\}$. Let $w_{ij}^k\triangleq\Pr(W_{ij}^k=1|X_j(k)=1)$ denote the conditional probability of the successful transmission from node j to node i at step k. Let $p_i(k)\triangleq\Pr(X_i(k)=1)$. Then the Equation (7) is equivalently represented by

$$\begin{aligned}p_i(k+1)=&\Big(1-\prod_{j\in E_i^{\circ k}}(1-w_{ij}^kp_j(k))\Big)\\&\times\prod_{j\in I_i^{\circ k}}(1-w_{ij}^kp_j(k)).\end{aligned}\tag{8}$$

Therefore, we obtain the following result.

Proposition 2. *Assume that (A1), (A3) and (A4) hold. Consider $X_i(k)$ the stochastic binary state of the dynamics (1) and $p_i(k)$ the probability state of the dynamics (8). If $\Pr(X_i(0)=1)=p_i(0)$ for all $i\in[n]$, then*

$$\Pr(X_i(k)=1)=p_i(k),\quad\forall i\in[n],\ 0\leq k\leq T.$$

2.3 Dynamics under state transformation

The state transformation that simplifies the model in [9–11] does not apply directly to the model (8) due to the existence of the inhibition. To address this issue, we introduce the probability of no inhibition at the previous step $k-1$ (from all neighboring neurons) at neuron i below:

$$\pi_i(k)\triangleq\prod_{j\in I_i^{\circ k-1}}(1-w_{ij}^{k-1}p_j(k-1))\in[0,1].\tag{9}$$

Then the probability update in (8) is equivalently given by

$$p_i(k+1) = \Big(1 - \prod_{j \in E_i^{\circ k}} (1 - w_{ij}^k p_j(k))\Big)\pi_i(k+1) \tag{10}$$

which yields

$$1 - \frac{p_i(k+1)}{\pi_i(k+1)} = \prod_{j \in E_i^{\circ k}} (1 - w_{ij}^k p_j(k)). \tag{11}$$

From (10), clearly $p_i(k) \leq \pi_i(k)$ holds for all $i \in [n]$ and all $k \geq 1$. Define the following states (of Shannon information)

$$s_i(k) \triangleq \begin{cases} -\log\left(1 - \frac{p_i(k)}{\pi_i(k)}\right), & \text{if } \pi_i(k) \in (0,1] \\ 0, & \text{if } \pi_i(k) = 0 \end{cases} \tag{12}$$

$$o_i(k) \triangleq -\log \pi_i(k) \tag{13}$$

where $o_i(k), s_i(k) \in [0, +\infty]$ and log function is defined by

$$\log(x) \triangleq \begin{cases} \ln(x), & x \in (0,1]; \\ -\infty, & x = 0. \end{cases} \tag{14}$$

Remark 5 (State Interpretation). The state $o_i(k)$ for a neuron i satisfies

$$o_i(k) = -\log\Big(\prod_{j \in I_i^{\circ k-1}} (1 - w_{ij}^{k-1} p_j(k-1))\Big).$$

and hence is the Shannon information associated with the absence of inhibition from the neighboring neurons at the previous step $k-1$, which can be intuitively understood as the level of inhibition from neighboring neurons. In particular, the state $o_i(k) = 0$ represents no active inhibition at $k-1$ and $o_i(k) = \infty$ represents active inhibition at $k-1$.

The state $s_i(k)$ for a neuron i is essentially the Shannon information associated with the absence of excitations from the excitatory neighboring neurons at the previous step $k-1$, that is

$$s_i(k) = -\log\Big(\prod_{j \in E_i^{\circ k-1}} (1 - w_{ij}^{k-1} p_j(k-1))\Big).$$

In particular, the state $s_i(k) = 0$ if the neuron i is resting with probability 1 at time k (i.e. $p_i(k) = 0$), and $s_i(k) = +\infty$ if it fires with probability 1 (i.e. $p_i(k) = 1$).

