2021

p-ADIC CELLULAR NEURAL NETWORKS

B. A. Zambrano-Luna

Wilson A. Zuniga-Galindo
The University of Texas Rio Grande Valley

p-ADIC CELLULAR NEURAL NETWORKS

B. A. ZAMBRANO-LUNA AND W. A. ZÚÑIGA-GALINDO

Abstract. In this article we introduce the p-adic cellular neural networks which are mathematical generalizations of the classical cellular neural networks (CNNs) introduced by Chua and Yang. The new networks have infinitely many cells which are organized hierarchically in rooted trees, and also they have infinitely many hidden layers. Intuitively, the p-adic CNNs occur as limits of large hierarchical discrete CNNs. More precisely, the new networks can be very well approximated by hierarchical discrete CNNs. Mathematically speaking, each of the new networks is modeled by one integro-differential equation depending on several p-adic spatial variables and the time. We study the Cauchy problem associated to these integro-differential equations and also provide numerical methods for solving them.

1. Introduction

In the late 80s Chua and Yang introduced a new natural computing paradigm called the cellular neural networks (or cellular nonlinear networks) CNN which includes the cellular automata as a particular case [6], [7], [9]. From the beginning the CNN paradigm was intended for applications as an integrated circuit. This paradigm has been extremely successful in various applications in vision, robotics and remote sensing, see e.g. [8], [25] and the references therein.

In this article we present a mathematical generalization of the CNNs of Chua and Yang called *p-adic cellular neural networks*. The p-adic continuous CNNs offer a theoretical framework to study the emergent patterns of hierarchical discrete CNNs having arbitrary many hidden layers.

Nowadays, it is widely accepted that the analysis on ultrametric spaces is the natural tool for formulating models where the hierarchy plays a central role. An ultrametric space (M,d) is a metric space M with a distance satisfying $d(A,B) \leq \max\{d(A,C), d(B,C)\}$ for any three points A, B, C in M. Ultrametricity in physics means the emergence of ultrametric spaces in physical models. Ultrametricity was discovered in the 80s by Parisi and others in the theory of spin glasses and by Frauenfelder and others in physics of proteins. In both cases, the space of states of a complex system has a hierarchical structure which play a central role in the physical behavior of the system, see e.g. [10], [11], [15], [20], [21], [24], [27], [28]-[30], and the references therein.

On the other hand, Khrennikov and his collaborators have studied neural network models where p-state neurons take their values in p-adic numbers, see [2],

Key words and phrases. Cellular neural networks, Hierarchies, Deep learning, p-Adic numbers.
The second author was partially supported by Conacyt Grant No. 250845 (Mexico) and by the Debnath Endowed Professorship (UTRGV, USA).

[19]. These models are completely different to the ones considered here. In addition, Khrennikov has developed non-Archimedean models of brain activity and mental processes, see e.g. [18] and the references therein.

Among the ultrametric spaces, the field of p-adic numbers $\mathbb{Q}_p$ plays a central role. A p-adic number is a series of the form

$$x = x_{-k}p^{-k} + x_{-k+1}p^{-k+1} + \ldots + x_0 + x_1 p + \ldots, \text{ with } x_{-k} \neq 0, \tag{1.1}$$

where p is a prime number, the x_js are p-adic digits, i.e. numbers in the set $\{0, 1, \ldots, p-1\}$. The set of all the possible series of form (1.1) constitutes the field of p-adic numbers $\mathbb{Q}_p$. There are natural field operations, sum and multiplication, on series of form (1.1), see e.g. [16]. There is also a natural norm in $\mathbb{Q}_p$ defined as $|x|_p = p^k$, for a nonzero p-adic number x of the form (1.1). The field of p-adic numbers with the distance induced by $|\cdot|_p$ is a complete ultrametric space. The ultrametric property refers to the fact that $|x-y|_p \leq \max\left\{|x-z|_p, |z-y|_p\right\}$ for any x, y, z in $\mathbb{Q}_p$.

We denote by G_M the set all the p-adic numbers of the form $\boldsymbol{i} = \boldsymbol{i}_{-M}p^{-M} + \boldsymbol{i}_{-M+1}p^{-M+1} + \cdots + \boldsymbol{i}_0 + \cdots + \boldsymbol{i}_{M-1}p^{M-1}$, where the $\boldsymbol{i}_j$s belong to $\{0, 1, \ldots, p-1\}$. Then $(G_M, |\cdot|_p)$ is a finite ultrametric space. Geometrically speaking, G_M is a regular rooted tree with $2M$ layers, here regular means that exactly p edges emanate from each vertex. A (1-dimensional) p-adic discrete CNN is a dynamical system of the form

$$\frac{\partial}{\partial t} X(\boldsymbol{i}, t) = -X(\boldsymbol{i}, t) + \sum_{\boldsymbol{j} \in G_M} \mathbb{A}(\boldsymbol{i}, \boldsymbol{j}) Y(\boldsymbol{j}, t) + \sum_{\boldsymbol{j} \in G_M} \mathbb{B}(\boldsymbol{i}, \boldsymbol{j}) U(\boldsymbol{j}) + Z(\boldsymbol{i}), \tag{1.2}$$

$\boldsymbol{i} \in G_M$, where $Y(\boldsymbol{j}, t) = f(X(\boldsymbol{j}, t))$, with $f(x) = \frac{1}{2}(|x+1| - |x-1|)$. Here $X(\boldsymbol{i}, t), Y(\boldsymbol{i}, t) \in \mathbb{R}$ are the state, respectively the output, of cell $\boldsymbol{i}$ at the time t. The function $U(\boldsymbol{i}) \in \mathbb{R}$ is the input of the cell $\boldsymbol{i}$, $Z(\boldsymbol{i}) \in \mathbb{R}$ is the threshold of cell $\boldsymbol{i}$, and the matrices $\mathbb{A}, \mathbb{B} : G_M^N \times G_M^N \to \mathbb{R}$ are the feedback operator and feedforward operator, respectively. Notice that matrices $\mathbb{A}$, $\mathbb{B}$ are functions on the Cartesian product of two rooted trees. The Chua-Yang CNNs are a particular case of (1.2). In this article we study N-dimensional, discrete hierarchical CNNs having arbitrary many layers. For the seek of simplicity, we focus on space-invariant networks, i.e. in the case in which

$$\mathbb{A}(\boldsymbol{i}, \boldsymbol{j}) = \mathbb{A}(|\boldsymbol{i} - \boldsymbol{j}|_p), \quad \mathbb{B}(\boldsymbol{i}, \boldsymbol{j}) = \mathbb{B}(|\boldsymbol{i} - \boldsymbol{j}|_p). \tag{1.3}$$

In this article we initiate the study of the emergent patterns produced by the p-adic discrete CNNs. Since we are interested in arbitrary large trees, the description of these networks requires literally of millions of integro-differential equations, consequently a numerical approach seems not suitable, instead of this, we construct a p-adic continuous model that can be very well approximated by (1.2).

The study of the qualitative behavior of differential equations on large graphs is a relevant matter due its applications. In [23] Nakao and Mikhailov proposed using continuous models to study reaction-diffusion systems on networks and the corresponding Turing patterns. In [28] the second author showed that p-adic analysis is the natural tool to carry out this program. Models constructed using energy landscapes naturally drive to a large systems of differential equations (the master equation of the system), see e.g. [5], [20], [21]. p-Adic continuous versions of some of these systems were constructed by Avetisov, Kozyrev and others in connection with models of protein folding, see e.g. [20], [21] for a general discussion. Another relevant system is the Eigen-Schuster model in biology. In [29] a p-adic continuous version of this model was introduced, this p-adic version allows to explain the Eigen paradox. Recently Hua and Hovestadt pointed out that the p-adic number system offers a natural representation of hierarchical organization of complex networks [15].

Intuitively, in the space-invariant case, the continuous model corresponding to (1.2) is obtained by taking the limit as M tends to infinity:

$$\frac{\partial X(x,t)}{\partial t}=-X(x,t)+\int\limits_{\mathbb{Q}_p}A(|x-y|_p)Y(y,t)dy+ \int\limits_{\mathbb{Q}_p}B(|x-y|_p)U(y)dy+Z(x), \tag{1.4}$$

with $Y(x,t)=f(X(x,t))$. For the sake of simplicity, in the introduction we discuss our results in dimension one. We study the case where $A(|x|_p)$, $B(|x|_p)$ are integrable, and U, Z are continuous functions vanishing at infinity. Under these hypotheses the initial value problem attached to (1.4), with initial datum X_0 (a continuous function vanishing at infinity) has a unique solution $X(x,t)$ which is a continuos function vanishing at infinity in x for any $t\geq 0$, satisfying $|X(x,t)|\leq X_{\max}$, where the constant $X_{\max}$ is completely determined by A, B, U, Z and f, see Theorem 2. An analog result is valid for discrete CNNs, see Theorem 3.

The solution $X(x,t)$ can be very well approximated in the $\|\cdot\|_\infty$-norm as

$$\sum_{\boldsymbol{j}\in G_M} X(\boldsymbol{j},t)\Omega\left(p^M|x-\boldsymbol{j}|_p\right).$$

By using standard techniques of approximation of semilinear evolution equations, we show that the solution of the Cauchy problem attached to (1.2), under condition (1.3), is arbitrarily closed in the $\|\cdot\|_\infty$-norm to the solution of the Cauchy problem attached to (1.4), if M is sufficiently large, see Theorem 4. This implies that the p-adic continuous CNNs have infinitely many hidden layers, and that they are continuous versions of suitable p-adic discrete CNNs. It is relevant to mention that equation (1.4) makes sense over the real numbers, i.e. by replacing $\mathbb{Q}_p$ by $\mathbb{R}$ in (1.4) we get an equation modeling a continuous network. But, there are no natural discretizations of the real version of (1.4) that can be interpreted as hierarchical CNNs, because the real numbers are a completely ordered field, and thus the natural hierarchy is only the linear one.

In practical applications it is natural to assume that radial functions A, B have compact support or that they are test functions. Under this hypothesis we study the patterns produced by p-adic continuous CNNs when U, Z and X_0 are test functions. The hypothesis that X_0 is a test functions means that at time $t=0$ only certain clusters of cells are excited. Each cluster corresponds to a p-adic ball centered at some cell with radius, say p^{-L}. The intensity of the excitation is the same for all cells in a given cluster. The fact that U, Z are test functions can be interpreted in an analogous way. Let B_{M_0} denote the ball centered at the origin with radius p^{M_0}, which the smallest ball containing the supports of A, B, U, Z, X_0. Then the solution $X(x,t)$ of the initial value problem attached to (1.4) is a test function supported in B_{M_0} of the form $\sum_{\boldsymbol{j}\in G_{M_0}} X(\boldsymbol{j},\boldsymbol{t})\Omega\left(p^{M_0}|x-\boldsymbol{j}|_p\right)$ for $t\geq 0$, with $M_0\geq L$, see Theorem 1. This means that a p-adic continuous CNN produces a pattern which is organized in a finite number of disjoint clusters, each of them supporting a time varying pattern. We also show the existence of two steady state patterns $X_{+}(x)$, $X_{-}(x)$, which are test functions, such that $X_{-}(x) \leq \lim_{t\rightarrow\infty} X(x,t) \leq X_{+}(x)$, see Theorem 2. We conjecture that for generic p-adic continuous CNNs, $\lim_{t\rightarrow\infty} X(x,t)$ is a test function, which means that the steady state pattern is organized in a finite number of disjoint clusters, each of them supporting a constant pattern. This is exactly the multistability property reported in [23], see also [28], for reaction-diffusion networks.

We have conducted a large number of numerical simulations. Such simulations require solving integro-differential equations on a tree. The numerical study of p-adic continuous CNNs offers two big challenges. The first, the need of dealing with matrices having millions of entries, the second, the visualization of functions depending on p-adic variables. Due to the first problem, we use small trees with 16 to 64 leaves. The p-adic numbers have a fractal nature, then, it is necessary to visualize real-valuated functions defined on the Cartesian product of a fractal times the real line. To deal with this problem we us systematically heat maps which allow us to get a glimpse of the hierarchical nature of the CNNs. Our numerical simulations show that the solutions of continuous CNNs exhibit a very complex behavior, including self-similarity and multistability, depending on the interaction of all the parameters defining the network and initial datum.

2. p-Adic Analysis: Essential Ideas

2.1. **The field of p-adic numbers.** Along this article p will denote a prime number. The field of $p-$adic numbers $\mathbb{Q}_{p}$ is defined as the completion of the field of rational numbers $\mathbb{Q}$ with respect to the $p-$adic norm $|\cdot|_{p}$, which is defined as

$$|x|_{p}=\begin{cases} 0 & \text{if} \quad x=0\\ p^{-\gamma} & \text{if} \quad x=p^{\gamma}\frac{a}{b}, \end{cases}$$

where a and b are integers coprime with p. The integer $\gamma:=ord(x)$, with $ord(0):=+\infty$, is called the *$p-$adic order of x*. We extend the $p-$adic norm to $\mathbb{Q}_{p}^{N}$ by taking

$$||x||_{p}:=\max_{1\leq i\leq N}|x_{i}|_{p}, \qquad \text{for } x=(x_{1},\dots,x_{N})\in\mathbb{Q}_{p}^{N}.$$

We define $ord(x)=\min_{1\leq i\leq N}\{ord(x_{i})\}$, then $||x||_{p}=p^{-ord(x)}$. The metric space $\left(\mathbb{Q}_{p}^{N},||\cdot||_{p}\right)$ is a complete ultrametric space. As a topological space $\mathbb{Q}_{p}$ is homeomorphic to a Cantor-like subset of the real line, see e.g. [1], [27].

Any $p-$adic number $x\neq0$ has a unique expansion of the form

$$x=p^{ord(x)}\sum_{j=0}^{\infty}x_{j}p^{j},$$

where $x_{j}\in\{0,1,\dots,p-1\}$ and $x_{0}\neq0$.

