



**HAL**  
open science

# Local Identifiability of Deep ReLU Neural Networks: the Theory

Joachim Bona-Pellissier, François Malgouyres, François Bachoc

► **To cite this version:**

Joachim Bona-Pellissier, François Malgouyres, François Bachoc. Local Identifiability of Deep ReLU Neural Networks: the Theory. 2022. hal-03687395v1

**HAL Id: hal-03687395**

**<https://hal.science/hal-03687395v1>**

Preprint submitted on 14 Jun 2022 (v1), last revised 21 Nov 2022 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

---

# Local Identifiability of Deep ReLU Neural Networks: the Theory

---

Joachim Bona-Pellissier<sup>a\*</sup>, François Malgouyres<sup>b</sup>, François Bachoc<sup>b</sup>

Institut de Mathématiques de Toulouse ; UMR 5219

Université de Toulouse ; CNRS

<sup>a</sup> UT1, F-31042 Toulouse, France

<sup>b</sup> UPS, F-31062 Toulouse Cedex 9, France

\*Corresponding author: Joachim Bona-Pellissier; joachim.bona-pellissier@univ-toulouse.fr

## Abstract

Is a sample rich enough to determine, at least locally, the parameters of a neural network? To answer this question, we introduce a new local parameterization of a given deep ReLU neural network by fixing the values of some of its weights. This allows us to define local lifting operators whose inverses are charts of a smooth manifold of a high dimensional space. The function implemented by the deep ReLU neural network composes the local lifting with a linear operator which depends on the sample. We derive from this convenient representation a geometrical necessary and sufficient condition of local identifiability. Looking at tangent spaces, the geometrical condition provides: 1/ a sharp and testable necessary condition of identifiability and 2/ a sharp and testable sufficient condition of local identifiability. The validity of the conditions can be tested numerically using backpropagation and matrix rank computations.

## 1 Introduction

### 1.1 Context and motivations

Neural networks are famous for their capacity to perform complex tasks in a wide variety of domains such as image classification [18], object recognition [31, 32], speech recognition [15, 34, 14], natural language processing [25, 24, 17], anomaly detection [30] or climate sciences [1].

A question that has recently drawn attention is the question of the identifiability of the parameters of neural networks. This question can be described as follows: for a given architecture and some given inputs  $x^i$ , do the responses  $f_\theta(x^i)$  of the network to these inputs uniquely characterize the weight and bias parameters  $\theta$ , up to neuron permutation and positive rescaling of the weights and biases? It is indeed well known, for ReLU networks, that the latter operations on  $\theta$  do not change the function  $f_\theta$  [29]. It is therefore impossible, knowing only the  $f_\theta(x^i)$ , to distinguish the elements within the equivalence class of  $\theta$  modulo these operations. Questions that are naturally related to identifiability, and that we do not address in this article, are inverse stability -is the characterization stable to small perturbation?- and stable recovery -are we able to stably recover  $\theta$  (up to equivalence) in practice?

Identifiability is important for different reasons. In the first place, model extraction attacks for neural networks have been a growing topic over the last years. Indeed, some algorithms are able to recover in practice the parameters of a neural network from queries [7, 33]. This can be a concern since neural network providers may wish to keep these parameters secret, for security [19], for privacy [11, 6], or for intellectual property [40].

A way of preventing such a recovery can be by guaranteeing that identifiability does not hold, that is, for a list of requests  $X$ , guaranteeing that  $\theta$  is not uniquely characterized by the answers  $f_\theta(X)$  to these requests. To do so, one needs to check that a necessary condition of identifiability is not met. On the opposite side, guaranteeing that identifiability holds is interesting in the position of an attacker. If the attacker has access to  $X$ , to  $f_\theta(X)$ , and is able to compute a  $\tilde{\theta}$  such that  $f_{\tilde{\theta}}(X) = f_\theta(X)$ , the question then becomes: does this guarantee that  $\tilde{\theta} \sim \theta$  or shall the attacker expand  $X$  with new queries? The attacker needs a sufficient condition of identifiability.

In these examples  $X$  represents requests to an already trained network. We can think of another context in which  $X$  represents a training database or a test set. In the former case, optimizing the empirical risk may be difficult in general and the model may generalize poorly. However, if successful, that is if the trained network matches the samples, a sufficient condition of identifiability can guarantee that the function implemented by the network only depends on its values at these samples. In particular, it does not depend on the choice of the optimizer, on its initialization or on stochastic parameters.