Remark 6. The inverse mappings of the state transformations in (13) and (12) are respectively given by

$$\pi_i(k) = e^{-o_i(k)} \quad \text{and} \quad p_i(k) = e^{-o_i(k)}(1 - e^{-s_i(k)}).$$

Taking logarithm and negation on both sides of (11) yields the following representation of the evolution of $s_i(k)$

$$s_i(k+1) = \sum_{j \in E_i^{\circ k}} \Psi(w_{ij}^k \pi_j(k), s_j(k)) \tag{15}$$

where $\Psi(\cdot,\cdot)$ is the Tuneable Log-Sigmoid (TLogSigmoid) activation function identified in [11] and defined by

$$\Psi(w,x) \triangleq -\log(1 - w + we^{-x}),\ w \in [0,1], x \in [0,+\infty]. \tag{16}$$

Replacing $\pi_j(k)$ by $e^{-o_j(k)}$ from the relation (13) yields

$$s_i(k+1) = \sum_{j \in E_i^{\circ k}} \Psi(w_{ij}^k e^{-o_j(k)}, s_j(k)), \quad i \in [n] \tag{17}$$

with initial condition $s_i(0) = -\log(1-p_i(0))$ for all $i \in [n]$. Furthermore, taking logarithm and negation on both sides of (9) yields dynamics of state o_i as follows:

$$o_i(k+1) = \sum_{j \in I_i^{\circ k}} \Psi(w_{ij}^k e^{-o_j(k)}, s_j(k)), \quad i \in [n] \tag{18}$$

with initial condition $o_i(0) = 0$ for all $i \in [n]$, if there is no inhibition before the initial step.

Equations (17) and (18) then completely characterize the evolution of the state $(s_i(k), o_i(k))$ over time step $k \geq 0$. The evolution can be computed as follows:

1. Start with $o_i(0) = 0$ for all $i \in [n]$, if there is no inhibition before step $k = 0$.
2. Compute the state $s_i(0) = -\log(1-p_i(0))$ based on the probability of excitation $p_i(0)$ for all node $i \in [n]$.
3. Iteratively compute the states s_i and o_i over time based on the dynamics specified in (17) and (18).

Remark 7 (Convexity). We highlight that if $w \in (0,1)$ and $x \in (0,\infty)$, the activation function $\Psi(w,x) \triangleq -\log(1-w+we^{-x})$ is strictly convex in w and strictly concave in x, and is strictly monotonically increasing with respect to x and w by evaluating the derivatives (see [11, Section V]). Let

$$g(z; v, s) \triangleq \Psi(ve^{-z}, s), \quad v \in (0,1), s \in (0,+\infty)$$

Then $\partial_z g = \partial_w \Psi(ve^{-z}, s) v(-1)e^{-z}$, and

$$\begin{aligned} \partial_{zz}^2 g &= \partial_w \Psi(ve^{-z}, s) ve^{-z} + \partial_{ww}^2 \Psi(ve^{-z}, s)(v(-1)e^{-z})^2 \\ &= ve^{-z}(\partial_w \Psi(ve^{-z}, s) + \partial_{ww}^2 \Psi(ve^{-z}, s) ve^{-z}) > 0 \end{aligned}$$

since $\partial_w \Psi(ve^{-z}, s) > 0$ and $\partial_{ww}^2 \Psi(ve^{-z}, s) > 0$ (see [11, Section V]). This implies that the function $\Psi(ve^{-z}, s)$ is strictly convex in z when $v \in (0,1), s \in (0,+\infty)$.

3 Models with neurotransmitter populations

To account for different realizations of the effective receptions of different neurotransmitters [21] over the same link, we introduce the following generalized model:

$$\begin{aligned} X_i(k+1) = &\Big(1 - \prod_{j \in E_i^{\circ k}} \prod_{\ell=1}^{a_{ij}^k} (1 - W_{ij(\ell)}^k X_j(k))\Big) \\ &\times \prod_{j \in I_i^{\circ k}} \prod_{\ell=1}^{a_{ij}^k} (1 - W_{ij(\ell)}^k X_j(k)) \end{aligned} \tag{19}$$

where a_{ij}^k denotes the number of neurotransmitters sent from neuron j to neuron i at step k, and $W_{ij(\ell)}^k$ is a binary variable that represents the successful reception of the ℓ^{th} neurotransmitter at step k from neuron j to neuron i when taking 1, otherwise 0. In this way, the successful reception of a neurotransmitter represented by a binary random variable $W_{ij(\ell)}^k$ is realized at each transmission ℓ at step k from neuron j to neuron i.

We introduce the following assumptions regarding independence of neurotransmissions to identify an approximate model for the probability of excitation.

(A5) (Memoryless Neurotransmission) For any $k > 0$, $W^k \triangleq [W_{ij(\ell)}^k] \in \mathrm{R}^{n \times n \times a_{ij}^k}$ is independent of $\{W^t : 0 \leq t < k\}$ and $\{X(t) : 0 \leq t < k\}$.

(A6) (Neurotransmission Conditional Independence) For any $k \geq 0$, the binary random variables $\{W^k_{ij(\ell)} : i, j \in [n], \ell \in [a^k_{ij}]\}$ representing the transmissions are jointly conditionally independent given the current states $X(k)$.