2.2. **Topology of $\mathbb{Q}_{p}^{N}$.** For $r\in\mathbb{Z}$, denote by $B_{r}^{N}(a)=\{x\in\mathbb{Q}_{p}^{N};||x-a||_{p}\leq p^{r}\}$ *the ball of radius p^{r} with center at $a=(a_{1},\dots,a_{N})\in\mathbb{Q}_{p}^{N}$*, and take $B_{r}^{N}(0):=B_{r}^{N}$. Note that $B_{r}^{N}(a)=B_{r}(a_{1})\times\cdots\times B_{r}(a_{N})$, where $B_{r}(a_{i}):=B_{r}^{1}(a_{i})=\{x\in\mathbb{Q}_{p};|x_{i}-a_{i}|_{p}\leq p^{r}\}$ is the one-dimensional ball of radius p^{r} with center at $a_{i}\in\mathbb{Q}_{p}$. The ball B_{0}^{N} equals the product of N copies of $B_{0}=\mathbb{Z}_{p}$, *the ring of $p-$adic integers.* We also denote by $S_{r}^{N}(a)=\{x\in\mathbb{Q}_{p}^{N};||x-a||_{p}=p^{r}\}$ *the sphere of radius p^{r} with center at $a=(a_{1},\dots,a_{N})\in\mathbb{Q}_{p}^{N}$*, and take $S_{r}^{N}(0):=S_{r}^{N}$. We notice that $S_{0}^{1}=\mathbb{Z}_{p}^{\times}$

(the group of units of $\mathbb{Z}_p$), but $\left(\mathbb{Z}_p^{\times}\right)^N \subsetneq S_0^N$. The balls and spheres are both open and closed subsets in $\mathbb{Q}_p^N$. In addition, two balls in $\mathbb{Q}_p^N$ are either disjoint or one is contained in the other.

As a topological space $\left(\mathbb{Q}_p^N, ||\cdot||_p\right)$ is totally disconnected, i.e. the only connected subsets of $\mathbb{Q}_p^N$ are the empty set and the points. A subset of $\mathbb{Q}_p^N$ is compact if and only if it is closed and bounded in $\mathbb{Q}_p^N$, see e.g. [27, Section 1.3], or [1, Section 1.8]. The balls and spheres are compact subsets. Thus $\left(\mathbb{Q}_p^N, ||\cdot||_p\right)$ is a locally compact topological space.

We will use $\Omega\left(p^{-r}||x-a||_p\right)$ to denote the characteristic function of the ball $B_r^N(a)$. For more general sets, we will use the notation 1_A for the characteristic function of a set A.

2.3. **The Bruhat-Schwartz space.** A real-valued function φ defined on $\mathbb{Q}_p^N$ is *called locally constant* if for any $x \in \mathbb{Q}_p^N$ there exists an integer $l(x) \in \mathbb{Z}$ such that

$$\varphi(x+x') = \varphi(x) \text{ for } x' \in B_{l(x)}^N. \tag{2.1}$$

A function $\varphi : \mathbb{Q}_p^N \to \mathbb{R}$ is called a *Bruhat-Schwartz function (or a test function)* if it is locally constant with compact support. Any test function can be represented as a linear combination, with real coefficients, of characteristic functions of balls. The $\mathbb{R}$-vector space of Bruhat-Schwartz functions is denoted by $\mathcal{D}(\mathbb{Q}_p^N)$. For $\varphi \in \mathcal{D}(\mathbb{Q}_p^N)$, the largest number $l = l(\varphi)$ satisfying (2.1) is called *the exponent of local constancy (or the parameter of constancy) of* φ.

If U is an open subset of $\mathbb{Q}_p^N$, $\mathcal{D}(U)$ denotes the space of test functions with supports contained in U, then $\mathcal{D}(U)$ is dense in

$$L^{\rho}(U) = \left\{ \varphi : U \to \mathbb{R}; \left(\int_{\mathbb{Q}_p^N} |\varphi(x)|^{\rho}\, d^N x \right)^{\frac{1}{\rho}} < \infty \right\},$$

where $d^N x$ is the Haar measure on $\mathbb{Q}_p^N$ normalized by the condition $vol(B_0^N) = 1$, for $1 \leq \rho < \infty$, see e.g. [1, Section 4.3]. In the case $U = \mathbb{Q}_p^N$, we will use the notation L^{ρ} instead of $L^{\rho}\left(\mathbb{Q}_p^N\right)$. For an in depth discussion about p-adic analysis the reader may consult [1], [17], [26], [27].

2.4. **The Spaces $\mathcal{X}_{\infty}$, $\mathcal{X}_M$.** We define $\mathcal{X}_{\infty}(\mathbb{Q}_p^N) := \mathcal{X}_{\infty} = \overline{\left(\mathcal{D}(\mathbb{Q}_p^N), \|\cdot\|_{\infty}\right)}$, where $\|\phi\|_{\infty} = \sup_{x \in \mathbb{Q}_p^N} |\phi(x)|$ and the bar means the completion with respect the metric induced by $\|\cdot\|_{\infty}$. Notice that all the functions in $\mathcal{X}_{\infty}$ are continuous and that

$$\mathcal{X}_{\infty} \subset \mathcal{C}_0 := \left(\left\{ f : \mathbb{Q}_p^N \to \mathbb{R}; f \text{ continuous with } \lim_{\|x\|_p \to \infty} f(x) = 0 \right\}, \|\cdot\|_{\infty} \right).$$

On the other hand, since $\mathcal{D}(\mathbb{Q}_p^N)$ is dense in $\mathcal{C}_0$, cf. [26, Chap. II, Proposition 1.3], we conclude that $\mathcal{X}_{\infty} = \mathcal{C}_0$.

For $M \geq 1$, we set $G_M^N := B_M^N / B_{-M}^N$, which is a finite additive group with $\# G_M^N := p^{2NM}$ elements. Any element $\boldsymbol{i} = (\boldsymbol{i}_1, \ldots, \boldsymbol{i}_N)$ of G_M^N can be represented as

$$\boldsymbol{i}_j = \boldsymbol{i}_{-M}^j p^{-M} + \boldsymbol{i}_{-M+1}^j p^{-M+1} + \ldots + \boldsymbol{i}_0^j + \boldsymbol{i}_1^j p + \ldots + \boldsymbol{i}_{M-1}^j p^{M-1}, \tag{2.2}$$

for $j=1,\ldots,N$, with $\boldsymbol{i}_k^j \in \{0,1,\ldots,p-1\}$. From now on, we fix a set of representatives in $\mathbb{Q}_p^N$ for G_M^N of the form (2.2). Notice that

$$\boldsymbol{i}_j = p^{-M}\left(\boldsymbol{a}_0^j + \boldsymbol{a}_1^j p + \cdots + \boldsymbol{a}_{2M-1}^j p^{2M-1}\right),$$

where $\boldsymbol{a}_0^j + \boldsymbol{a}_1^j p + \cdots + \boldsymbol{a}_{2M-1}^j p^{2M-1} \in \mathbb{Z}_p/p^{2M}\mathbb{Z}_p = B_0/B_{-2M}$.

The functions

$$\left\{\Omega\left(p^M \|x-\boldsymbol{i}\|_p\right)\right\}_{\boldsymbol{i}\in G_M^N} \tag{2.3}$$

are orthogonal with respect to the standard L^2 inner product, since

$\Omega\left(p^M \|x-\boldsymbol{i}\|_p\right)\Omega\left(p^M \|x-\boldsymbol{j}\|_p\right) = 0$, for $\boldsymbol{i},\boldsymbol{j}\in G_M^N$, $\boldsymbol{i}\neq\boldsymbol{j}$ and for any $x\in B_M^N$.

We denote by $\mathcal{D}^M\left(\mathbb{Q}_p^N\right) := \mathcal{D}^M$ the $\mathbb{R}$-vector space spanned by (2.3). We set

$$\mathcal{X}_M := \left(\mathcal{D}^M, \|\cdot\|_\infty\right) \text{ for } M\geq 1.$$

Notice that $\mathcal{X}_M$ is isomorphic as a Banach space to $\left(\mathbb{R}^{\#G_M^N}, \|\cdot\|_{\mathbb{R}}\right)$, where

$$\left\|\left(t_1,\ldots,t_{\#G_M^N}\right)\right\|_{\mathbb{R}} = \max_{1\leq j\leq \#G_M^N} |t_j|.$$

2.5. **Tree-like structures and p-adic numbers.** Take $N=1$ and fix $M\in\mathbb{N}\smallsetminus\{0\}$, then $G_M^1 := G_M = p^{-M}\mathbb{Z}_p/p^M\mathbb{Z}_p$ is an additive group consisting of elements of the form

$$\boldsymbol{i} = \boldsymbol{i}_{-M}p^{-M} + \boldsymbol{i}_{-M+1}p^{-M+1} + \cdots + \boldsymbol{i}_0 + \cdots + \boldsymbol{i}_{M-1}p^{M-1}, \tag{2.4}$$

where the $\boldsymbol{i}_j$s belong to $\{0,1,\ldots,p-1\}$. Furthermore, the restriction of $|\cdot|_p$ to G_M induces an absolute value such that $|G_M|_p = \left\{0, p^{-(M-1)}, \cdots, p^{-1}, 1, \cdots, p^M\right\}$. We endow G_M with the metric induced by $|\cdot|_p$, and thus G_M becomes a finite ultrametric space. In addition, G_M can be identified with the set of branches (vertices at the top level) of a rooted tree with $2M+1$ levels and p^{2M} branches. Any element $\boldsymbol{i}\in G_M$ can be uniquely written as $p^{-M}\widetilde{\boldsymbol{i}}$, where

$$\widetilde{\boldsymbol{i}} = \widetilde{\boldsymbol{i}}_0 + \widetilde{\boldsymbol{i}}_1 p + \cdots + \widetilde{\boldsymbol{i}}_{2M-1}p^{2M-1} \in \mathbb{Z}_p/p^{2M}\mathbb{Z}_p,$$

with the $\widetilde{\boldsymbol{i}}_j$s belonging to $\{0,1,\ldots,p-1\}$. The elements of the $\mathbb{Z}_p/p^{2M}\mathbb{Z}_p$ are in bijection with the vertices at the top level of the above mentioned rooted tree. By definition the root of the tree is the only vertex at level 0. There are exactly p vertices at level 1, which correspond with the possible values of the digit $\widetilde{\boldsymbol{i}}_0$ in the p-adic expansion of $\widetilde{\boldsymbol{i}}$. Each of these vertices is connected to the root by a non-directed edge. At level ℓ, with $1\leq\ell\leq 2M$, there are exactly p^ℓ vertices, each vertex corresponds to a truncated expansion of $\widetilde{\boldsymbol{i}}$ of the form $\widetilde{\boldsymbol{i}}_0+\cdots+\widetilde{\boldsymbol{i}}_{\ell-1}p^{\ell-1}$. The vertex corresponding to $\widetilde{\boldsymbol{i}}_0+\cdots+\widetilde{\boldsymbol{i}}_{\ell-1}p^{\ell-1}$ is connected to a vertex $\widetilde{\boldsymbol{i}}_0'+\cdots+\widetilde{\boldsymbol{i}}_{\ell-2}'p^{\ell-2}$ at the level $\ell-1$ if and only if $\left(\widetilde{\boldsymbol{i}}_0+\cdots+\widetilde{\boldsymbol{i}}_{\ell-1}p^{\ell-1}\right) - \left(\widetilde{\boldsymbol{i}}_0'+\cdots+\widetilde{\boldsymbol{i}}_{\ell-2}'p^{\ell-2}\right)$ is divisible by $p^{\ell-1}$.

In conclusion, $\mathbb{Z}_p/p^{2M}\mathbb{Z}_p$ is a rooted tree, and $\mathbb{Z}_p$ is an infinite rooted tree. Now, the 1-dimensional unit sphere $\mathbb{Z}_p^\times$ is the disjoint union of sets of the form $j+p\mathbb{Z}_p$, for $j\in\{1,\ldots,p-1\}$. Each set of the form $j+p\mathbb{Z}_p$ is an infinite rooted tree. Then, $\mathbb{Z}_p^\times$ is a forest formed by the disjoint union of $p-1$ infinite rooted trees. On the other hand, $\mathbb{Q}_p^\times = \mathbb{Q}_p\smallsetminus\{0\}$ is a countable disjoint union of scaled versions of the forest $\mathbb{Z}_p^{\times}$, more precisely, $\mathbb{Q}_p^{\times} = \bigsqcup_{k=-\infty}^{k=+\infty} p^k \mathbb{Z}_p^{\times}$. The field of p-adic numbers has a fractal structure, see e.g. [1], [27].

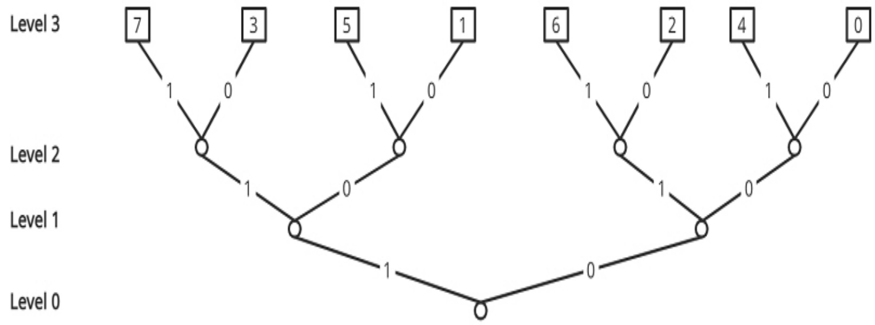

FIGURE 1. The rooted tree associated with the group $\mathbb{Z}_2/2^3\mathbb{Z}_2$. We identify the elements of $\mathbb{Z}_2/2^3\mathbb{Z}_2$ with the set of integers $\{0,\ldots,7\}$ with binary representation $\boldsymbol{i} = \boldsymbol{i}_0 + \boldsymbol{i}_1 2 + \boldsymbol{i}_3 2^2$, $\boldsymbol{i}_0, \boldsymbol{i}_1, \boldsymbol{i}_2 \in \{0,1\}$. Two leaves $\boldsymbol{i}, \boldsymbol{j} \in \mathbb{Z}_2/2^3\mathbb{Z}_2$ have a common ancestor at level 2 if and only if $\boldsymbol{i} \equiv \boldsymbol{j} \mod 2^2$, i.e., $\boldsymbol{i} = \boldsymbol{a}_0 + \boldsymbol{a}_1 2 + \boldsymbol{i}_2 2^2$ and $\boldsymbol{j} = \boldsymbol{a}_0 + \boldsymbol{a}_1 2 + \boldsymbol{j}_2 2^2$ with $\boldsymbol{i}_2, \boldsymbol{j}_2 \in \{0,1\}$. Now, for $\boldsymbol{i}, \boldsymbol{j} \in \mathbb{Z}_2/2^3\mathbb{Z}_2$ have a common ancestor at level 1 if and only if $\boldsymbol{i} \equiv \boldsymbol{j} \mod 2$. Notice that that the p-adic distance satisfies $-\log_2 |\boldsymbol{i} - \boldsymbol{j}|_2 = -$ (level of the first common ancestor of $\boldsymbol{i}$, $\boldsymbol{j}$).

