Ludwig-Maximilians-Universität at München
Department "Institut für TODO"
Lehr- und Forschungseinheit TODO
Prof. Dr. PROFESSORS NAME

Bachelor's Thesis

THESIS TITLE

Valentin Herrmann
EMAIL

Bearbeitungszeitraum:	START bis END
Betreuer:	SUPERVISOR
Verantw. Hochschullehrer:	Prof. Dr. PROFESSORS NAME

Aufgabenstellung

THESIS TITLE

Problem Statement TO BE ADDED

Scope of the Thesis TO BE ADDED

Tasks TO BE ADDED

Requirements TO BE ADDED

Keywords TO BE ADDED

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig angefertigt, alle Zitate als solche kenntlich gemacht sowie alle benutzten Quellen und Hilfsmittel angegeben habe.

München, August 21, 2025 ....

Acknowledgments

I would like to appreciate ...

Abstract

This thesis proposes ...

Contents

1 Introduction

While the hype on AI is still ongoing there are still people wondering if current AI models are fundamentally able of reasoning (TODO: Apple Paper). Many other possible models could be better fitted to the task of reasoning. Among others Spiking neural networks present a model closer to the workings of the brain. In fact, they have been called the 3rd generation of AI models, after the 2nd generation that currently drives most successful models.

While the idea behind SNNs is quite old, they have not been as much researched, since it is much more inefficient and harder to train them. Therefore there still remain a lot of open questions about them. In this paper we shall extend on the work done in [Nguyen et al., 2025]. We will add a decaying factor to the input of the neurons and allow recursive connections between neurons in a layer.

We will roughly follow the structure of [Nguyen et al., 2025]. In the second chapter we will formally introduce discrete time leaky-integrate-and-fire SNN, d.t. LIF-SNNs. In section 3 we will give theorems about approximation of continuous functions on compact domains by d.t. LIF-SNNs. The main part will be the following section in which we will see that the number of distinct values a d.t. LIF-SNN can take on only depends on the first hidden layer and grows in particular only quadratically in time. We will further support our findings with experimental data in the last section.

2 Definitions

Our type of SNN should be thought of as a composition of an initial input layer, a number of hidden spiking layers with internal state and an affine-linear layer mapping spikes activations over time, so called spike trains, to the value of the output layer. The following definitions deviate slightly from [Nguyen et al., 2025], since we already bake in some of the assumptions (direct encoding and membrane-potential outputs) the paper makes at a later point. We shall first define the membrane potential $u^l(t)$ and the spike activations $s^l(t)$ of the hidden layers:

Definition 2.1. *The **input vector** $i^l(t) \in \{0,1\}^{n_l}$, the **spike vector** $s^l(t) \in \{0,1\}^{n_l}$ and the **membrane potential vector** $u^l(t)$ of a hidden layer $l \in [L]$ are recursively defined as*

$$i^l(t) = \alpha^l i^l(t-1) + W^l s^{l-1}(t) + V^l s^l(t-1) \tag{1}$$
$$p^l(t) = \beta^l u^l(t-1) + i^l(t) + b^l \tag{2}$$
$$s^l(t) = H(p^l(t) - \vartheta 1_{n_l}) \tag{3}$$
$$u^l(t) = p^l(t) - \vartheta s^l(t) \tag{4}$$

with $s^l(0) = 0$ and given

- ***first layer spike activations**: $(s^0(t))_{t\in[T]} \in \{0,1\}^{n_0 \times T}$*
- ***initial membrane potential**: $u^l(0) \in \mathbb{R}^{n_l}$*
- ***initial input**: $i^l(0) \in \mathbb{R}^{n_l}$*
- ***weight matrices**: $W^l \in \mathbb{R}^{n_l \times n_{l-1}}$*
- ***bias vectors**: $b^l \in \mathbb{R}^{n_l}$*
- ***leaky terms**: $\alpha^l, \beta^l \in [0,1]$*
- ***threshold**: $\vartheta^l \in (0,\infty)$*

where $H := \mathbb{1}_{[0,\infty)}$ is a step function and $T \in \mathbb{N}$ is the number of simulated time steps.