Following similar steps of the previous section, under assumptions (A5) and (A6), we obtain

$$\begin{aligned}\mathbb{E}(X_i(k+1)|(X(t))_{t\in[k]}) &= \mathbb{E}(X_i(k+1)|X(k)) \\ &= \Big(1 - \prod_{j\in E_i^{\circ k}} \prod_{\ell=1}^{a^k_{ij}} \mathbb{E}(1 - W^k_{ij(\ell)} X_j(k)|X(k))\Big) \\ &\quad \times \prod_{j\in I_i^{\circ k}} \prod_{\ell=1}^{a^k_{ij}} \mathbb{E}(1 - W^k_{ij(\ell)} X_j(k)|X(k)).\end{aligned} \tag{20}$$

Proposition 3. *Assume (A5) and (A6) hold. Given a state configuration $x \in \{0,1\}^n$ at step k, the transition probability to a state configuration $q \in \{0,1\}^n$ is given by*

$$\begin{aligned}&\Pr(X(k+1) = q|X(k) = x) \\ &\quad = \prod_{i=1}^{n} \Big(q_i\rho_i(k+1) + (1-q_i)(1-\rho_i(k+1))\Big)\end{aligned} \tag{21}$$

where

$$\begin{aligned}\rho_i(k+1) = \Big(1 - \prod_{j\in E_i^{\circ k}} (1 - w^k_{ij}x_j)^{a^k_{ij}}\Big) \\ \times \prod_{j\in I_i^{\circ k}} (1 - w^k_{ij}x_j)^{a^k_{ij}}, \quad i \in [n].\end{aligned} \tag{22}$$

The proof is similar to that of Prop. 1 and hence omitted.

Remark 8. Compared to Proposition 1, the difference lies in the representation of the probability $\rho_i(k+1)$ in (22) which now involves the number of neurotransmitters a^k_{ij} with $i, j \in [n]$ and $k \geq 0$.

Such results that evaluate transition probabilities are needed for computing optimal control solutions under the framework of Markov decision processes [9, 10].

We introduce the following assumptions to further simplify the representation.

(A7) (Neurotransmission Independence at Step k) At step $k \geq 0$, $\{W^k_{ij(\ell)} : i, j \in [n], \ell \in [a^k_{ij}]\}$ are independent, and for each $i, j \in [n]$ and each $\ell \in [a^k_{ij}]$, $W^k_{ij(\ell)}$ is independent of $\{X_q(k) : q \in [n], q \neq j\}$.

(A8) (State Independence among Neurons) Up to some terminal step T, for each $k \leq T$, the underlying binary random variables $\{X_i(k) : i \in [n]\}$ are independent.

Under (A5), (A7) and (A8), taking the expectation on both sides of the Equation (19) above yields

$$\begin{aligned}\mathbb{E}X_i(k+1) &= \mathbb{E}\Big(1-\prod_{j\in E_i^{\circ k}}\prod_{\ell=1}^{a_{ij}^k}(1-W_{ij(\ell)}^k X_j(k))\Big)\\ &\quad\times\mathbb{E}\prod_{j\in I_i^{\circ k}}\prod_{\ell=1}^{a_{ij}^k}(1-W_{ij(\ell)}^k X_j(k)))\\ &= \Big(1-\prod_{j\in E_i^{\circ k}}\mathbb{E}\prod_{\ell=1}^{a_{ij}^k}(1-W_{ij(\ell)}^k X_j(k))\Big)\\ &\quad\times\prod_{j\in I_i^{\circ k}}\mathbb{E}\prod_{\ell=1}^{a_{ij}^k}(1-W_{ij(\ell)}^k X_j(k))).\end{aligned} \tag{23}$$

Denote $p_i(k)\triangleq \Pr(X_i(k)=1)$. Let $w_{ij}^k\triangleq \Pr(W_{ij(\ell)}^k=1|X_j(k)=1)$ denote the conditional probability of the successful reception of each neurotransmitter from node j to node i at step k. We note that

$$\begin{aligned}&\mathbb{E}\prod_{\ell=1}^{a_{ij}^k}(1-W_{ij(\ell)}^k X_j(k)))\\ &=\Pr\Big(\prod_{\ell=1}^{a_{ij}^k}(1-W_{ij(\ell)}^k X_j(k)))=1\Big)\\ &=\Pr(\prod_{\ell=1}^{a_{ij}^k}(1-W_{ij(\ell)}^k X_j(k)))=1|X_j(k)=1)p_j(k)\\ &\quad+\Pr(\prod_{\ell=1}^{a_{ij}^k}(1-W_{ij(\ell)}^k X_j(k)))=1|X_j(k)=0)(1-p_j(k))\\ &=(1-w_{ij}^k)^{a_{ij}^k}p_j(k)+(1-p_j(k))=1-\omega_{ij}^{k,\text{eff}}p_j(k)\end{aligned} \tag{24}$$