3. p-ADIC CNNS: BASIC DEFINITIONS

We say that a function $f : \mathbb{R} \to \mathbb{R}$ is called a Lipschitz function if there exists a real constant $L(f) > 0$ such that, for all $x, y \in \mathbb{R}$, $|f(x) - f(y)| \leq L(f)|x - y|$. A relevant example is

$$f(x) = \frac{1}{2}\left(|x+1| - |x-1|\right).$$

3.1. **p-Adic discrete CNNs.** By considering G_M^N as a subset of $\mathbb{Q}_p^N$, $\left(G_M^N, \|\cdot\|_p\right)$ becomes a finite ultrametric space.

Definition 1. *An element $\boldsymbol{i}$ of G_M^N is called a cell. A p-adic discrete CNN is a dynamical system $CNN_d(\mathbb{A}, \mathbb{B}, U, Z)$ on G_M^N. The state $X_{\boldsymbol{i}}(t) \in \mathbb{R}$ of cell $\boldsymbol{i}$ is described by the following differential equations:*

(i) *state equation:*

$$\frac{dX(\boldsymbol{i},t)}{dt} = -X(\boldsymbol{i},t) + \sum_{\boldsymbol{j} \in G_M^N} \mathbb{A}(\boldsymbol{i},\boldsymbol{j}) Y(\boldsymbol{j},t) + \sum_{\boldsymbol{j} \in G_M^N} \mathbb{B}(\boldsymbol{i},\boldsymbol{j}) U(\boldsymbol{j}) + Z(\boldsymbol{i}),\ \boldsymbol{i} \in G_M^N,$$

(ii) *output equation:*

$$Y(\boldsymbol{j},t) = f(X(\boldsymbol{j},t)),$$

where $Y(\boldsymbol{i},t) \in \mathbb{R}$ is the output of cell $\boldsymbol{i}$ at the time t, $f : \mathbb{R} \to \mathbb{R}$ is a bounded Lipschitz function satisfying $f(0) = 0$. The function $U(\boldsymbol{i}) \in \mathbb{R}$ is the input

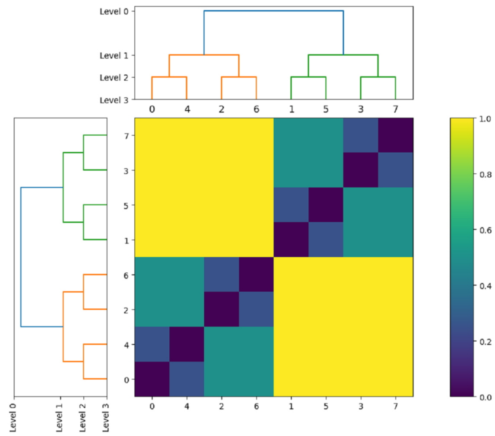

FIGURE 2. The heat map associated with the p-adic distance function on $\mathbb{Z}_2/2^3\mathbb{Z}_2$.

of the cell $\boldsymbol{i}$, $Z(\boldsymbol{i}) \in \mathbb{R}$ *is the threshold of cell* $\boldsymbol{i}$, *and* $\mathbb{A}, \mathbb{B} : G_M^N \times G_M^N \to \mathbb{R}$ *are the feedback operator and feedforward operator, respectively.*

Not all the cells of G_M^N are active. A cell $\boldsymbol{i}$ is connected with cell $\boldsymbol{j}$ if $\mathbb{A}(\boldsymbol{i},\boldsymbol{j}) \neq 0$ or $\mathbb{B}(\boldsymbol{i},\boldsymbol{j}) \neq 0$ for some $\boldsymbol{j} \in G_M^N$. Then, a p-adic discrete CNN is a dynamical system on

$$C_{N,M} := \left\{\boldsymbol{i} \in G_M^N; \mathbb{A}(\boldsymbol{i},\boldsymbol{j}) \neq 0 \text{ or } \mathbb{B}(\boldsymbol{i},\boldsymbol{j}) \neq 0 \text{ for some } \boldsymbol{j} \in G_M^N\right\}.$$

The topology of a p-adic discrete CNN depends on the functions $\mathbb{A}$, $\mathbb{B} : G_M^N \times G_M^N \to \mathbb{R}$. For general matrices $\mathbb{A}$, $\mathbb{B}$, it is difficult to give a graph-type description of the topology of the network. Our p-adic CNNs contain as a particular case the CNNs of Chua and Yang, see e.g. [8], [25]. In this article we focus on p-adic CNNs satisfying

$$\mathbb{A}(\boldsymbol{i},\boldsymbol{j}) = \mathbb{A}(\|\boldsymbol{i}-\boldsymbol{j}\|_p), \; \mathbb{B}(\boldsymbol{i},\boldsymbol{j}) = \mathbb{B}(\|\boldsymbol{i}-\boldsymbol{j}\|_p), \tag{3.1}$$

which are discrete CNNs having the space-invariant property. The fact that $\mathbb{A}$ and $\mathbb{B}$ are radial functions of $\|\cdot\|_p$ implies that the cells are organized in a tree like-structure with many layers.

3.2. p-Adic continuous CNNs.

Definition 2. *Given* $A(x,y)$, $B(x,y) \in L^1(\mathbb{Q}_p^N \times \mathbb{Q}_p^N)$, *and* U, $Z \in \mathcal{X}_\infty$, *a p-adic continuous CNN, denoted as CNN(A,B,U,Z), is the dynamical system given by the following differential equations: (i) state equation:*

$$\frac{\partial X(x,t)}{\partial t} = -X(x,t) + \int\limits_{\mathbb{Q}_p^N} A(x,y)Y(y,t)d^N y + \int\limits_{\mathbb{Q}_p^N} B(x,y)U(y)d^N y + Z(x), \tag{3.2}$$

where $x \in \mathbb{Q}_p^N$, $t \geq 0$, *and (ii) output equation:* $Y(x,t) = f(X(x,t))$. *We say that* $X(x,t) \in \mathbb{R}$ *is the state of cell* x *at the time* t, $Y(x,t) \in \mathbb{R}$ *is the output of cell* x *at the time* t. *Function* $A(x,y)$ *is the kernel of the feedback operator, while function* $B(x,y)$ *is the kernel of the feedforward operator. Function* U *is the input of the CNN, while function* Z *is the threshold of the CNN.*

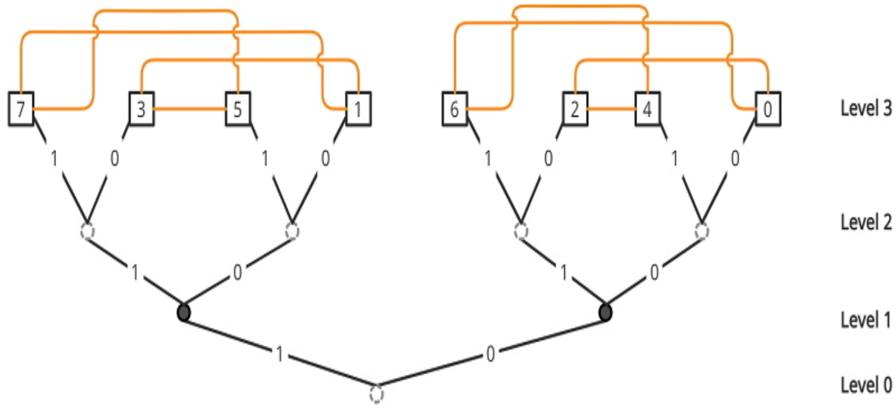

FIGURE 3. A 1-dimensional discrete 2-adic CNN with 8 cells: $C_{1,3} = \{0,1,2,3,4,5,7\} \subset \mathbb{Z}_2/2^3\mathbb{Z}_2 \subset 2^{-3}\mathbb{Z}_2/2^3\mathbb{Z}_2$. We set $\mathbb{B} = 0$ and $\mathbb{A}(\boldsymbol{i},\boldsymbol{j}) = [a_{\boldsymbol{i},\boldsymbol{j}}]$, with $a_{\boldsymbol{i},\boldsymbol{j}} \neq 0$ if $|\boldsymbol{i} - \boldsymbol{j}|_2 = 1/2$ and $\boldsymbol{i}$, $\boldsymbol{j} \in C_{1,3}$; $a_{\boldsymbol{i},\boldsymbol{j}} = 0$ otherwise.

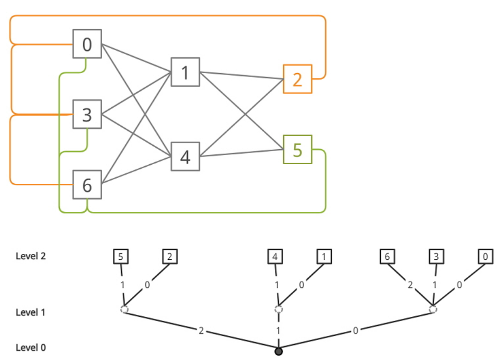

FIGURE 4. A 1-dimensional 3-adic CNN with 7 cells, $C_{1,2} = \{0,1,2,3,4,5,6\} \subset \mathbb{Z}_3/3^2\mathbb{Z}_3 \subset 3^{-2}\mathbb{Z}_3/3^2\mathbb{Z}_3$. We set $\mathbb{B} = 0$ and $\mathbb{A}(\boldsymbol{i},\boldsymbol{j}) = [a_{\boldsymbol{i},\boldsymbol{j}}]$, with $a_{\boldsymbol{i},\boldsymbol{j}} \neq 0$ if $|\boldsymbol{i} - \boldsymbol{j}|_3 = 1$ and $\boldsymbol{i},\boldsymbol{j} \in C_{1,2}$; $a_{\boldsymbol{i},\boldsymbol{j}} = 0$ otherwise.

We focus mainly in continuous CNNs having the space invariant property, i.e. $A(x,y) = A(\|x-y\|_p)$ and $B(x,y) = B(\|x-y\|_p)$ for some $A, B \in L^1$, however our results are valid for general p-adic continuous CNNs. Along this article the function $f(x) = \frac{1}{2}(|x+1| - |x-1|)$ will be fixed, for this reason it does not appear in the list of parameters of the CNNs.

3.3. **Discretization of p-adic continuous CNNs.** A central result of the present work is the fact that p-adic continuous CNNs are 'continuous versions' of p-adic discrete CNNs. More precisely, p-adic discrete CNNs are very good approximations of

p-adic continuous CNNs for sufficiently large M. We discuss here the discretization process in an intutive way (a formal theorem will be provided later on).

Intuitively, a discretization of a p-adic continuous CNN(A,B,U,Z) is obtained assuming that $X(\cdot,t)$, A, $Y(\cdot,t)$, B, U and Z belong to $\mathcal{D}^M$, i.e.

$$X(x,t)=\sum_{\boldsymbol{i}\in G_M^N}X(\boldsymbol{i},t)\Omega\left(p^M\left\|x-\boldsymbol{i}\right\|_p\right),\ \ Y(x,t)=\sum_{\boldsymbol{i}\in G_M^N}Y(\boldsymbol{i},t)\Omega\left(p^M\left\|x-\boldsymbol{i}\right\|_p\right),$$
$$U(x)=\sum_{\boldsymbol{i}\in G_M^N}U(\boldsymbol{i})\Omega\left(p^M\left\|x-\boldsymbol{i}\right\|_p\right),\ \ Z(x)=\sum_{\boldsymbol{i}\in G_M^N}Z(\boldsymbol{i})\Omega\left(p^M\left\|x-\boldsymbol{i}\right\|_p\right),$$
$$A(x,y)=\sum_{\boldsymbol{i}\in G_M^N}\sum_{\boldsymbol{j}\in G_M^N}A(\boldsymbol{i},\boldsymbol{j})\Omega\left(p^M\left\|x-\boldsymbol{i}\right\|_p\right)\Omega\left(p^M\left\|y-\boldsymbol{j}\right\|_p\right),$$
$$B(x,y)=\sum_{\boldsymbol{i}\in G_M^N}\sum_{\boldsymbol{j}\in G_M^N}B(\boldsymbol{i},\boldsymbol{j})\Omega\left(p^M\left\|x-\boldsymbol{i}\right\|_p\right)\Omega\left(p^M\left\|y-\boldsymbol{j}\right\|_p\right).$$

Notice that if $f:\mathbb{R}\to\mathbb{R}$, then

$$f\left(X(x,t)\right)=\sum_{\boldsymbol{i}\in G_M^N}f(X(\boldsymbol{i},t))\Omega\left(p^M\left\|x-\boldsymbol{i}\right\|_p\right)=Y(x,t).$$

Now,

$$\frac{\partial}{\partial t}X(x,t)=\sum_{\boldsymbol{i}\in G_M^N}\frac{\partial}{\partial t}X(\boldsymbol{i},t)\Omega\left(p^M\left\|x-\boldsymbol{i}\right\|_p\right),$$

and

$$\int_{\mathbb{Q}_p^N}A(x,y)f\left(X(y,t)\right)d^Ny$$
$$=\sum_{\boldsymbol{i}\in G_M^N}\left\{\sum_{\boldsymbol{j}\in G_M^N}A(\boldsymbol{i},\boldsymbol{j})f(X(\boldsymbol{j},t))\int_{\mathbb{Q}_p^N}\Omega\left(p^M\left\|y-\boldsymbol{j}\right\|_p\right)d^Ny\right\}\Omega\left(p^M\left\|x-\boldsymbol{i}\right\|_p\right)$$
$$=p^{-MN}\sum_{\boldsymbol{i}\in G_M^N}\left\{\sum_{\boldsymbol{j}\in G_M^N}A(\boldsymbol{i},\boldsymbol{j})Y(\boldsymbol{j},t))\right\}\Omega\left(p^M\left\|x-\boldsymbol{i}\right\|_p\right).$$

Similarly,

$$\int_{\mathbb{Q}_p^N}B(x,y)U(y)d^Ny=p^{-MN}\sum_{\boldsymbol{i}\in G_M^N}\left\{\sum_{\boldsymbol{j}\in G_M^N}B(\boldsymbol{i},\boldsymbol{j})U(\boldsymbol{j}))\right\}\Omega\left(p^M\left\|x-\boldsymbol{i}\right\|_p\right).$$

Therefore,

$$\frac{\partial}{\partial t}X(\boldsymbol{i},t)=-X(\boldsymbol{i},t)+\sum_{\boldsymbol{j}\in G_M^N}p^{-MN}A(\boldsymbol{i},\boldsymbol{j})Y(\boldsymbol{j},t)$$
$$+\sum_{\boldsymbol{j}\in G_M^N}p^{-MN}B(\boldsymbol{i},\boldsymbol{j})U(\boldsymbol{j})+Z(\boldsymbol{i}),\text{ for }\boldsymbol{i}\in G_M^N,$$

and $Y(\boldsymbol{i},t)=f(X(\boldsymbol{i},t))$, for $\boldsymbol{i}\in G_M^N$. This is exactly a p-adic discrete CNN with $\mathbb{A}(\boldsymbol{i},\boldsymbol{j})=p^{-MN}A(\boldsymbol{i},\boldsymbol{j})$, $\mathbb{B}(\boldsymbol{i},\boldsymbol{j})=p^{-MN}B(\boldsymbol{i},\boldsymbol{j})$.