## 1.2 Existing work on identifiability, inverse stability and stable recovery

Even though it has regained interest recently, the question of identifiability for neural networks is not new. Indeed, in the 1990s, interesting results on identifiability of networks with smooth activation functions (tanh, logistic sigmoid, Gaussian...) have been established [38, 2, 20, 16, 10].

When it comes to shallow [28, 36] as well as deep [29, 4] ReLU neural networks, some results have been recently established. They show that under some conditions, the function implemented by the network uniquely characterizes its parameters, up to neuron permutation and rescaling operations.

All these results assume the function implemented by the network to be known on the whole input space, or at least on an open subset of it. As far as we know, there exists only one identifiability result for deep ReLU networks assuming the knowledge of this function on a *finite* sample. Stock and Gribonval [37] give a theoretical condition for the existence of a finite set which locally identifies the parameters of a deep neural network. The construction in [37] shares similarities with previous works on deep structured matrix factorization [22]. The present article lies in this line of research.

Closely related to identifiability are the topics of inverse stability and stable recovery of the parameters of a network. Some negative [27] as well as positive [9, 21, 22, 23] results of inverse stability exist for ReLU and identity activation functions. Several stable recovery algorithms have also been proposed, for shallow networks in a first place, for smooth activation function [12], as well as ReLU in the fully-connected case [13, 42, 43, 44] or in the convolutional case [5, 41]. These references provide a large sample complexity under which minimizing the empirical risk allows to recover the parameters of the network.

For deep networks, some stable recovery algorithms exist, for instance for Heavyside activation function [3], or for the first layer with sparsity assumptions [35] in the ReLU case. For deep ReLU networks, when one has full access to the function implemented by the network, a practical algorithm [33] sequentially constructs a sample and approximately recovers the architecture and the parameters modulo permutation and rescaling. Similarly, formulating the problem as a cryptanalytic problem, [7] reconstructs a functionally equivalent network with fewer requests.

## 1.3 Contributions

**1/** We establish a necessary and sufficient geometrical condition of local identifiability from a finite sample  $X$  for deep fully-connected ReLU networks. The condition is that the intersection between a smooth manifold and an affine space is reduced to a single point. **2/** Considering tangent spaces, we then provide a computable necessary condition of local identifiability which, since global identifiability implies local identifiability, is also a computable necessary condition of identifiability. **3/** We also establish a computable sufficient condition of local identifiability, which is close to the necessary condition. To the best of our knowledge, these are the first testable conditions of local identifiability for any finite input sample. In particular, [37] provides a theoretical condition equivalent to the existence of a finite sample for which local identifiability holds. The existence of the sample is proved in a constructive manner. The authors do not characterize local identifiability for any given sample.

4/ To prove these results, we develop geometrical tools which can be of independent interest for theoretically understanding deep ReLU networks as well as for possible applications. Namely, we introduce local reparameterizations  $\rho_\theta$  of the network by fixing some weight values as constants. Building on these local parameterizations, we introduce local lifting operators  $\psi^\theta$  and we decompose the function implemented by the network  $f_\theta(x)$  as a composition of  $\psi^\theta$ , which only depends on the parameters, and a piecewise constant operator  $\alpha$  which depends on  $\theta$  and the inputs  $x^i$ . For almost any parameterization  $\theta$ , the operator  $\alpha$  is constant in a neighborhood of  $\theta$  and consists in applying a linear function to  $\psi^\theta$ . We show that in fact, the operators  $\psi^\theta$  are the inverses of coordinate charts of a smooth manifold  $\Sigma_1^*$ , contained in a high dimensional space. We find  $\Sigma_1^*$  to be of particular interest in representing geometrically some properties of the network parameters (in particular to establish 1/, 2/ and 3/ above).

## 1.4 Overview of the article

This work is structured as follows. We start by introducing basic tools and already known results in Section 2. We then introduce the local parameterizations  $\rho_\theta$  and the set  $\Sigma_1^*$ , and we show that it is a smooth manifold in Section 3. This allows us to state our main results in Section 4, that is the geometrical and the numerically testable conditions of local identifiability. Finally we discuss in Section 5 the numerical computations needed to test the latter conditions. All the proofs are provided in the supplementary material.

## 2 ReLU networks, lifting operator and rescaling of the parameters

### 2.1 ReLU networks

Let us introduce our notations for deep fully-connected ReLU networks. In this paper, a network is a graph  $(E, V)$  of the following form.