where

$$\omega_{ij}^{k,\text{eff}}\triangleq 1-(1-w_{ij}^k)^{a_{ij}^k}.$$

Therefore, the Equation (23) is equivalent to

$$\begin{aligned}p_i(k+1) &= \Big(1-\prod_{j\in E_i^{\circ k}}(1-\omega_{ij}^{k,\text{eff}}p_j(k))\Big)\\ &\quad\times\prod_{j\in I_i^{\circ k}}(1-\omega_{ij}^{k,\text{eff}}p_j(k))).\end{aligned} \tag{25}$$

Hence we obtain the following result.

Proposition 4. *Assume that (A5), (A7) and (A8) hold. Consider $X_i(k)$ the stochastic binary state of the dynamics (19) and $p_i(k)$ the probability state of the dynamics (25). If $\Pr(X_i(0)=1)=p_i(0)$ for all $i\in[n]$, then*

$$\Pr(X_i(k)=1)=p_i(k),\quad \forall i\in[n],\ 0\le k\le T.$$

In what follows, we simplify the dynamics in (25) using the same state transformation in the previous section. With a slight abuse of notation, define

$$\pi_i(k)\triangleq\prod_{j\in I_i^{\circ k-1}}(1-\omega_{ij}^{k-1,\text{eff}}p_j(k-1))^{a_{ij}^{k-1}}\in[0,1] \tag{26}$$

$$s_i(k) \triangleq \begin{cases} -\log\left(1-\frac{p_i(k)}{\pi_i(k)}\right), & \text{if } \pi_i(k)\in(0,1] \\ 0, & \text{if } \pi_i(k)=0 \end{cases} \tag{27}$$

$$o_i(k) \triangleq -\log \pi_i(k) \in [0,+\infty] \tag{28}$$

with $o_i(k), s_i(k) \in [0,+\infty]$. Following the same analysis in the previous section, we obtain the following dynamics

$$s_i(k+1) = \sum_{j\in E_i^{\circ k}} \Psi(\omega_{ij}^{k,\text{eff}} e^{-o_j(k)}, s_j(k)), \quad i\in[n] \tag{29}$$

$$o_i(k+1) = \sum_{j\in I_i^{\circ k}} \Psi(\omega_{ij}^{k,\text{eff}} e^{-o_j(k)}, s_j(k)), \quad i\in[n] \tag{30}$$

where $\Psi(\cdot,\cdot)$ is TLogSigmoid function in (16) and $\omega_{ij}^{k,\text{eff}} \triangleq 1-(1-w_{ij}^k)^{a_{ij}^k}$ explicitly include the numbers of neurotransmitters $\{a_{ij}^k : i,j\in[n], k\geq 0\}$ into the evolution dynamics.

The initial conditions can be given by $o_i(0)=0$ (if no inhibition exists before the starting time), and $s_i(0) = -\log(1-p_i(0))$, for all $i\in[n]$.

The pair of Equations (17) and (18) can now be viewed as special cases of the pair of Equations (29) and (30) by setting $a_{ij}^k = 1$ for all $i,j\in[n]$ and $k\geq 0$.

Remark 9 (Biological Interpretation). An important feature of the model above is that the connection weights a_{ij}^k and w_{ij}^k are non-negative and admit direct biological interpretations (i.e. the number of neurotransmitters and the conditional probability of the successful reception of each neurotransmitter, respectively), in contrast to conventional neural networks that allow negative weights [3, 24].

Remark 10 (Relation with Graph Neural Networks). Such dynamics in (29) and (30) are related to graph neural networks [25], in particular the message-passing graph neural networks [14, Chp. 5], in which each node has a continuous state (or nodal feature) of dimension two. The state s_i (resp. o_i) summarizes the incoming influence of excitatory (resp. inhibitory) connections.