Intuitively a p-adic continuous CNN has infinitely many layers, each layer corresponds to some M, which are organized in a hierarchical structure. For practical purposes, a p-adic continuous CNN is realized as a p-adic discrete CNN for M sufficiently large.

4. Stability of p-adic continuous CNN

Lemma 1. *Let f be a Lipschitz functions with $f(0) = 0$ and let E be a radial function in $L^1(\mathbb{Q}_p^N)$. Then, the mappings*

$$F_0 : g \to \int_{\mathbb{Q}_p^N} E(\|x-y\|_p) f(g(y)) d^N y$$

$$F_1 : g \to \int_{\mathbb{Q}_p^N} E(\|x-y\|_p) g(y) d^N y$$

are well defined bounded operators from $\mathcal{X}_\infty$ into itself.

Proof. We first notice that for all $g \in \mathcal{X}_\infty$, $F_0(g)(x)$ exists for all $x \in \mathbb{Q}_p^N$, since

$$|E(\|y\|_p)| \; |f(g(x-y))| \le L(f) \|g\|_\infty \; |E(\|y\|_p)| \,, \tag{4.1}$$

where $E(\|y\|_p) \in L^1(\mathbb{Q}_p^N)$. To show the continuity of $F_0(g)(x)$, we take a sequence $\{x_m\}_{m\in\mathbb{N}} \subset \mathbb{Q}_p^N$ such that $x_m \to x$. By using (4.1) and the dominated convergence theorem, $\lim_{m\to\infty} F_0(g)(x_m) = F_0(g)(x)$. Finally, we show that $F_0(g) \in \mathcal{X}_\infty$. By contradiction, assume that $F_0(g) \notin \mathcal{X}_\infty$. Then, there is a sequence $\{x_m\}_{m\in\mathbb{N}} \subset \mathbb{Q}_p^N$ such that $\lim_{m\to\infty} \|x_m\|_p = \infty$ and $\epsilon > 0$ such that $F_0(g)(x_m) > \epsilon$ for all $m \in \mathbb{N}$. By using (4.1) and the dominated convergence theorem, we have

$$\begin{aligned}\epsilon \le \lim_{m\to\infty} |F_0(g)(x_m)| &= \lim_{m\to\infty} \left| \int_{\mathbb{Q}_p^N} E(\|y\|_p) f(g(x_m - y)) d^N y \right| \\ &= \left| \int_{\mathbb{Q}_p^N} E(\|y\|_p) \left\{ \lim_{m\to\infty} f(g(x_m - y)) \right\} d^N y \right| = 0\end{aligned}$$

which contradicts the fact $\epsilon > 0$. The same argument allow us to show that $F_1(g) \in \mathcal{X}_\infty$ for any $g \in \mathcal{X}_\infty$. □

Lemma 2. *Assume $A, B \in L^1(\mathbb{Q}_p^N)$ are radial functions and that U, $Z \in \mathcal{X}_\infty$. For $g \in \mathcal{X}_\infty$, set*

$$\boldsymbol{H}(g) := \int_{\mathbb{Q}_p^N} A(\|x-y\|_p) f\left(g(y)\right) d^N y + \int_{\mathbb{Q}_p^N} B(\|x-y\|_p) U(y) d^N y + Z(x).$$

Then $\boldsymbol{H} : \mathcal{X}_\infty \to \mathcal{X}_\infty$ is a well-defined operator satisfying

$$\|\boldsymbol{H}(g) - \boldsymbol{H}(g')\|_\infty \le L(f) \|A\|_1 \|g - g'\|_\infty, \textit{ for } g,\, g' \in \mathcal{X}_\infty,$$

where $L(f)$ is the Lipschitz constant of f.

Proof. By Lemma 1, $\boldsymbol{H}:\mathcal{X}_\infty\to\mathcal{X}_\infty$ is a well-defined operator. Take $g,\ g'\in\mathcal{X}_\infty$, then

$$|\boldsymbol{H}(g)(x)-\boldsymbol{H}(g')(x)|=\left|\int_{\mathbb{Q}_p^N}A(\|x-y\|_p)\left(f\left(g(y)\right)-f\left(g'(y)\right)\right)d^Ny\right|$$

$$\leq\int_{\mathbb{Q}_p^N}|A(\|x-y\|_p)||f\left(g(y)\right)-f\left(g'(y)\right)|d^Ny\leq L(f)\|g-g'\|_\infty\int_{\mathbb{Q}_p^N}|A(\|x-y\|_p)|d^Ny$$

$$=L(f)\|A\|_1\|g-g'\|_\infty.$$

□

Remark 1. *(i) Lemma 1 remains valid if we replace the condition E is radial and integrable by the condition $E(x,y)$ is a continuous function with compact support. (ii) Under the hypothesis of part (i), Lemma 2 is valid for operators of the form*

$$\boldsymbol{L}g=\int_{\mathbb{Q}_p^N}A(x,y)f\left(g(y)\right)d^Ny+\int_{\mathbb{Q}_p^N}B(x,y)U(y)d^Ny+Z(x),$$

for $g\in\mathcal{X}_\infty$.

Proposition 1. *Assume that A, B, f satisfy hypotheses of Lemma 2 and that U, $Z\in\mathcal{X}_\infty$. Let τ be a fixed positive real number. Then for each $X_0\in\mathcal{X}_\infty$ there exists a unique $X\in C([0,\tau],\mathcal{X}_\infty)$ which satisfies*

$$X(x,t)=e^{-t}X_0\left(x\right)+\int_0^t e^{-(t-s)}\boldsymbol{H}(X(x,s))ds \tag{4.2}$$

where

$$\boldsymbol{H}X(x,t)=\int_{\mathbb{Q}_p^N}A(\|x-y\|_p)f(X(y,t))d^Ny+\int_{\mathbb{Q}_p^N}B(\|x-y\|_p)U(y)d^Ny+Z(x). \tag{4.3}$$

The function $X(x,t)$ is differentiable in t for all x, and it is a solution of equation (3.2) with initial datum X_0.

Proof. The result follows from Lemma 2, by using standard techniques in PDEs, see e.g. [22, Theorem 5.1.2]. To make the treatment comprehensive to a general audience, we provide some details here. First, define

$$\boldsymbol{T}(Y)=X_0e^{-t}+\int_0^t e^{-(t-s)}\boldsymbol{H}(Y(x,s))ds,$$

and $\mathcal{Y}=C([0,\tau],\mathcal{X}_\infty)$ which is a Banach space with the norm $\|\cdot\|_\infty$. By Lemma 2, $\boldsymbol{T}:\mathcal{Y}\to\mathcal{Y}$. If Y, $Y_1\in\mathcal{Y}$, then

$$\|\boldsymbol{T}(Y)(t)-\boldsymbol{T}(Y_1)(t)\|_\infty=\left\|\int_0^t e^{-(t-s)}\left\{\boldsymbol{H}(Y)(s)-\boldsymbol{H}(Y_1)(s)\right\}ds\right\|_\infty$$

$$\leq\int_0^t e^{-(t-s)}\|\boldsymbol{H}(Y)(s)-\boldsymbol{H}(Y_1)(s)\|_\infty\,ds\leq L(f)\|A\|_1\int_0^t\|Y-Y_1\|_\infty ds.$$

And hence,

$$\|\boldsymbol{T}^M(Y)(t)-\boldsymbol{T}^M(Y_1)(t)\|_\infty\leq\frac{\tau^M L(f)^M\|A\|_1^M}{M!}\|Y-Y_1\|_\infty,$$

for $M \geq 1$. By the contraction mapping theorem, there is a unique unique $X \in \mathcal{Y}$ which $\boldsymbol{T}(X) = X$. Moreover, since the right-hand side of (4.2) is differentiable in t, X is a solution of (3.2) with initial condition X_0. $\square$

Remark 2. *The contraction mapping theorem provides an iterative formula for $X(x,t)$. Set $X_1(x,t) = X_0(x)$ and*

$$X_{L+1}(x,t) = e^{-t}X_0(x) + \int_0^t e^{-(t-s)} H(X_L(x,s))ds, \text{ for } L = 1, 2, \dots,$$

then $\lim_{L\to\infty} \|X_L(\cdot,t) - X(\cdot,t)\|_\infty = 0$ *for each $t \leq \tau$, see e.g.* [22, Theorem 5.2.2].

Theorem 1. *Assume A, $B \in L^1(p^{-M_0}\mathbb{Z}_p^N)$ are radial functions, for some $M_0 \in \mathbb{N}$, and that U, Z, $X_0 \in \mathcal{X}_{M_0}$. We also assume that f is a Lipschitz functions with $f(0) = 0$. Then there is a unique $X \in C([0,\tau], \mathcal{X}_{M_0}) \cap C^1([0,\tau], \mathcal{X}_{M_0})$ satisfying (4.2), which is a solution of equation (3.2) with initial datum X_0.*

Remark 3. *This theorem remains valid if $A(x,y)$, $B(x,y)$ are continuous functions with compact support, see Remark 1.*

Proof. Since $\mathcal{X}_{M_0}$ is a subspace of $\mathcal{X}_\infty$, by applying Proposition 1, there exists a unique $X \in C([0,\tau], \mathcal{X}_\infty) \cap C^1([0,\tau], \mathcal{X}_\infty)$ that satisfies all the announced properties. By Remark 2, $\lim_{L\to\infty} \|X_L(\cdot,t) - X(\cdot,t)\|_\infty = 0$, where

$$X_{L+1}(x,t) = e^{-t}X_0(x) + \int_0^t e^{-(t-s)} H(X_L(x,s))ds, \text{ for } L = 1, 2, \dots.$$

By induction on L, if $X_L(\cdot,s) \in \mathcal{X}_{M_0}$, i.e. if

$$X_L(x,s) = \sum_{\boldsymbol{i}\in G_{M_0}^N} X_L(\boldsymbol{i},s)\Omega\left(p^{M_0}\|x-\boldsymbol{i}\|_p\right),$$
$$f(X_L(x,s)) = \sum_{\boldsymbol{i}\in G_{M_0}^N} Y_L(\boldsymbol{i},s)\Omega\left(p^{M_0}\|x-\boldsymbol{i}\|_p\right)$$

by using that

$$\int_0^t e^{-(t-s)} H(X_L(x,s))ds$$
$$= \sum_{\boldsymbol{i}\in G_{M_0}^N} \left(\int_0^t e^{-(t-s)} Y_L(\boldsymbol{i},s)ds\right) \left(\int_{\mathbb{Q}_p^N} A(\|x-y\|_p)\Omega\left(p^{M_0}\|y-\boldsymbol{i}\|_p\right)d^N y\right)$$
$$+ \sum_{\boldsymbol{i}\in G_{M_0}^N} U(\boldsymbol{i})(1-e^{-t}) \int_{\mathbb{Q}_p^N} B(\|x-y\|_p)\Omega\left(p^{M_0}\|y-\boldsymbol{i}\|_p\right)d^N y$$
$$+ \sum_{\boldsymbol{i}\in G_{M_0}^N} (1-e^{-t})Z(\boldsymbol{i})\Omega\left(p^{M_0}\|x-\boldsymbol{i}\|_p\right),$$

and that for any $E \in L^1(p^{-M_0}\mathbb{Z}_p^N)$ are radial function, with the convention that the support of E is the ball $p^{-M_0}\mathbb{Z}_p^N$,

$$\int_{\mathbb{Q}_p^N} E(\|x-y\|_p)\Omega\left(p^{M_0}\|y-\boldsymbol{i}\|_p\right)d^Ny = \int_{\boldsymbol{i}+p^{M_0}\mathbb{Z}_p^N} E(\|x-y\|_p)d^Ny$$

$$= \begin{cases} 0 & \text{if} \quad x \notin p^{-M_0}\mathbb{Z}_p^N \\ \int_{p^{M_0}\mathbb{Z}_p^N} E(\|z\|_p)d^Nz & \text{if} \quad x \in \boldsymbol{i}+p^{M_0}\mathbb{Z}_p^N \\ p^{-M_0N}E(\|\boldsymbol{i}-\boldsymbol{j}\|_p) & \text{if} \quad x \in \boldsymbol{j}+p^{M_0}\mathbb{Z}_p^N,\ \boldsymbol{i}\neq\boldsymbol{j}, \end{cases}$$

we conclude that

$$X_{L+1}(x,t) = e^{-t}X_0(x) + \tag{4.4}$$

$$\sum_{\boldsymbol{j}\in G_{M_0}^N}\left\{\sum_{\substack{\boldsymbol{i}\in G_{M_0}^N\\ \boldsymbol{i}\neq\boldsymbol{j}}} a(\boldsymbol{i},t)\,p^{-M_0N}A(\|\boldsymbol{i}-\boldsymbol{j}\|_p)\right\}\Omega\left(p^{M_0}\|y-\boldsymbol{j}\|_p\right)+$$

$$\sum_{\boldsymbol{j}\in G_{M_0}^N} a(\boldsymbol{j},t)\left(\int_{p^{M_0}\mathbb{Z}_p^N} A(\|z\|_p)d^Nz\right)\Omega\left(p^{M_0}\|y-\boldsymbol{j}\|_p\right)+$$

$$\sum_{\boldsymbol{i}\in G_{M_0}^N}\left\{\sum_{\substack{\boldsymbol{i}\in G_{M_0}^N\\ \boldsymbol{i}\neq\boldsymbol{j}}} U(\boldsymbol{i})(1-e^{-t})B(\|\boldsymbol{i}-\boldsymbol{j}\|_p)\right\}\Omega\left(p^{M_0}\|y-\boldsymbol{j}\|_p\right)+$$

$$\sum_{\boldsymbol{i}\in G_{M_0}^N} U(\boldsymbol{j})(1-e^{-t})\left(\int_{p^{M_0}\mathbb{Z}_p^N} B(\|z\|_p)d^Nz\right)\Omega\left(p^{M_0}\|y-\boldsymbol{j}\|_p\right)+(1-e^{-t})Z(\boldsymbol{i}),$$

i.e. $X_{L+1}(\cdot,s) \in \mathcal{X}_M$. And consequently, $\{X_L(\cdot,t)\}_{L\in\mathbb{N}\smallsetminus\{0\}}$ is a sequence in $\mathcal{X}_M$. Since $\mathcal{X}_M$ is closed in $\mathcal{X}_\infty$, $X(\cdot,t) \in \mathcal{X}_M$ for any $t \leq \tau$. □

Remark 4. *By using that*

$$p^{MN}\int_{p^M\mathbb{Z}_p^N} A(\|z\|_p)d^Nz \to A(0),\ p^{MN}\int_{p^M\mathbb{Z}_p^N} B(\|z\|_p)d^Nz \to B(0)$$

as $M \to \infty$, see e.g. [26, Theorem 1.14], *(4.4) provides an explicit approximation of the continuous CNN described in Theorem 1.*

Lemma 3. *Let τ be a fixed positive real number, let $X(x,t)$ be the solution given in Proposition 1, with $X(x,0) = X_0$. Then, for all $x, y \in \mathbb{Q}_p^N$ and $t \in (0,\tau)$,*

$$|X(x,t) - X(y,t)| \leq |X_0(x) - X_0(y)|e^{\|A\|_1 L(f)t}.$$

Moreover, if X_0 is a locally-constant function, i.e. $X_0(x) = X_0(y)$ for $y \in B_l(x)$, with $l = l(x) \in \mathbb{Z}$, for any $x \in \mathbb{Q}_p^N$, then $X(\cdot,t)$ is a locally-constant function and $X(x,t) = X(y,t)$ for $y \in B_l(x)$ for any $x \in \mathbb{Q}_p^N$.