In the definition of $s^l(t)$ we first check whether the additional activation by spikes $W^l s^{l-1}(t)$ of the previous layers and the bias b^l plus the decayed previous activation $\beta^l u^l(t-1)$ passes over a certain threshold $\vartheta^l 1_{n_l}$. If that is not the case we use the just computed value for $u^l(t)$. If the neuron activates, we remove a threshold worth of value. Of course, we do this for the neurons of the whole layer in parallel.

We further define d.t. LIF-SNN and the function the network realizes:

Definition 2.2. *A **discrete-time LIF-SNN** of **depth** L with **layer-widths** $(n_0,\dots,n_{L+1})$ is given by*

$$\Phi := ((W^l, b^l, u^l(0), i^l(0), \alpha^l, \beta^l, \vartheta^l)_{l\in[L]}, T, ((a_t)_{t\in[T]}, c, V)$$

where $(a_t)_{t\in[T]} \in \mathbb{R}^T$, $c \in \mathbb{R}^{n_{L+1}}$ and $V \in \mathbb{R}^{n_{L+1}\times n_L}$ are the parameters of the output layer.

Definition 2.3. *A discrete-time LIF-SNN ϕ **realizes** the function $R(\Phi): \mathbb{R}^{n_0} \to \mathbb{R}^{n_{L+1}}$:*

$$R(\Phi)(x) = D((s^L(t))_{t\in[T]}) \quad \text{with } s^0(t) := x$$

where the last layer D is defined by $D((s(t))_{t\in[T]}) = \sum_{t=1}^T a_t(Vs(t)+c)$.

Let us now take a look at some simple examples:

Example 2.1. *Let $T, L \in \mathbb{N}$. Then there exists a d.t. LIF-SNN with $\forall_{t\in[T]} s^L(t) = s^0(t)$ for any $s^0 \in \{0,1\}^{n_0 \times T}$.*

We can proof this by using constant width $n_l = n$, weights $W^l = I_n$, biases $b^l = 0$, initial membrane potential $u^l(0) = 0$, leaky term $\beta^l = 0$ and threshold $\vartheta^l = 1$ for all $l \in [L]$.

We then get by definition:

$$\begin{aligned} s^l(t) &= H(\, s^{l-1}(t) - 1_{n_l}) = \; s^{l-1}(t) \\ u^l(t) &= \quad s^{l-1}(t) - s^l(t) = 0 \end{aligned}$$

Example 2.2. *Let $T, L \in \mathbb{N}$. Then there exists a d.t. LIF-SNN with $\forall_{t\in T} s^L(t) = \max_{t'\in[t-1]}(s^0(t'))$ for any $s^0 \in \{0,1\}^{n_0 \times T}$: In this network an output neuron switches on when the corresponding input neurons fires and does not switch off later.*

We can proof this by using constant width $n_l = n$, weights $W^l = T \cdot I_n$, biases $b^l = 0$, initial membrane potential $u^l(0) = 0$, leaky term $\beta^l = 1$ and threshold $\vartheta^l = 1$ for all $l \in [L]$.

We then get by definition:

$$\begin{aligned} s^l(t) &= H(\, u^l(t-1) + T \cdot s^{l-1}(t) - 1_{n_l}) \\ u^l(t) &= \quad u^l(t-1) + T \cdot s^{l-1}(t) - s^l(t) \end{aligned}$$

By adding over all timesteps we obtain

$$\begin{aligned} s^l(t) &= H(T \sum_{i=1}^{t} s^{l-1}(i) - (\sum_{t=i}^{t-1} s^l(i)) + 1_{n_l})) \\ u^l(t) &= \sum_{i=1}^{t} (T \cdot s^{l-1}(i) - s^l(i)) \end{aligned}$$

Since $\sum_{i=1}^{t-1} s^l(i) + 1_{n_l} \leq T$, once there is any $t_0 \in [T]$ with $s_i^{l-1}(t_0) = 1$ for an i, we get $\forall_{t \geq t_0} s_i^l(t) = 1$. By induction over the layers we clearly get the required property.

3 Structure of computations in d.t. LIF-SNNs

This section concern itself with the approximation of continuous functions by d.t. LIF-SNN.