Remark 11 (Offset). One option is to consider self-excitatory neurons. A finite offset in the dynamics (29) and (30) can be created by introducing one node $n_0\in[n]$ that has (i) an excitatory autapse with $w_{n_0n_0}^k = 1$ for all $k\geq 0$, (ii) no incoming links, (iii) the outgoing transmission probability $w_{in_0}^k < 1$, for $i\in[n]\backslash n_0$ and $k\geq 0$, and (iv) initial state $s_{n^0}(0)=\infty$. Another option is to consider neurons activated by additional inputs. For instance, for a neuron $n_0\in[n]$ with outgoing transmission probability $w_{in_0}^k<1$ for all $i\in[n]\backslash n_0$, setting its state to $s_{n^0}(k)=\infty$ by an additional input will create a finite offset at step $k\geq 0$ for neurons in the outgoing neighborhood of n_0.

To obtain upper bounds for the states of (29) and (30), we introduce the following notation

$$\Omega^k \triangleq [\omega_{ij}^{k,\text{eff}}] \in \mathrm{R}^{n\times n}, \text{ with } \omega_{ij}^{k,\text{eff}} \triangleq 1-(1-w_{ij}^k)^{a_{ij}^k}.$$

Proposition 5 (Upper Bound). *Let the initial states be given by $s(0)=[s_1(0),\cdots,s_n(0)]^\intercal$ with $s_i(0) = -\log(1-p_i(0))$ and $o(0)=[o_1(0),\cdots,o_n(0)]^\intercal$. Then the states of (29) and (30) for $k\geq 1$ satisfy that for all $i\in[n]$,*

$$s_i(k) \leq [\mathcal{T}_E(k,0)s(0)]_i \tag{31}$$

$$o_i(k) \leq [(B_I^{k-1}\odot\Omega^{k-1})\mathcal{T}_E(k-1,0)s(0)]_i \tag{32}$$

with $\mathcal{T}_E(k,0) \triangleq (B_E^{k-1}\odot\Omega^{k-1})\cdots(B_E^0\odot\Omega^0)$.

Proof. By the concavity of $\Psi(w,x) \triangleq -\log(1-w+we^{-x})$ in x (see [11, Sec. V]), we have for any $z,z^*\in[-\infty,+\infty]$,

$$\Psi(w,z) \leq \Psi(w,z^*) + \partial_x\Psi(w,z^*)(z-z^*), \quad w\in[0,1].$$

In particular, taking $z^* = 0$ yields $\partial_x \Psi(w, 0) = w$ and hence $\Psi(w, z) \leq wz$. Applying this inequality to (29) yields

$$s_i(k+1) \leq \sum_{j \in E_i^{\circ k}} \omega_{ij}^{k,\text{eff}} e^{-o_j(k)} s_j(k) \leq \sum_{j \in E_i^{\circ k}} \omega_{ij}^{k,\text{eff}} s_j(k)$$

for $o_j(k) \in [0, +\infty]$. That is, the state s is element-wise upper bounded by the state of the discrete-time linear system

$$z(k+1) = (B_E^k \odot \Omega^k) z(k), \quad z(0) = s(0), \quad z(k) \in \mathrm{R}^n,$$

the solution of which is given by $z(k) = \mathcal{T}_E(k, 0)s(0)$. Thus, for all $k \geq 0$, $s_i(k) \leq z_i(k) = [\mathcal{T}_E(k, 0)s(0)]_i$. In addition, from the dynamics (30) of the state o, we obtain similarly

$$o_i(k+1) \leq \sum_{j \in I_i^{\circ k}} \omega_{ij}^{k,\text{eff}} s_j(k), \quad \forall i \in [n]. \tag{33}$$

Therefore,

$$o_i(k+1) \leq [(B_I^k \odot \Omega^k) s(k)]_i \leq [(B_I^k \odot \Omega^k) z(k)]_i.$$

Replacing $z(k)$ by its solution $\mathcal{T}_E(k, 0)s(0)$ and shifting the time index from $k+1$ to k yield the desired result. □

Proposition 6 (Stability). *Assume that $B_E^k = B_E$, $B_I^k = B_I$ and $\Omega^k = \Omega$ are invariant with respect to the step $k \geq 0$. Then the system (29) and (30) is asymptotically and exponentially stable with respect to k at the origin if*

$$\max_{i \in [n]} |\lambda_i(B_E \odot \Omega)| < 1$$

with $\lambda_i(B_E \odot \Omega), i \in [n]$ as the i^{th} eigenvalue of $B_E \odot \Omega$.

Proof. Prop. 5 together with the condition for the stability of discrete-time linear systems implies the desired result. □

4 Limit model with infinite neurotransmitters

Since the number of released neurotransmitters could be very large (see e.g. [21, Part III, Chp. 11]), the number of receptors at the post-synaptic neurons are relatively moderate and are assumed not to change over a short period of time, the probability of a successful transmission w_{ij} of each neurotransmitter from node j to node i may decrease with respect to the number of neurotransmitters a_{ij}. Hence we introduce the following assumption.