Proof. Fix $x, y \in \mathbb{Q}_p^N$, the by Proposition 1 and Lemma 2, for all $t \in (0, \tau]$

$$|X(x,t) - X(y,t)| \leq e^{-t}|X_0(x) - X_0(y)| + \int_0^t e^{-(t-s)}|\boldsymbol{H}(X(x,s)) - \boldsymbol{H}(X(y,s))|ds$$
$$\leq |X_0(x) - X_0(y)| + L(f)\|A\|_1 \int_0^t |X(x,s) - X(y,s)|ds.$$

Thus, by Gronwall theorem, see [22, Theorem 5.1.1],

$$|X(x,t) - X(y,t)| \leq |X_0(x) - X_0(y)|e^{L(f)\|A\|_1 t}$$

for all $t \in (0, \tau)$. □

Definition 3. *A function $X_{stat}(x) := X_{stat}(x; A, B, U, Z) \in \mathcal{X}_\infty$ is called a stationary state of a p-adic continuous CNN(A, B, U, Z), if*

$$X_{stat}(x) = \int_{\mathbb{Q}_p^N} A(\|x-y\|_p)Y(y)d^N y + \int_{\mathbb{Q}_p^N} B(\|x-y\|_p)U(y)d^N y + Z(x),$$

where $Y(x) = f(X_{stat}(x))$ and $x \in \mathbb{Q}_p^N$.

Remark 5. *If a p-adic continuous CNN(A, B, U, Z) satisfies that $\|A\|_1 L(f) < 1$, then the CNN(A, B, U, Z) has a unique stationary state. This follows by the fact that, under this condition, $\boldsymbol{H}(X)$ becomes a contraction map in $\mathcal{X}_\infty$, cf. Lemmas 1, 2.*

Theorem 2. *All the states $X(x,t)$ of a p-adic continuous CNN(A, B, U, Z) are bounded for all time $t \geq 0$. More precisely, if*

$$X_{\max} := \|X_0\|_\infty + \|f\|_\infty \|A\|_1 + \|U\|_\infty \|B\|_1 + \|Z\|_\infty,$$

then

$$|X(x,t)| \leq X_{\max} \text{ for all } t \geq 0 \text{ and for all } x \in \mathbb{Q}_p^N. \tag{4.5}$$

In addition

$$X_-(x) := \lim\inf_{t\to\infty} X(x,t) \leq X(x,t) \leq \lim\sup_{t\to\infty} X(x,t) =: X_+(x),$$

for $x \in \mathbb{Q}_p^N$. If $X_-(x) = X_+(x) := X^*(x)$,then $X^*(x)$ is a stationary solution of the CNN(A, B, U, Z) and

$$X^*(x) \geq -\|f\|_\infty \|A\|_1 - \|U\|_\infty \|B\|_1 + Z(x). \tag{4.6}$$

Remark 6. *Condition (4.5) implies that $X(x,t)$ does not blow-up at finite time. The existence of a stationary state $X^*(x)$ means that the state of each cell of a p-adic continuous CNN most settle at stable equilibrium point after the transient has decayed to zero.*

Proof. By Proposition 1, see (4.2)-(4.3), by using that $|Y(y,t)| = |f(X(x,t))| \leq \|f\|_\infty$, we have

$$|\boldsymbol{H}(X(x,t))| \leq \int_{\mathbb{Q}_p^N} |A(\|x-y\|_p)||Y(y,t)|d^N y + \int_{\mathbb{Q}_p^N} |B(\|x-y\|_p)||U(y)|d^N y + |Z(x)|$$
$$\leq \|f\|_\infty \|A\|_1 + \|B\|_1 \|U\|_\infty + \|Z\|_\infty.$$

Therefore

$$\begin{aligned}\|X(x,t)\|_\infty &\le e^{-t}\|X_0\|_\infty + \int_0^t e^{-(t-s)}\left\|\boldsymbol{H}(X(x,s))\right\|_\infty ds\\ &\le \|X_0\|_\infty + \|f\|_\infty\|A\|_1 + \|B\|_1\|U\|_\infty + \|Z\|_\infty.\end{aligned}$$

This bound is valid for any $t \in [0,\tau]$, but τ is an arbitrary, the bound is valid for any $t \ge 0$.

The bound (4.5) implies existence of the functions:

$$\begin{aligned}X_+(x) &= \limsup_{t\to\infty} X(x,t) = \lim_{M\to\infty} \sup\{X(x,t); t > M\},\\ X_-(x) &= \liminf_{t\to\infty} X(x,t) = \lim_{M\to\infty} \inf\{X(x,t); t > M\}.\end{aligned}$$

Now assume that $\lim_{t\to\infty} X(x,t) = X^*(x)$ exists. By using that

$$\begin{aligned}\int_0^t e^{-(t-s)}\boldsymbol{H}(X(x,s))ds &= \int_0^t e^{-u}\boldsymbol{H}(X(x,t-u))du\\ &= \int_0^\infty 1_{[0,t]}(u)\, e^{-u}\boldsymbol{H}(X(x,t-u))du,\end{aligned}$$

and

$$\left|1_{[0,t]}(u)\, e^{-u}\boldsymbol{H}(X(x,t-u))\right| \le (\|f\|_\infty\|A\|_1 + \|B\|_1\|U\|_\infty + \|Z\|_\infty)\, e^{-u} \in L^1(\mathbb{R}),$$

and the dominated convergence and Lemma 2, it follows from (4.2) that

$$\begin{aligned}\lim_{t\to\infty} X(x,t) &= \int_0^\infty e^{-u} \lim_{t\to\infty}\left\{1_{[0,t]}(u)\,\boldsymbol{H}(X(x,t-u))\right\} du = \int_0^\infty e^{-u}\boldsymbol{H}(X^*(x,))du\\ &= \int_{\mathbb{Q}_p^N} A(\|x-y\|_p) f(X^*(x)) d^N y + \int_{\mathbb{Q}_p^N} B(\|x-y\|_p) U(y) d^N y + Z(x).\end{aligned}$$

□

5. Stability of p-adic discrete CNN and Approximation of Continuous CNNs

5.1. **The operators $\boldsymbol{P}_M$, $\boldsymbol{E}_M$.** We now define for $M \ge 1$, $\boldsymbol{P}_M : \mathcal{X}_\infty \to \mathcal{X}_M$ as

$$\boldsymbol{P}_M\varphi(x) = \sum_{\boldsymbol{i}\in G_M^N} \varphi(\boldsymbol{i})\,\Omega\left(p^M\|x-\boldsymbol{i}\|_p\right).$$

Therefore $\boldsymbol{P}_M$ is a linear bounded operator, indeed, $\|\boldsymbol{P}_M\| \le 1$.

We denote by $\boldsymbol{E}_M$, $M \ge 1$, the embedding $\mathcal{X}_M \to \mathcal{X}_\infty$. The following result is a consequence of the above observations. If $\mathcal{Z}$, $\mathcal{Y}$ are real Banach spaces, we denote by $\mathfrak{B}(\mathcal{Z},\mathcal{Y})$, the space of all linear bounded operators from $\mathcal{Z}$ into $\mathcal{Y}$.

Lemma 4. [30, Lemma 2] *With the above notation, the following assertions hold true:*
(i) $\mathcal{X}_\infty$, $\mathcal{X}_M$ *for* $M \ge 1$, *are real Banach spaces, all with the norm* $\|\cdot\|_\infty$;
(ii) $\boldsymbol{P}_M \in \mathfrak{B}(\mathcal{X}_\infty,\mathcal{X}_M)$ *and* $\|\boldsymbol{P}_M\varphi\|_\infty \le \|\varphi\|_\infty$ *for any* $M \ge 1$, $\varphi \in \mathcal{X}_\infty$;
(iii) $\boldsymbol{E}_M \in \mathfrak{B}(\mathcal{X}_M,\mathcal{X}_\infty)$ *and* $\|\boldsymbol{E}_M\varphi\|_\infty = \|\varphi\|_\infty$ *for any* $M \ge 1$, $\varphi \in \mathcal{X}_M$;
(iv) $\boldsymbol{P}_M\boldsymbol{E}_M\varphi = \varphi$ *for* $M \ge 1$, $\varphi \in \mathcal{X}_M$;
(v) $\lim_{M\to\infty}\|\varphi - \boldsymbol{P}_M\varphi\|_\infty = 0$ *for any* $\varphi \in \mathcal{X}_\infty$;
(vi) $\lim_{M\to\infty}\|\boldsymbol{E}_M\boldsymbol{P}_M\phi - \phi\|_\infty = 0$ *for all* $\phi \in \mathcal{X}_\infty$.

Proposition 2. *Assume that* $A(\|x\|_p)$, $B(\|x-y\|_p)$, $U(x)$, $Z(x) \in \mathcal{X}_M$, $M \geq 1$. *Let* τ *be a fixed positive real number. Consider the initial value problem:*

$$(5.1)\qquad \begin{cases} X \in C([0,\tau],\mathcal{X}_M) \cap C^1([0,\tau],\mathcal{X}_M) \\ \frac{\partial X(x,t)}{\partial t} = -X(x,t) + \int\limits_{\mathbb{Q}_p^N} A(\|x-y\|_p) f(X(x,t)) d^N y \\ + \int\limits_{\mathbb{Q}_p^N} B(\|x-y\|_p) U(y) d^N y + Z(x),\ x \in B_M^N,\ t \geq 0 \\ X(x,0) = X_0 \in \mathcal{X}_M. \end{cases}$$

There exists a unique $X \in C([0,\tau],\mathcal{X}_M)$ *which satisfies*

$$X(x,t) = e^{-t} X_0(x) + \int_0^t e^{-(t-s)} \boldsymbol{H}(X(x,s)) ds$$

where

$$\boldsymbol{H}(X)(x,t) = \int\limits_{\mathbb{Q}_p^N} A(\|x-y\|_p) f(X(x,t)) d^N y + \int\limits_{\mathbb{Q}_p^N} B(\|x-y\|_p) U(y) d^N y + Z(x).$$

The function $X(x,t)$ *is a solution of equation 5.1 with initial datum* X_0.

Proof. The result is established by using the argument given in the proof of Theorem 1. □

By the discussion presented in section 3.3, (5.1) describes a p-adic discrete CNN. Furthermore, Theorem 2 is also valid for discrete CNN in $\mathcal{X}_M$.

Remark 7. *By using the discretization procedure given in Section 3.3 and in the proof of Theorem 1, Proposition 2 implies that the initial value problem*

$$\begin{cases} X_M \in C([0,\tau],\mathcal{X}_M) \cap C^1([0,\tau],\mathcal{X}_M) \\ \frac{\partial X_M}{\partial t} = -X_M + \boldsymbol{P}_M \boldsymbol{H}(\boldsymbol{E}_M X_M) \\ X_M(0) = \boldsymbol{P}_M(X_0) \end{cases}$$

has a unique solution for an arbitrary $\tau > 0$.

Theorem 3. *All the states* $X(\boldsymbol{i},t)$, $\boldsymbol{i} \in G_M^N$, *in a p-adic discrete CNN are bounded for all time* $t \geq 0$. *More precisely, if*

$$X_{\max} := \max_{\boldsymbol{i}\in G_M^N} |X_0(\boldsymbol{i})| + p^{-MN} \left(\max_{\boldsymbol{i}\in G_M^N} |f(\boldsymbol{i})| \right) \sum_{\boldsymbol{i}\in G_M^N} |A(\boldsymbol{i})| + p^{-MN} \left(\max_{\boldsymbol{i}\in G_M^N} |U(\boldsymbol{i})| \right) \sum_{\boldsymbol{i}\in G_M^N} |A(\boldsymbol{i})| + \max_{\boldsymbol{i}\in G_M^N} |Z(\boldsymbol{i})|,$$

then

$$|X(\boldsymbol{i},t)| \leq X_{\max} \textit{ for all } t \geq 0 \textit{ and for all } \boldsymbol{i} \in G_M^N.$$

In addition

$$X_-(\boldsymbol{i}) := \lim\inf_{t\to\infty} X(\boldsymbol{i},t) \leq X(\boldsymbol{i},t) \leq \lim\sup_{t\to\infty} X(\boldsymbol{i},t) =: X_+(\boldsymbol{i}),$$

for $\boldsymbol{i} \in G_M^N$. *If* $X_-(\boldsymbol{i}) = X_+(\boldsymbol{i}) := X^*(\boldsymbol{i})$, *then*

$$X^*(\boldsymbol{i}) = \sum_{\boldsymbol{j}\in G_M^N} p^{-MN} A(\|\boldsymbol{i}-\boldsymbol{j}\|_p) f(X^*(\boldsymbol{i})) + \sum_{\boldsymbol{j}\in G_M^N} p^{-MN} B(\|\boldsymbol{i}-\boldsymbol{j}\|_p) U(\boldsymbol{j}) + Z(\boldsymbol{i}), \text{ for } \boldsymbol{i} \in G_M^N,$$

and

$$X^*(\boldsymbol{i}) \geq -p^{-MN}\left(\max_{\boldsymbol{i}\in G_M^N} |f(\boldsymbol{i})|\right) \sum_{\boldsymbol{i}\in G_M^N} |A(\boldsymbol{i})| - p^{-MN}\left(\max_{\boldsymbol{i}\in G_M^N} |U(\boldsymbol{i})|\right) \sum_{\boldsymbol{i}\in G_M^N} |A(\boldsymbol{i})| + Z(\boldsymbol{i}) \text{ , for all } \boldsymbol{i} \in G_M^N.$$

Theorem 4. *Let X be the solution of a continuous p-adic CNN given by Theorem 1 with initial condition X_0. Let X_M be the solution of the Cauchy problem*

$$\begin{cases} \frac{dX_M}{dt} = -X_M + \boldsymbol{P}_M \boldsymbol{H}(\boldsymbol{E}_M X_M) \\ X_M(0) = \boldsymbol{P}_M(X_0), \end{cases} \tag{5.2}$$

cf. Proposition 2 and Remark 7. Then

$$\lim_{M\to\infty} \sup_{0\leq t\leq\tau} \|X_M(t) - X(t)\|_\infty = 0.$$

Proof. The result follows from Lemma 4, Propositions 1, 2, by using standard techniques of approximation for evolution equations, see e.g. [22, Theorem 5.4.7]. See also [30, Section 9.1 and Theorem 7] for an in-depth discussion of similar matters. □

6. Numerical Simulations of p-Adic Continuous CNNs

In this section we present some numerical simulations of the solutions of several p-adic continuous CNNs in dimension 1. We give two numerical schemes for the numerical approximation of the solutions.