In [Nguyen et al., 2025] the following theorem was proved:

Theorem 3.1. *Let f be a continuous function on a compact set $\Omega \subset \mathbb{R}^{n_0}$. For all $\varepsilon > 0$, there exists a d.t. LIF-SNN Φ with direct encoding, membrane potential output, $L = 2$ and $T = 1$ such that*

$$\|R(\Phi) - f\|_\infty \leq \varepsilon$$

Moreover, if f is Γ-Lipschitz, then Φ can be chosen with width parameter $n = (n_1, n_2)$ given by

$$\begin{aligned} n_1 &= \left(\max\left\{\left\lceil \frac{\operatorname{diam}_\infty(\Omega)}{\varepsilon}\Gamma \right\rceil, 1\right\} + 1\right) n_0 \\ n_2 &= \max\left\{\left\lceil \frac{\operatorname{diam}_\infty(\Omega)}{\varepsilon}\Gamma \right\rceil^{n_0}, 1\right\} \end{aligned}$$

where $\operatorname{diam}_\infty(\Omega) = \sup_{x,y\in\Omega} \|x - y\|_\infty$.

Proof. See [Nguyen et al., 2025]. □

We will now provide an alternative version of this theorem which uses an increased latency to reduce the number of neurons:

Theorem 3.2. *Let $f \in \mathcal{C}^1(\Omega, \mathbb{R}^m)$ be a continuously differentiable function on a compact set $\Omega \subset \mathbb{R}^n$. For all $\varepsilon > 0$, there exists a d.t. LIF-SNN Φ with direct encoding, membrane potential output, $L = 2$ and*

$$\begin{aligned} T &= TODO \\ n_1 &= (K-1) \cdot n + n \\ n_2 &= K^n \cdot m \end{aligned}$$

such that

$$\|R(\Phi) - f\|_\infty \leq \varepsilon$$

Lemma 3.1. *Let $\Omega \subset \mathbb{R}^n$ be compact. Then there exists a half-open cube C with width $\operatorname{diam}_\infty(\Omega)$ such that $\Omega \subset \overline{C}$.*

Proof of Lemma 3.1. We can first define $a := (\min_{x\in\Omega} x_i)_i$ since Ω is compact We further have $b := a + \operatorname{diam}_\infty(\Omega) \cdot \mathbb{1}_n$. We now have $C := [a, b)$. Suppose now a point $y \in \Omega \setminus \overline{C}$ exists. By definition of a, we have $a \leq y$. By definition of C we further get $y \not\leq b$, and therefore $\exists_i b_i < y_i$ by definition of C. But this means $\|a - y\|_\infty > \operatorname{diam}_\infty(\Omega)$ □

Lemma 3.2. *Let $f \in \mathcal{C}^1(\Omega, \mathbb{R}^m)$ be a continuously differentiable function on a compact set $\Omega \subset \mathbb{R}^n$ and $\varepsilon > 0$. Then there exists a half-open cube C with $\Omega \subset C$ that can be composed in*

$$K^n := \min_{\substack{\xi,\theta>0 \\ \xi\theta=\varepsilon}} \left\{ \left\lceil \frac{\operatorname{diam}_\infty(\Omega)}{\frac{2}{\sqrt{n}} \min(\delta(\xi), \theta)} \right\rceil \right\}^n$$

half-open subcubes $(C_i)_{i=1..K^n}$ such that $\Omega \subset C$ and linear functions $g_i : C_i \to \mathbb{R}^m$ exists, such that $\|f - g|_\Omega\|_\infty < \varepsilon$ for $g := \sum_{i=1}^m g_i \mathbb{1}_{C_i}$. Here $\delta(\varepsilon)$ is the modulus of uniform continuity of the total derivative dF.

Proof of Lemma 3.2. By Lemma 3.1 we have a half-open cube C with width $\operatorname{diam}_\infty(\Omega)$ and $\Omega \subset \overline{C}$. Let further $\varepsilon > 0$ be given. Since Ω is compact and $f \in \mathcal{C}^1(\Omega, \mathbb{R}^m)$, dF is uniformly continuous on Ω. Let $\delta(\varepsilon)$ be dFs modulus of continuity. We will now partition C in

$$K^n := \min_{\substack{\xi,\theta>0\\ \xi\theta=\varepsilon}} \left\{ \left\lceil \frac{\operatorname{diam}_\infty(\Omega)}{\frac{2}{\sqrt{n}}\min(\delta(\xi),\theta)} \right\rceil \right\}^n$$