(A9) The probability of transmission w_{ij}^k depends on the number of neurotransmitters a_{ij}^k as follows:

$$w_{ij}^k = {a_{ij}^k}^{-1} \lambda_{ij}^k, \text{ with } \lambda_{ij}^k \text{ fixed } \forall k \geq 0, \ \forall i, j \in [n].$$

Based on (24), we obtain that under (A9),

$$\begin{aligned}
&\lim_{a_{ij}^k \to \infty} \mathbb{E} \prod_{q=1}^{a_{ij}^k} \left(1 - W_{ij(q)}^k X_j(k)\right) \\
&\quad = \lim_{a_{ij}^k \to \infty} [1 - (1 - (1 - w_{ij}^k)^{a_{ij}^k}) p_j(k)] \\
&\quad = \lim_{a_{ij}^k \to \infty} \left[1 - \left(1 - \left(1 - {a_{ij}^k}^{-1} \lambda_{ij}^k\right)^{a_{ij}^k}\right) p_j(k)\right] \\
&\quad = 1 - (1 - e^{-\lambda_{ij}^k}) p_j(k) = 1 - \lambda_{ij}^{k,\text{eff}} p_j(k)
\end{aligned} \tag{34}$$

with $\lambda_{ij}^{k,\text{eff}} \triangleq 1 - e^{-\lambda_{ij}^k}$. Therefore, under (A5), (A7), (A8) and (A9), the expected state in (19) then satisfies

$$\begin{aligned}\mathbb{E}X_i(k+1) = \lim_{a_{ij}^k \to \infty} \Big[\Big(1 - \prod_{j \in E_i^{\circ k}} \mathbb{E} \prod_{q=1}^{a_{ij}^k} \big(1 - W_{ij(q)}^k X_j(k)\big)\Big) \\ \times \prod_{j \in I_i^{\circ k}} \mathbb{E} \prod_{\ell=1}^{a_{ij}^k} (1 - W_{ij(\ell)}^k X_j(k))\Big].\end{aligned} \tag{35}$$

That is,

$$\begin{aligned}p_i(k+1) = \Big(1 - \prod_{j \in E_i^{\circ k}} (1 - \lambda_{ij}^{k,\text{eff}} p_j(k))\Big) \\ \times \prod_{j \in I_i^{\circ k}} (1 - \lambda_{ij}^{k,\text{eff}} p_j(k)),\end{aligned} \tag{36}$$

for all $i \in [n]$, with $\lambda_{ij}^{k,\text{eff}} \triangleq 1 - e^{-\lambda_{ij}^k}$ where λ_{ij}^k is the rate of the Poisson distribution at time k for the synaptic connection from neuron j to neuron i.

To simplify the representation of the dynamics, let

$$\bar{\pi}_i(k) \triangleq \prod_{j \in I_i^{\circ k-1}} (1 - \lambda_{ij}^{k-1,\text{eff}} p_j(k-1)) \in (0, 1] \tag{37}$$

represent the probability of no inhibition at node i from its neighbouring neurons in the previous step. Let

$$\bar{s}_i(k) \triangleq -\log\left(1 - \frac{p_i(k)}{\bar{\pi}_i(k)}\right) \in [0, +\infty] \tag{38}$$

$$\bar{o}_i(k) \triangleq -\log \bar{\pi}_i(k) \in [0, +\infty). \tag{39}$$

From (36), we obtain the dynamics for $(\bar{s}, \bar{o})$, given by

$$\bar{s}_i(k+1) = \sum_{j \in E_i^{\circ k}} \Psi(\lambda_{ij}^{k,\text{eff}} e^{-\bar{o}_j(k)}, \bar{s}_j(k)) \tag{40}$$

$$\bar{o}_i(k+1) = \sum_{j \in I_i^{\circ k}} \Psi(\lambda_{ij}^{k,\text{eff}} e^{-\bar{o}_j(k)}, \bar{s}_j(k)) \tag{41}$$

where $\lambda_{ij}^{k,\text{eff}} \triangleq 1 - e^{-\lambda_{ij}^k}$ and $\Psi(\cdot, \cdot)$ is TLogSigmoid function in (16). The initial conditions are given by $\bar{o}_i(0) = 0$, representing no inhibitions before the first step, and $\bar{s}_i(0) = -\log(1 - p_i(0))$, for all $i \in [n]$.