6.1. Numerical Scheme A.

Lemma 5. *Let $H(|\cdot|_p) \in L^1(\mathbb{Q}_p)$ and let $g \in \mathcal{X}_\infty$. We set $G_k = p^{-k}\mathbb{Z}_p/p^k\mathbb{Z}_p$, $k \in \mathbb{N}$. Then*

$$\int_{\mathbb{Q}_p} H(|x-y|)g(y)dy = \lim_{k\to\infty} \sum_{\boldsymbol{i}\in G_k;\ \boldsymbol{i}\neq x} g(\boldsymbol{i})p^{-k}H(|x-\boldsymbol{i}|_p) + g(x)(1-p^{-1})\sum_{l=k}^{\infty} H(p^{-l})p^{-l}.$$

Proof. By Lemma 4-(v), $\lim_{k\to\infty}\sum_{\boldsymbol{i}\in G_k} g(\boldsymbol{i})\Omega(p^k|x-\boldsymbol{i}|_p)=g(x)$, now by the dominated convergence theorem,

$$\int_{\mathbb{Q}_p} H(|x-y|_p)g(y)dy=\lim_{k\to\infty}\sum_{\boldsymbol{i}\in G_k} g(\boldsymbol{i})\int_{\mathbb{Q}_p} H(|x-y|_p)\Omega(p^k|y-\boldsymbol{i}|_p)dy$$

$$=\lim_{k\to\infty}\sum_{\boldsymbol{i}\in G_k} g(\boldsymbol{i})\int_{x-\boldsymbol{i}+p^k\mathbb{Z}_p} H(|z|_p)dz.$$

Now, if $|x-\boldsymbol{i}|_p>p^{-k}$, i.e. $x\neq\boldsymbol{i}$ in G_k, then

$$\int_{x-\boldsymbol{i}+p^k\mathbb{Z}_p} H(|z|_p)dz=p^{-k}H(|x-\boldsymbol{i}|_p).$$

And if $|x-\boldsymbol{i}|_p\leq p^{-k}$, i.e. $x=\boldsymbol{i}$ in G_k, then

$$\int_{x-\boldsymbol{i}+p^k\mathbb{Z}_p} H(|z|_p)dz=\sum_{l=k}^{\infty} H(p^{-l})(1-p^{-1})p^{-l}.$$

□

We now assume that A, B are radial integrable functions, and that U, Z, $X_0\in\mathcal{X}_\infty$. Based on the continuity of operators $\boldsymbol{A},\boldsymbol{B}:\mathcal{X}_\infty\to\mathcal{X}_\infty$ and the formula given in Lemma 5, we can approximate the solution $X(x,t)$ of a p-adic continuous CNN(A,B,U,Z) by p^{2k} ODEs, $k\geq 1$, of the form

$$\frac{d}{dx}X(\boldsymbol{i},t)=-X(\boldsymbol{i},t)+\sum_{\boldsymbol{j}\in G_k;\,\boldsymbol{j}\neq\boldsymbol{i}} f(X(\boldsymbol{j},t))p^{-k}A(|\boldsymbol{i}-\boldsymbol{j}|_p)+$$
$$f(X(\boldsymbol{i},t))(1-p^{-1})\sum_{l=k}^{k_{\max}} A(p^{-l})p^{-l}+\sum_{\boldsymbol{j}\in G_k;\,\boldsymbol{j}\neq\boldsymbol{i}} U(\boldsymbol{i})p^{-k}B(|\boldsymbol{i}-\boldsymbol{j}|_p)$$
$$+U(\boldsymbol{j})(1-p^{-1})\sum_{l=k}^{k_{\max}} B(p^{-l})p^{-l}+Z(\boldsymbol{i}),\text{ for }\boldsymbol{i}\in G_k.$$

In the simulations the parameters k, k_{max} were chosen by trial and error on a case by case approach. The sum $\sum_{l=k}^{k_{\max}} A(p^{-l})p^{-l}$ can be approximated by $A(p^{-k})p^{-k}$ in the cases were $A(p^{-k})p^{-k}$ is the dominant term in $\sum_{l=k}^{k_{\max}} A(p^{-l})p^{-l}$.

6.2. **Numerical Scheme B.**

Lemma 6. *Let $H(x)=\sum_{l=0}^{m} H_l\Omega(p^{k_l}|x-b_l|_p)$ be a test function and let $g\in\mathcal{X}_\infty$. Take $G_k=p^{-k}\mathbb{Z}_p/p^k\mathbb{Z}_p$, $k\in\mathbb{N}$, as before. Then*

$$\int_{\mathbb{Q}_p} H(x-y)g(y)dy=\lim_{k\to\infty}\sum_{\boldsymbol{i}\in G_k} g(\boldsymbol{i})\sum_{l=0}^{m} H_l\int_{\mathbb{Q}_p}\Omega(p^{k_l}|x-\boldsymbol{i}-b_l-y|_p)\Omega(p^k|y|_p)dy$$

$$=\lim_{k\to\infty}\sum_{\boldsymbol{i}\in G_k} g(\boldsymbol{i})\sum_{l=0}^{m} H_l p^{\min(-k,-k_l)}\Omega\left(p^{-\max(-k,-k_l)}|x-\boldsymbol{i}-b_l|_p\right)$$

$$=\lim_{k\to\infty}\sum_{\boldsymbol{i}\in G_k} g(\boldsymbol{i})\sum_{l=0}^{m} H_l p^{-\max(k,k_l)}\Omega\left(p^{\min(k,k_l)}|x-\boldsymbol{i}-b_l|_p\right).$$

Proof. It is sufficient to consider the case where $H(x)=\Omega(p^{k_H}|x-b_H|_p)$ for some $k_H\in\mathbb{Z}$ and $b_H\in\mathbb{Q}_p$. Since $g(x)=\lim_{k\to\infty}\sum_{\boldsymbol{i}\in G_k}g(\boldsymbol{i})\Omega(p^k|x-\boldsymbol{a}|_p)$, we have

$$\begin{aligned}\int_{\mathbb{Q}_p}H(x-y)g(y)dy&=\lim_{k\to\infty}\sum_{\boldsymbol{i}\in G_k}g(\boldsymbol{i})\int_{\mathbb{Q}_p}\Omega(p^{k_H}|x-b_H-y|_p)\Omega(p^k|y-\boldsymbol{i}|_p)dy\\&=\lim_{k\to\infty}\sum_{\boldsymbol{i}\in G_k}g(\boldsymbol{i})\int_{\mathbb{Q}_p}\Omega(p^{k_H}|(x-b_H-\boldsymbol{i})-y|_p)\Omega(p^k|y|_p)dy.\end{aligned}$$

Without loss of generality, we may assume that $k_H\leq k$, and since any two balls are disjoint or one contains the other, then $B_{-k}\cap B_{-k_H}(x-b_H-a)=\emptyset$ or $B_{-k}\cap B_{-k_H}(x-b_H-\boldsymbol{i})=B_{-k}$. The latter case occurs if and only if $0\in B_{-k_H}(x-b_H-\boldsymbol{i})$, i.e. when $|x-b_H-a|_p\leq p^{-k_H}$. Therefore

$$\int_{\mathbb{Q}_p}\Omega(p^{k_H}|x-b_H-\boldsymbol{i}-y|_p)\Omega(p^k|y|_p)dy=p^{-k}\Omega\left(p^{k_H}|x-b_H-\boldsymbol{i}|_p\right).$$

□

We now assume that U, Z, $X_0\in\mathcal{X}_\infty$ and that A, B are test functions of the form

$$A(x)=\sum_{l=0}^{m_A}A_l\Omega(p^{k_l}|x-a_l|_p),\quad B(x)=\sum_{l=0}^{m_B}B_l\Omega(p^{k_l}|x-b_l|_p.$$

Based on the continuity of operators $\boldsymbol{A},\boldsymbol{B}:\mathcal{X}_\infty\to\mathcal{X}_\infty$ and the formula given in Lemma 6, we can approximate the solution $X(x,t)$ of a p-adic continuous CNN by p^{2k} ODEs, $k\geq1$, of the form

$$\begin{aligned}\frac{d}{dx}X(\boldsymbol{i},t)&=-X(\boldsymbol{i},t)+\sum_{\boldsymbol{j}\in G_k}f(X(\boldsymbol{j},t))\sum_{l=0}^{m_A}A_lp^{-\max(k,k_l)}\Omega\left(p^{\min(k,k_l)}|\boldsymbol{i}-\boldsymbol{j}-a_l|_p\right)\\&+\sum_{\boldsymbol{j}\in G_k}U(\boldsymbol{j})\sum_{l=0}^{m_B}B_lp^{-\max(k,k_l)}\Omega\left(p^{\min(k,k_l)}|\boldsymbol{i}-\boldsymbol{j}-b_l|_p\right)+Z(\boldsymbol{i}),\text{ for }\boldsymbol{i}\in G_k.\end{aligned}$$

It is possible to combine the approximations given in numeric schemes A, B.

6.3. **A remark on the visualization of finite rooted trees.** The discretizations of the kernels A, B are functions on $G_k\times G_k$, while the input U and X_0 are functions on G_k. We use systematically heat maps to present these functions. We always include a plot of the tree G_k. By convention we identify the leaves of the tree G_k with the set of rational numbers $\{0,1/p^k,2/p^k,\ldots,(p^{2k}-1)/p^k\}$. Furthermore, we label the levels of G_k with integers from the set $\{-k,-k+1,\ldots,0,1,\ldots k-1\}$. The level l consists of the cells $\boldsymbol{i}$, $\boldsymbol{j}$ such that

$$-\log_p(|\boldsymbol{i}-\boldsymbol{j}|_p)=(\text{the level of the first common ancestor of }\boldsymbol{i},\boldsymbol{j})=l.$$

6.4. **First Simulation.** In this example, we take $k=2$, $p=2$, which means that we use a tree with $2^4=16$ leaves and 4 levels. A basic application of the classical CNNs is image processing, see e.g. [8]. In this example we present a one-dimensional edge detector, which is a p-adic, one-dimensional analog of the examples 3.1 and 3.2 in [8]. The input U is a image having three levels:

$$U(x)=\sum_{\boldsymbol{i}\in G_2}U_{\boldsymbol{i}}\Omega(2^2|x-\boldsymbol{i}|_2),\ \ U_{\boldsymbol{i}}=\begin{cases}-1 & \text{if} \qquad\qquad \boldsymbol{i}=1,2,1/4,13/4\\ 0 & \text{if} \qquad\qquad \boldsymbol{i}=1/2,9/4,5/4\\ 1 & \text{otherwise,}\end{cases}$$

$x\in G_2=2^{-2}\mathbb{Z}_2/2^2\mathbb{Z}_2$. As in [8] we take $X_0(x)=0$, $A(x)=0$. To construct template B, we identify a matrix with a test function. We use

$$B(x)=64\Omega(2^2|x|_2)-4\sum_{\boldsymbol{i}\in G_2;\ \boldsymbol{i}\neq 0}\Omega(2^2|x-\boldsymbol{i}|_2),\ x\in G_2.$$

Finally, we take $Z(x)=-\Omega(2^{-2}|x|_2)$, $\quad f(x)=\frac{1}{2}\left(|x+1|-|x-1|\right)$. The output $Y(x,t)$ consists of the edges on the input U, see Figure 7.

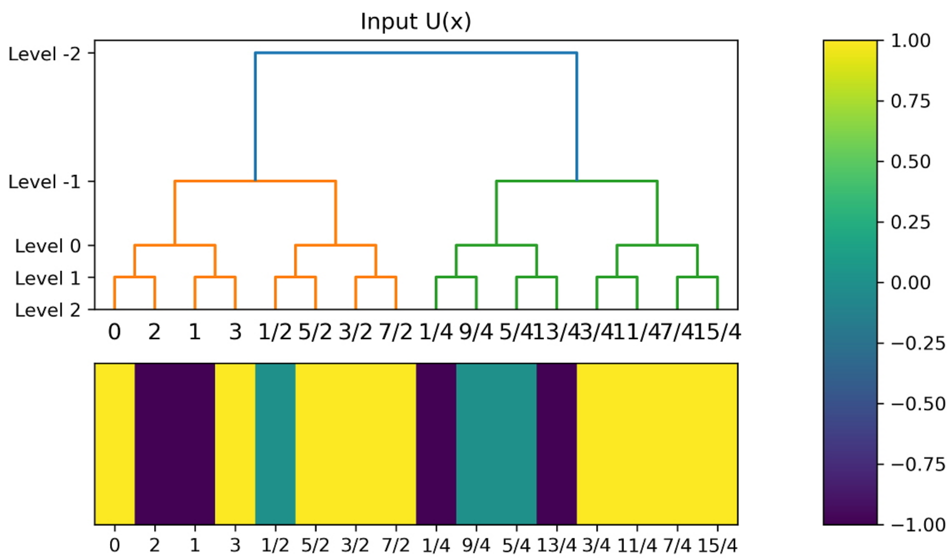

FIGURE 5. Simulation 1. Heat map $U(x)$.