smaller half-open cubes. There are ξ, θ such that the minimum in the definition of K^n is obtained, since we take it over the set of natural numbers. The subcubes have width $w := \frac{\operatorname{diam}_\infty(\Omega)}{K} \leq \frac{2}{\sqrt{n}}\min(\delta(\xi),\theta)$. Let us further define $g_i : C_i \to \mathbb{R}^m$ by $g_i(x) := f(c_i) + df_{c_i}(x - c_i)$ where c_i is the center of C_i. It suffices now to show $\left\| f|_{C_i} - g_i \right\|_\infty < \varepsilon$. Let $x \in C_i$. We get:

$$\begin{aligned}
\left\| f(x) - f(c_i) - df_{c_i}(x-c_i) \right\|_p &= \left\| f(x) - f(c_i) - df_{c_i}(x-c_i) \right\|_p \\
&= \left\| \int_0^1 df_{c_i+(x-c_i)t}(x-c_i)dt - df_{c_i}(x-c_i) \right\|_p \\
&\leq \int_0^1 \left\| df_{c_i+(x-c_i)t}(x-c_i) - df_{c_i}(x-c_i) \right\|_p dt \\
&= \int_0^1 \left\| (df_{c_i+(x-c_i)t} - df_{c_i})(x-c_i) \right\|_p dt \\
&\leq \int_0^1 \left\| (df_{c_i+(x-c_i)t} - df_{c_i}) \right\| \left\| x-c_i \right\|_p dt \\
&\leq \int_0^1 \xi \left\| x-c_i \right\|_p dt \\
&= \xi \left\| x-c_i \right\|_p \\
&\leq \xi\theta \\
&= \varepsilon
\end{aligned}$$

□

Proof of Theorem 3.2. Let a continuously differentiable function $f \in \mathcal{C}^1(\Omega, \mathbb{R}^m)$ on a compact set $\Omega \subset \mathbb{R}^n$ be given. Let there further be a $\varepsilon > 0$. By Lemma 3.2 we have a half-open cube C with $\Omega \subset C$ with composition K^n half-open subcubes $(C_i)_i = 1..K^n$ and linear functions $g_i : C_i \to \mathbb{R}^m$, such that $\left\| f - g|_\Omega \right\|_\infty < \varepsilon\tau$ for a $g := \sum_{i=1}^m g_i \mathbb{1}_{C_i}$. We will now define a d.t. LIF-SNNΦ with direct input encoding and membrane-potential outputs such that $\left\| R(\Phi)|_\Omega - g \right\|_\infty < \varepsilon(1-\tau)$. Before anything else we shall set the following basic parameters: $u^l(0) = i^l(0) = 0$, $\alpha^l = 0$ and $\beta^l = \vartheta^l = 1$. We will now define the first layer in a similar fashion to [Nguyen et al., 2025] in the proof of Theorem 3.1. Let us choose $K-1$ neurons in the first hidden layer for each dimension representing hyperplanes separating the subcubes in the input space. □

comparison: sin mit kleiner ampitude, extrem hoher frequenz

The proof works by first showing that a continuous function can be arbitrarily approximated by step functions, in particular by step functions constant on hypercubes in Ω. Then the d.t. LIF-SNN is constructed by using the first layer to cut the input space by hyperplanes along the cubes and using the second layer to represent the hypercubes.

Figure 3.1 shows a sinus function getting approximated in this way.