Remark 12. We note that $\bar{\pi}_i(k)$ cannot be zero following its definition in (37), since λ_{ij}^k is assumed to be finite, and $p_j(k-1) \in [0, 1]$. Therefore, we excluded $+\infty$ in (39).

Remark 13 (Poisson Approximation). Consider n independent Bernoulli random variables $Z_1, \cdots, Z_n$, each of which takes 1 with probability $\frac{\lambda}{n}$ for some finite $\lambda \geq 0$. Then the sum $S_n = \sum_{i=1}^n Z_i$ follows a binomial distribution Binomial$(n, \frac{\lambda}{n})$, which has mean λ and variance $\lambda(1 - \frac{\lambda}{n})$. As $n \to \infty$, the Binomial distribution Binomial$(n, \frac{\lambda}{n})$ converges Poisson distribution with rate λ by Poisson limit theorem, and hence

$$\lim_{n \to \infty} \Pr\left(\prod_{i=1}^n (1 - Z_i) = 1\right) = \Pr(S_n = 0) = e^{-\lambda}.$$

Applying the Poisson limit gives rise to

$$\begin{aligned}\lim_{a_{ij}^k \to \infty} \Pr\Big[\prod_{q=1}^{a_{ij}^k} \Big(1 - W_{ij(q)}^k X_j(k)\Big) = 1\Big] \\ = e^{-\lambda_{ij}^k} p_j(k) + (1 - p_j(k))\end{aligned}$$

which is the same limit as that we established in (34). The last equality uses the conditional independence of the random variables $W^k_{ij(q)}X_j(k)$ with $q \in [a^k_{ij}]$ given $X_j(k)$.

Heuristically, this limit model approximates the finite neurotransmitter model well when a^k_{ij} is large, w^k_{ij} is small, and $\lambda^k_{ij} = a^k_{ij}w^k_{ij}$ is moderate.

The only difference between the finite model in (29) and (30) and the limit model in (40) and (41) is the parameters representing transmission probabilities. Hence results in Prop. 5 and Prop. 6 apply to the limit model by simply replacing the parameter matrix $\Omega^k \triangleq [\omega^{k,\text{eff}}_{ij}] \in \mathrm{R}^{n\times n}$ there by $\Lambda^k \triangleq [\lambda^k_{ij}] \in \mathrm{R}^{n\times n}$.

Corollary 1 (Upper Bound). *Let the initial states be given by* $\bar{s}(0) = [\bar{s}_1(0), \cdots, \bar{s}_n(0)]^\intercal$ *with* $\bar{s}_i(0) = -\log(1 - p_i(0))$ *and* $\bar{o}(0) = [\bar{o}_1(0), \cdots, \bar{o}_n(0)]^\intercal$. *Then the states of the system (40) and (41) for any step* $k \geq 1$ *satisfy that for all* $i \in [n]$,

$$\bar{s}_i(k) \leq [\Gamma_E(k, 0)\bar{s}(0)]_i \tag{42}$$

$$\bar{o}_i(k) \leq [(B^{k-1}_I \odot \Lambda^{k-1})\Gamma_E(k-1, 0)\bar{s}(0)]_i \tag{43}$$

with $\Gamma_E(k, 0) \triangleq (B^{k-1}_E \odot \Lambda^{k-1}) \cdots (B^0_E \odot \Lambda^0)$.

Corollary 2 (Stability). *Assume* $B^k_E = B_E$, $B^k_I = B_I$ *and* $\Lambda^k = \Lambda$ *are invariant with respect to the step* $k \geq 0$. *Then the system (40) and (41) is asymptotically and exponentially stable with respect to the step* k *at the origin if*

$$\max_{i\in[n]} |\lambda_i(B_E \odot \Lambda)| < 1$$

with $\lambda_i(B_E \odot \Lambda), i \in [n]$ *as the* i^{th} *eigenvalue of* $B_E \odot \Lambda$.

5 Conclusion

We have generalized the TransNN model to include inhibitions and shown that under technical assumptions, the probability of neuronal excitations in TransNNs with both inhibitory and excitatory connections can be equivalently represented by neural networks where each neuron has a two-dimensional continuous state and each link has the TLogSigmoid activation function. Moreover, neurotransmitter populations were considered in an extended model, and Poisson approximation was applied to establish limit models when the number of neurotransmitters at each link goes to infinity. Sufficient conditions for stability of proposed models have been established.

Future work should investigate non-trivial equilibria, the integration of the neuronal dynamics for action potentials and spiking sequences, the control of such networks, and the game-theoretic analysis of such dynamics with costs such as energy and resources for firing.