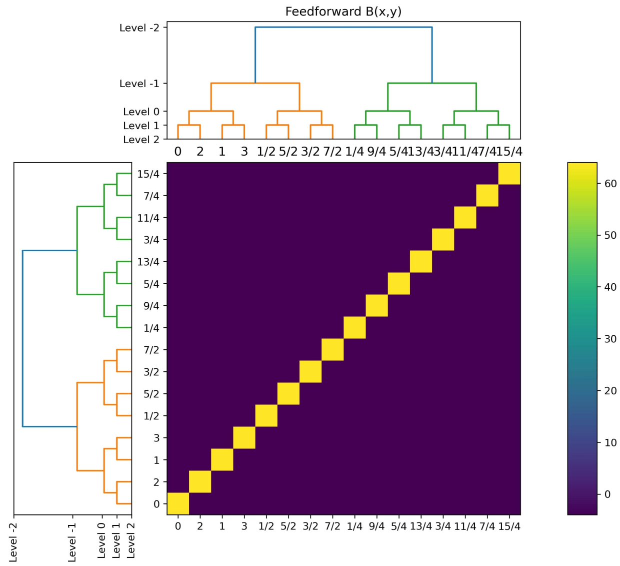

FIGURE 6. Simulation 1. Heat map of $B\left(|x-y|_2\right)$, x, $y \in G_2$.

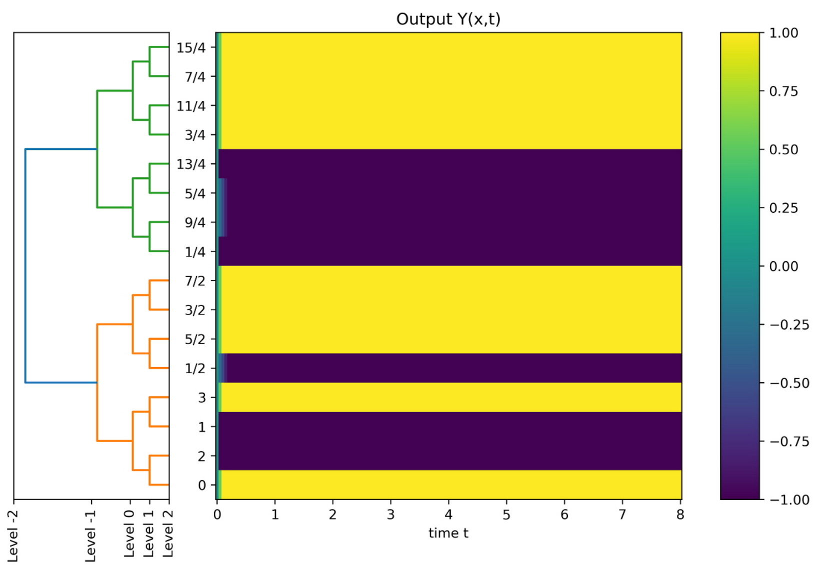

FIGURE 7. Simulation 1. Step 0.05.

6.5. **Second Simulation.** In this example, we take $k=2$, $p=2$, which means that we use a tree with $2^4=16$ leaves and 4 levels. We consider a CNN with the followin parameters:

$$A(x)=\Omega(2^2|x-2^{-2}|_2),\quad B(x)=U(x)=\Omega(2^2|x|_2),\ Z(x)=0,\ x\in G_2.$$

We set $X_0(x)=0$ and $\quad f(x)=\frac{1}{2}\left(|x+1|-|x-1|\right)$.

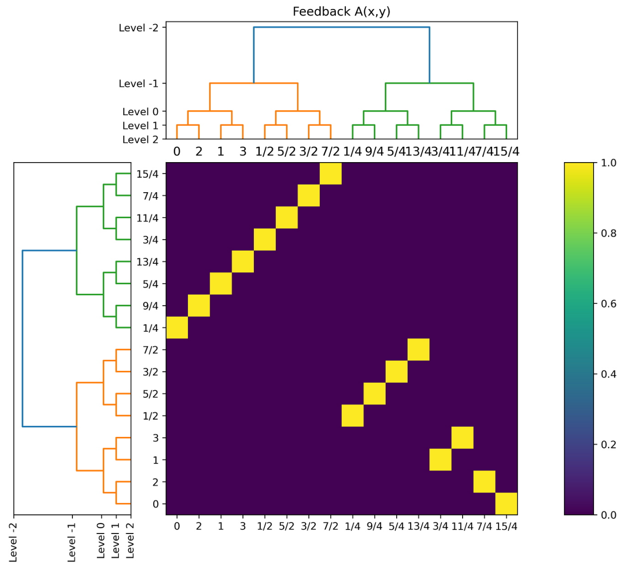

FIGURE 8. Simulation 2. Heat map $A(x-y)$ for $x, y \in G_2$.

In this network, we have $A(\boldsymbol{i},\boldsymbol{j}) = A(\boldsymbol{i}-\boldsymbol{j}) = \Omega\left(2^2 \left|\boldsymbol{i}-\boldsymbol{j}-2^{-1}\right|_2\right)$, $B(\boldsymbol{i},\boldsymbol{j}) = B(|\boldsymbol{i}-\boldsymbol{j}|_2) = \Omega\left(2^2 |\boldsymbol{i}-\boldsymbol{j}|_2\right) = \delta_{\boldsymbol{i},\boldsymbol{j}}$, where $\delta_{\boldsymbol{i},\boldsymbol{j}}$ denotes the Konecker delta function. This network does not have the space-invariant property because $A(\boldsymbol{i},\boldsymbol{j}) = \Omega\left(2^2 \left|\boldsymbol{i}-\boldsymbol{j}-2^{-1}\right|_2\right)$ is not a radial function. Due to this fact, $A(\boldsymbol{i},\boldsymbol{j})$ is not a symmetric matrix. For instance:

$$A(\frac{15}{4},0)=0,\ A(0,\frac{15}{4})=1,\quad A(\frac{1}{4},0)=0,\ A(0,\frac{1}{4})=1.$$

Our interpretation is that there is a connection from cell $\frac{15}{4}$ to cell 0, and a connection from cell 0 to cell $\frac{1}{4}$. This assertion is confirm by the ouput $Y(x,t)$, see Figure 11. Notice that $Y(\frac{1}{2},t)\neq 0$ and $A(\frac{1}{2},0)=A(0,\frac{1}{2})=0$. But $A(\frac{1}{4},\frac{1}{2})=0$, $A(\frac{1}{2},\frac{1}{4})=1$, then there is a connection from cell $\frac{1}{4}$ to cell $\frac{1}{2}$, which explains the fact that $Y(\frac{1}{2},t)\neq 0$.

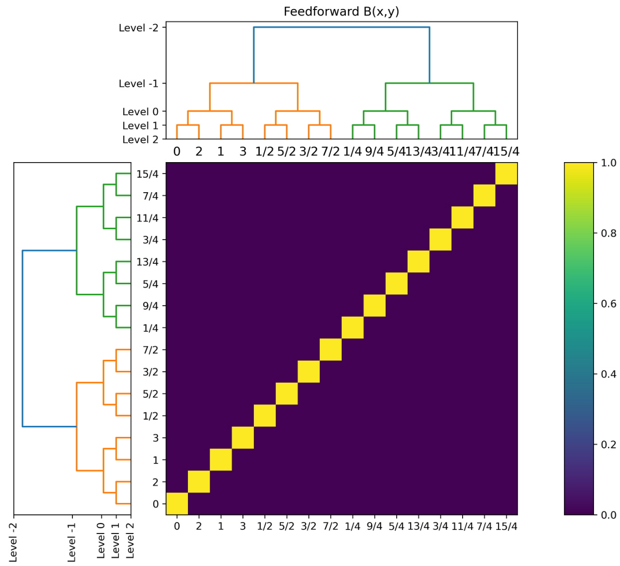

FIGURE 9. Simulation 2. Heat map of $B\left(|x-y|_2\right)$ for $x, y \in G_2$.

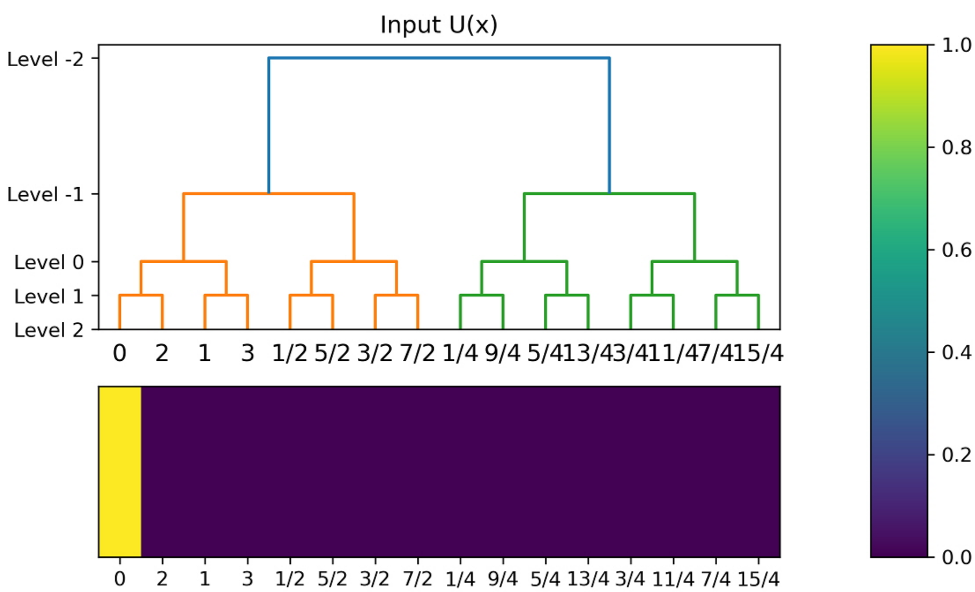

FIGURE 10. Simulation 2. Heat map of $U(x)$.

The numerical solutions is given in Figure 11. We now take $A(x) = B(x) = \Omega(2^2|x|_2)$. In this case the output $Y(x,t)$ changes completely, see Figure 12.

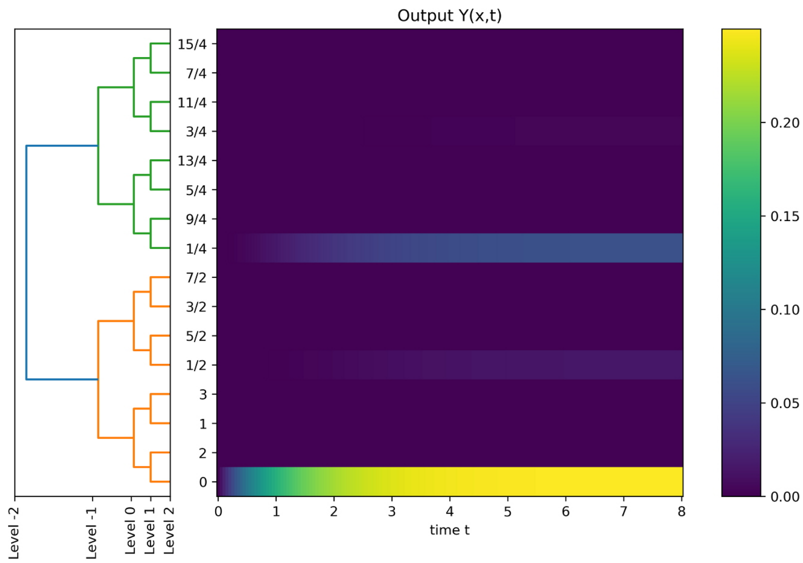

FIGURE 11. Simulation 2. Step 0.05.

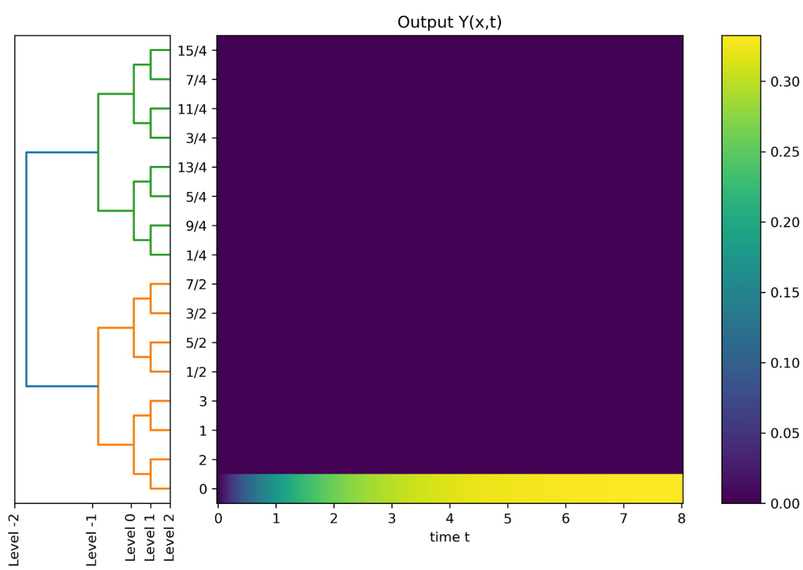

FIGURE 12. Simulation 2. Output with $A(x) = B(x) = \Omega(2^2|x|_2)$ and step 0.05.

6.6. **Third Simulation.** In this example, we take $k = 3$, $p = 2$, which means that we use a tree with $2^6 = 64$ leaves and 12 levels. We consider a CNN with the following parameters: $A(x) = \Omega(2^3|x - 2^{-2}|_2)$, $B(|x|_2) = \Omega(2^3|x|_2)$, $U(x) = \sin(p^4|x|_2)$, $Z(x) = 0.15\Omega(2^{-2}|x|_2)$ for $x \in G_3 = 2^{-3}\mathbb{Z}_3/2^3\mathbb{Z}_3$. We set $X_0(x) = 0$ and $f(x) = \frac{1}{2}(|x+1| - |x-1|)$.

As a consequence of the fractal nature of the p-adic numbers, the p-adic CNNs exhibit self-similarity in several ways. For instance, the graph of the kernel $A(x, y)$ is a self-similar set, this follows by comparing the graphs given in simulations 2 and 3 for this kernel. In addition, the output $Y(x, t) = 0$ when the norm $|x|_2$ is sufficiently large. In this simulation the CNN produces a pattern similar to the input.