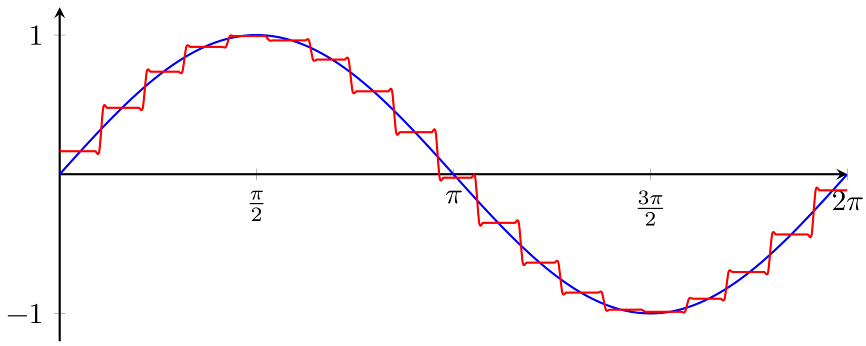

Figure 3.1: A d.t. LIF-SNNapproximating a sinus wave

4 Complexity of input partitions

In the following we will analyze how many different values a d.t. LIF-SNN with recurrent edges dependent on parameters can obtain. Since the first hidden layer only communicates by spike trains to the following layers, the maximal number of different output values is bound by the number of different spike trains the first hidden layer can emit. We therefore simplify the notation by writing $n, W, b, V, \beta, \vartheta, u, s$ for $n_1, W^{(1)}, b^{(1)}, V^{(1)}, \beta^{(1)}, \vartheta^{(1)}, u^{(1)}, s^{(1)}$ respectively. Further we can assume $n_0 = n_1 = n$, $W = I_{n_1}$ and $b = 0_{n_1}$, since we can instead just precompose our SNN Φ with $x \mapsto Wx + b$. Since precomposition can only decrease the number of different values, we can instead just study Φ by itself. We have now

$$u(t;x) = \beta u(t-1;x) + x + (V - \vartheta I_n)s(t-1;x)$$
$$s(t;x) = H(u(t;x) - \vartheta \mathbb{1}_n)$$

By repeatedly substituting $u(t-1;x)$ we get

$$u(t;x) = \beta^t u(0) + (\sum_{i=0}^{t-1} \beta^i)x + (V - \vartheta \cdot \mathbb{1}_n) \sum_{i=1}^{t-1} \beta^i s(i;x)$$

Lemma 4.1. *The constant regions of a d.t. LIF-SNN with recurrent edges with $W = I_{n_1}$, $b = 0_{n_1}$ are half-open cuboids.*

Proof. Let $x, y \in \mathbb{R}^n$. We will first proof that the constant regions are convex. Assume that $s(\cdot;x)$ and $s(\cdot;y)$ are equal. Let us assume that there is a $z = x + (y-x)\tau$ with $\tau \in [0,1]$ and $s(\cdot, z) \neq s(\cdot;x)$. In that case there exists a minimal $t \in [T]$ with $\exists_i s_i(t;z) \neq s_i(t;x)$. Let us now regard any i with $s_i(t;z) \neq s_i(t;x)$. We can assume w.l.o.g. that $x_i \leq y_i$. Due to minimality of t we have $\forall_{t' \in [t-1]} s(t';z) = s(t';x) = s(t';y)$. Since further $\beta \geq 0$ we get that $u_i(t;x)$ is monotone in x_i. Since further H is monotone, we conclude

$$s_i(t;x) \leq s_i(t;z) \leq s_i(t;y)$$

Since $s_i(t;x) = s_i(t;y)$, we get $s_i(t;x) = s_i(t;z) = s_i(t;y)$.

Let us now consider a maximal constant region $C \subset \mathbb{R}^n$ with spiketrain. We now get $C = \cap_{t \in [T], i \in [n]} \{x \mid s_i(t;x) =\}$ We will now proof that borders between constant regions are along coordinate hyperplanes. Let us yet again consider $x \in \mathbb{R}^n$, but this time with $s(\cdot, x) \neq s(\cdot, y)$. Let us assume there is $z = x + (y-x)\tau_0$ with $\tau_0 \in [0,1]$ minimal so that $s(\cdot, x) \neq s(\cdot, z)$. Let $t \in [T]$ be yet again minimal with $\exists_i s_i(t;x) \neq s_i(t;y)$. $\square$

Lemma 4.2. *For $y \in \mathbb{R}^{n_0}$ the mapping $f_y : \mathbb{R} \to \{0,1\}^T$ defined by $x \mapsto s_i$ with $s^{(0)}(t) = y + xe_i$ is monotone regarding lexical ordering.*

Theorem 4.1. *For $\beta = 1$ a d.t. LIF-SNN with recurrent edges can obtain $O(T^{2n})$ different values.*

Bibliography

References

[Nguyen et al., 2025] Nguyen, D. A., Araya, E., Fono, A., and Kutyniok, G. (2025). Time to spike? understanding the representational power of spiking neural networks in discrete time.