Use of artificial intelligence

ChatGPT and Claude Code were used to review the manuscript and correct an error in (24) in a previous version.

References

[1] David H Ackley, Geoffrey E Hinton, and Terrence J Sejnowski. A learning algorithm for boltzmann machines. Cognitive Science, 9(1):147–169, 1985.

[2] John M Bekkers. Synaptic transmission: functional autapses in the cortex. Current Biology, 13(11):R433–R435, 2003.

[3] Hans-Dieter Block. The perceptron: A model for brain functioning. i. Reviews of Modern Physics, 34(1):123, 1962.

[4] Francesco Bullo. Lectures on Neural Dynamics. 2025.

[5] Marie Cottrell. Mathematical analysis of a neural network with inhibitory coupling. Stochastic Processes and Their Applications, 40(1):103–126, 1992.

[6] Matthieu Courbariaux, Yoshua Bengio, and Jean-Pierre David. Binaryconnect: Training deep neural networks with binary weights during propagations. Advances in Neural Information Processing Systems, 28, 2015.

[7] John C Eccles, P Fatt, and K Koketsu. Cholinergic and inhibitory synapses in a pathway from motor-axon collaterals to motoneurones. The Journal of Physiology, 126(3):524, 1954.

[8] Joaquin M Fuster and Garrett E Alexander. Neuron activity related to short-term memory. Science, 173(3997):652–654, 1971.

[9] Shuang Gao and Peter E Caines. Transmission neural networks: Approximate receding horizon control for virus spread on networks. In Proceedings of IEEE Conference on Decision and Control (CDC), pages 6208–6215, 2025.

[10] Shuang Gao and Peter E. Caines. Transmission neural networks: Approximation and optimal control. IFAC-PapersOnLine, 59(4):31–36, 2025. 10th IFAC Conference on Networked Systems (NecSys).

[11] Shuang Gao and Peter E Caines. Transmission neural networks: From virus spread models to neural networks. Les Cahiers du GERAD, 2026. arXiv:2208.03616.

[12] Wulfram Gerstner, Werner M Kistler, Richard Naud, and Liam Paninski. Neuronal dynamics: From single neurons to networks and models of cognition. Cambridge University Press, 2014.

[13] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep learning. MIT press, 2016.

[14] William L Hamilton. Graph representation learning. Synthesis Lectures on Artifical Intelligence and Machine Learning, 14(3):1–159, 2020.

[15] H K Hartline, Henry G Wagner, and Floyd Ratliff. Inhibition in the eye of limulus. The Journal of General Physiology, 39(5):651–673, 1956.

[16] Geoffrey E Hinton and Terrence J Sejnowski. Optimal perceptual inference. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, volume 448, pages 448–453. Washington, 1983.

[17] 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.

[18] John J Hopfield. Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences, 79(8):2554–2558, 1982.

[19] John J Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. Proceedings of the National Academy of Sciences, 81(10):3088–3092, 1984.

[20] Itay Hubara, Matthieu Courbariaux, Daniel Soudry, Ran El-Yaniv, and Yoshua Bengio. Binarized neural networks. Advances in Neural Information Processing Systems, 29, 2016.

[21] Eric R. Kandel, John D. Koester, Sarah H. Mack, and Steven A. Siegelbaum. Principles of Neural Science. McGraw-Hill Education, New York, 6 edition, 2021.

[22] Eric R Kandel, James H Schwartz, Thomas M Jessell, Steven Siegelbaum, A James Hudspeth, Sarah Mack, et al. Principles of Neural Science, volume 4. McGraw-hill New York, 2000.

[23] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. Nature, 521(7553):436–444, 2015.

[24] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. Nature, 323(6088):533–536, 1986.

[25] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. The graph neural network model. IEEE Transactions on Neural Networks, 20(1):61–80, 2009.

[26] Philip K Shiu, Gabriella R Sterne, Nico Spiller, Romain Franconville, Andrea Sandoval, Joie Zhou, Neha Simha, Chan Hyuk Kang, Seongbong Yu, Jinseop S Kim, et al. A drosophila computational brain model reveals sensorimotor processing. Nature, 634(8032):210–219, 2024.

[27] Tatyana Turova. Stochastic dynamics of a neural network with inhibitory and excitatory connections. BioSystems, 40(1-2):197–202, 1997.

[28] Hugh R Wilson and Jack D Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. Biophysical Journal, 12(1):1–24, 1972.

[29] Hugh R Wilson and Jack D Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik, 13(2):55–80, 1973.