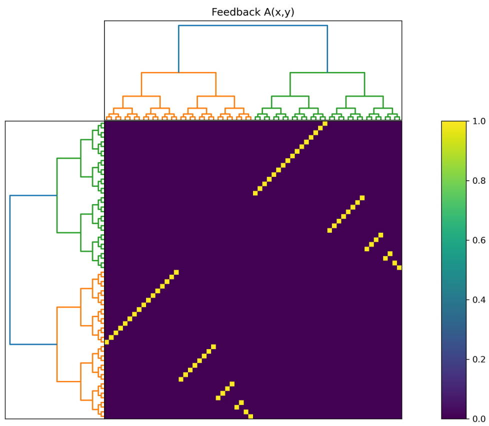

FIGURE 13. Simulation 3. Heat map of $A(x - y)$ for x, $y \in G_3$.

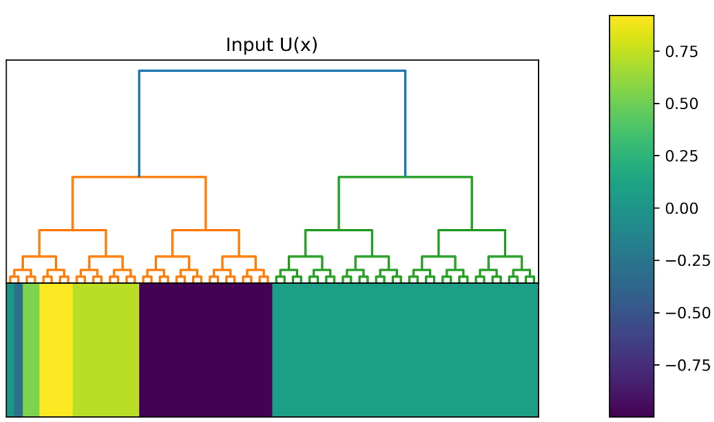

FIGURE 14. Simulation 3. Heat map of $U(x)$.

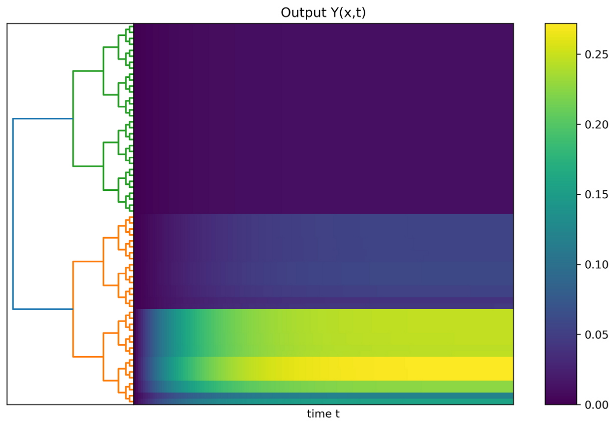

FIGURE 15. Simulation 3. Step 0.05.

6.7. **Fourth Simulation.** In this example, we take $k = 2$, $p = 2$, which means that we use a tree with $2^4 = 16$ leaves and 4 levels. The parameters of the CNN are $A(x) = \Omega\left(2^2\left|x - 2^{-2}\right|_2\right)$, $B(x) = U(x) = Z(x) = 0$, we set $X_0\left(x\right) = \Omega\left(2^2\left|x\right|_2\right)$, $f(x) = \frac{1}{2}\left(|x+1| - |x-1|\right)$ for $x \in G_2$.

In this example, at time zero the cells near the origin are excited. Which causes all the cells of the network to activate. The activation can be seen in the Fourier transform of the output. After some time the network returns to a state of rest.

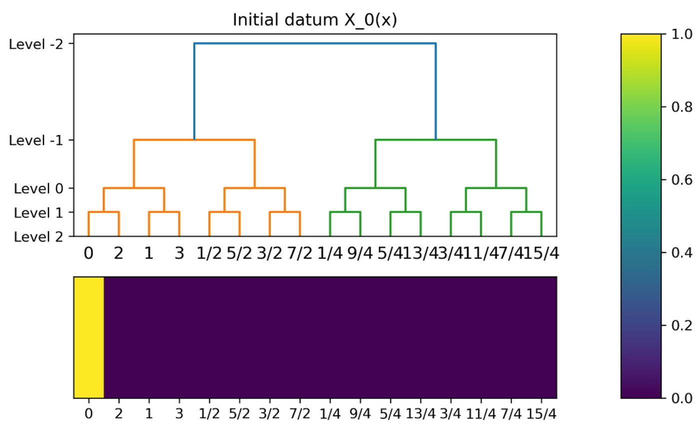

FIGURE 16. Simulation 4. Heat map of $X_0(x)$.

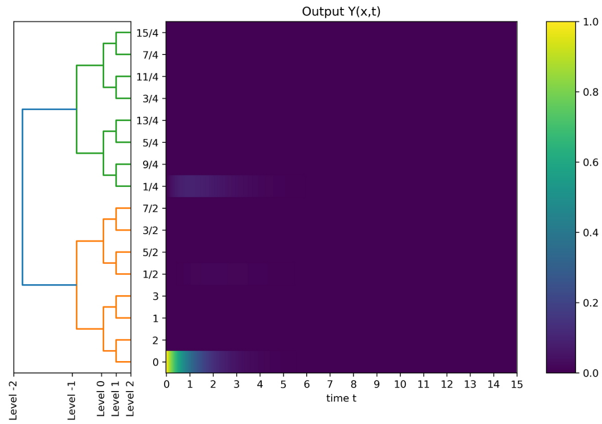

Figure 17. Simulation 4. Step 0.05.

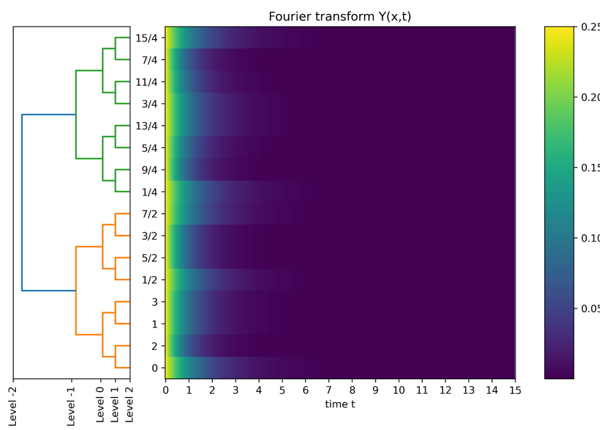

Figure 18. Simulation 4.

7. CONCLUSIONS

In this article, we present a p-adic generalization of Chua-Yang CNNs. In the p-adic framework, a continuous CNN is modeled by just one integro-differential equation depending on several p-adic variables and the time. In contrast, the classical CNNs are described by a discrete system of integro-differential equations. The need of constructing continuous models of discrete CNNs whose modeling requires millions of integro-differential equations is quite natural.

A one-dimensional p-adic continuous CNN has infinitely many cells which are hierarchically organized in rooted trees, also a such network has infinitely many hidden layers. The topology of the network, which lately controls the interaction of the cells, depends on the supports of the kernels of the feedback and feedforward operators. Under mild hypotheses, there is a natural discretization process of p-adic continuous CNNs that produces standard discrete CNNs. The solutions of the continuous CNNs can be very well approximated by the solutions of discrete

CNNs. Then, for practical purposes, a p-adic continuous CNN is a hierarchical discrete CNN with many hidden layers.

Our numerical simulations show that the solutions of continuous CNN exhibit a very complex behavior, including self-similarity and multistability, depending on the interaction of the parameters defining the network and the initial datum.

In the p-adic framework, the class of continuous CNNs is huge, for instance, consider equations of type

$$\frac{\partial X(x,t)}{\partial t}=-\boldsymbol{L}X(x,t)+\int\limits_{\mathbb{Q}_p^N}A(x,y)Y(y,t)d^Ny+\int\limits_{\mathbb{Q}_p^N}B(x,y)U(y)d^Ny+Z(x),$$

where $\frac{\partial X(x,t)}{\partial t}=-\boldsymbol{L}X(x,t)$ is a p-adic heat equation, i.e. the fundamental solution of a such equation is the transition probability density of a Markov process on $\mathbb{Q}_p^N$. The class of p-adic heat equations is extremely large, see e.g. [20], [31]. By incorporating a 'diffusion term' is natural to expect that the corresponding network will produce more complex patterns. We plan to study these networks in a forthcoming article. In the classical framework the reaction-diffusion CNNs have been studied intennsively, see e.g. [12, 13, 14].

References

[1] Albeverio S., Khrennikov A. Yu., Shelkovich V. M., Theory of p-adic distributions: linear and nonlinear models. London Mathematical Society Lecture Note Series, 370. Cambridge University Press, Cambridge, 2010.

[2] Albeverio, Sergio; Khrennikov, Andrei; Tirozzi, Brunello p-adic dynamical systems and neural networks, Math. Models Methods Appl. Sci. 9, no. 9, 1417–1437, 1999.

[3] Avetisov V. A., Bikulov A. Kh., Osipov V. A., p-adic description of characteristic relaxation in complex systems, J. Phys. A 36, no. 15, 4239–4246, 2003.

[4] Avetisov V. A., Bikulov A. H., Kozyrev S. V., Osipov V. A., p-adic models of ultrametric diffusion constrained by hierarchical energy landscapes, J. Phys. A 3 , no. 2, 177–189, 2002.

[5] Becker O. M., Karplus M., The topology of multidimensional protein energy surfaces: theory and application to peptide structure and kinetics, J. Chem.Phys. 106, 1495–1517, 1997.

[6] Chua Leon O., Yang Lin, Cellular neural networks: theory, IEEE Trans. Circuits and Systems 35, no. 10, 1257–1272, 1988.

[7] Chua Leon, Yang Lin, Cellular Neural Networks: Applications, IEEE Trans. on Circuits and Systems, 35, no. 10, 1273-1290, 1988.

[8] Chua Leon O, Roska, Tamas, Cellular neural networks and visual computing: foundations and applications. Cambridge university press, 2002.

[9] Chua L. O., CNN: A Paradigm for Complexity, World Scientific Series on Nonlinear Science (Series A), Vol. 31, Singapore: World Scientific Publishing Company, 1998.

[10] Dragovich B., Khrennikov A. Yu., Kozyrev S. V., Volovich, I. V., On p-adic mathematical physics, p-Adic Numbers Ultrametric Anal. Appl. 1, no. 1, 1–17, 2009.

[11] Frauenfelder H, Chan S. S., Chan W. S. (eds), The Physics of Proteins. Springer-Verlag, 2010.

[12] L. Goras, L. Chua, and D. Leenearts, Turing Patterns in CNNs – Part I: Once Over Lightly, IEEE Trans. on Circuits and Systems – I, 42, no. 10, 602-611, 1995.

[13] L. Goras, L. Chua, and D. Leenearts, Turing Patterns in CNNs – Part II: Equations and Behavior, IEEE Trans. on Circuits and Systems – I, 42, no. 10, 612-626, 1995.

[14] L. Goras, L. Chua, and D. Leenearts, Turing Patterns in CNNs – Part III: Computer Simulation Results, IEEE Trans. on Circuits and Systems – I, 42, no. 10, 627-637, 1995.

[15] Hua H., Hovestadt L., p-Adic numbers encode complex networks, Sci Rep 11, no. 17, 2021. https://doi.org/10.1038/s41598-020-79507-4.

[16] Neal Koblitz, *p-adic Numbers, p-adic Analysis, and Zeta-Functions*, Graduate Texts in Mathematics No. 58, Springer-Verlag, 1984.
[17] Kochubei Anatoly N., Pseudo-differential equations and stochastics over non-Archimedean fields. Marcel Dekker, Inc., New York, 2001.
[18] Khrennikov A, Information Dynamics in Cognitive, Psychological, Social and Anomalous Phenomena; Springer: Berlin/Heidelberg, Germany, 2004.
[19] Khrennikov, Andrei; Tirozzi, Brunello Learning of p-adic neural networks. Stochastic processes, physics and geometry: new interplays, II (Leipzig, 1999), 395–401, CMS Conf. Proc., 29, Amer. Math. Soc., Providence, RI, 2000.
[20] Khrennikov Andrei, Kozyrev Sergei, Zúñiga-Galindo W. A., Ultrametric Equations and its Applications. Encyclopedia of Mathematics and its Applications (168). Cambridge University Press, 2018.
[21] Kozyrev S. V., Methods and Applications of Ultrametric and p-Adic Analysis: From Wavelet Theory to Biophysics, Sovrem. Probl. Mat., 12, Steklov Math. Inst., RAS, Moscow, 2008, 3–168.
[22] Miklavčič Milan, Applied functional analysis and partial differential equations. World Scientific Publishing Co., Inc., River Edge, NJ, 1998.
[23] Nakao Hiroya and Mikhailov Alexander S., Turing patterns in network-organized activator-inhibitor systems, *Nature Physics* **6** (2010), 544-550.
[24] Rammal R., Toulouse G., Virasoro M. A., Ultrametricity for physicists, Rev. Modern Phys. 58 (1986), no. 3, 765–788.
[25] Slavova Angela, Cellular neural networks: dynamics and modelling. Mathematical Modelling: Theory and Applications, 16. Kluwer Academic Publishers, Dordrecht, 2003.
[26] Taibleson M. H., Fourier analysis on local fields. Princeton University Press, 1975.
[27] Vladimirov V. S., Volovich I. V., Zelenov E. I., p-adic analysis and mathematical physics. World Scientific, 1994.
[28] Zúñiga-Galindo, W. A., Reaction-diffusion equations on complex networks and Turing patterns, via p-adic analysis, J. Math. Anal. Appl. 491 (2020), no. 1, 124239, 39 pp.
[29] Zúñiga-Galindo, W. A., Non-archimedean replicator dynamics and Eigen's paradox. J. Phys. A 51 (2018), no. 50, 505601, 26 pp.
[30] Zúñiga-Galindo, W. A. Non-Archimedean reaction-ultradiffusion equations and complex hierarchic systems, Nonlinearity 31 (2018), no. 6, 2590–2616.
[31] Zúñiga-Galindo W. A., Pseudodifferential equations over non-Archimedean spaces. Lectures Notes in Mathematics 2174, Springer, 2016.

Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Departamento de Matemáticas, Av. Instituto Politécnico Nacional # 2508, Col. San Pedro Zacatenco, CDMX. CP 07360, México.

Email address: `bazambrano@math.cinvestav.mx`

University of Texas Rio Grande Valley, School of Mathematical & Statistical Sciences, One West University Blvd, Brownsville, TX 78520, United States and Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Departamento de Matemáticas, Unidad Querétaro, Libramiento Norponiente #2000, Fracc. Real de Juriquilla. Santiago de Querétaro, Qro. 76230, México.

Email address: `wilson.zunigagalindo@utrgv.edu, wazuniga@math.cinvestav.edu.mx`