- $V$  is a set of neurons, which is divided in  $L + 1$  layers, with  $L \geq 2$ :  $V = (V_l)_{l \in \llbracket 0, L \rrbracket}$ .  $V_0$  is the input layer,  $V_L$  the output layer and the layers  $V_l$  with  $1 \leq l \leq L - 1$  are the hidden layers. Using the notation  $|C|$  for the cardinal of a finite set  $C$ , we denote, for all  $l \in \llbracket 0, L \rrbracket$ ,  $N_l = |V_l|$  the size of the layer  $V_l$ .
- $E$  is the set of all oriented edges  $v \rightarrow v'$  between neurons in consecutive layers, that is

$$E = \{v \rightarrow v', v \in V_l, v' \in V_{l+1}, \text{ for } l \in \llbracket 0, L - 1 \rrbracket\}.$$

A network is parameterized by weights and biases, gathered in its parameterization  $\theta$ , with

$$\theta = ((w_{v \rightarrow v'})_{v \rightarrow v' \in E}, (b_v)_{v \in B}) \in \mathbb{R}^E \times \mathbb{R}^B,$$

where  $B = \bigcup_{l=1}^L V_l$ . It is also convenient to consider the weights and biases in matrix/vector form: for a given  $\theta$ , we denote, for  $l \in \llbracket 1, L \rrbracket$ ,

$$W_l = (w_{v \rightarrow v'})_{v' \in V_l, v \in V_{l-1}} \in \mathbb{R}^{N_l \times N_{l-1}} \quad \text{and} \quad b_l = (b_v)_{v \in V_l} \in \mathbb{R}^{N_l}.$$

When dealing with two parameterizations  $\theta$  and  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ , we take as a convention that  $w_{v \rightarrow v'}$  and  $b_v$  as well as  $W_l$  and  $b_l$  denote the weights and biases associated to  $\theta$ , and  $\tilde{w}_{v \rightarrow v'}$  and  $\tilde{b}_v$  as well as  $\tilde{W}_l$  and  $\tilde{b}_l$  denote those associated to  $\tilde{\theta}$ .

The activation function, denoted  $\sigma$ , is always ReLU: for any  $p \in \mathbb{N}^*$  and any vector  $x = (x_1, \dots, x_p)^T \in \mathbb{R}^p$ , it is defined as  $\sigma(x) = (\max(x_1, 0), \dots, \max(x_p, 0))^T$ .

For a given  $\theta$ , we define recursively  $f_l : \mathbb{R}^{V_0} \rightarrow \mathbb{R}^{V_l}$  (we omit the dependency in  $\theta$  in the notation for simplicity), for  $l \in \llbracket 0, L \rrbracket$ , by

- $\forall x \in \mathbb{R}^{V_0}, \quad f_0(x) = x$ ;
- $\forall l \in \llbracket 1, L - 1 \rrbracket, \forall x \in \mathbb{R}^{V_0}, \quad f_l(x) = \sigma(W_l f_{l-1}(x) + b_l)$ ;
- $\forall x \in \mathbb{R}^{V_0}, \quad f_L(x) = W_L f_{L-1}(x) + b_L$ .

We define  $f_\theta : \mathbb{R}^{V_0} \rightarrow \mathbb{R}^{V_L}$  as  $f_\theta = f_L$  and we refer to it as the function implemented by the network of parameter  $\theta$ .

## 2.2 The lifting operator $\phi$

For all  $l \in \llbracket 1, L-1 \rrbracket$ , for all  $v \in V_l$ , we denote, for all  $\theta \in \mathbb{R}^E \times \mathbb{R}^B$  and  $x \in \mathbb{R}^{V_0}$ ,

$$a_v(x, \theta) = \begin{cases} 1 & \text{if } (W_l f_{l-1}(x) + b_l)_v \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

which is the activation indicator of neuron  $v$ . For all  $l \in \llbracket 0, L-1 \rrbracket$ , we define

$$\mathcal{P}_l = V_l \times \cdots \times V_{L-1},$$

which is the set of all paths in the network starting from layer  $l$  and ending in layer  $L-1$ . We consider an additional element  $\beta$  which can be interpreted as an empty path and whose role will be clear once  $\phi$  has been defined and Proposition 1 stated. We define

$$\mathcal{P} = \left( \bigcup_{l=0}^{L-1} \mathcal{P}_l \right) \cup \{\beta\}.$$

In a similar way to [37], we define a ‘lifting operator’

$$\begin{aligned} \phi : \mathbb{R}^E \times \mathbb{R}^B &\longrightarrow \mathbb{R}^{\mathcal{P} \times V_L} \\ \theta &\longmapsto (\phi_{p,v}(\theta))_{p \in \mathcal{P}, v \in V_L} \end{aligned} \quad (1)$$

by:

- for all  $l \in \llbracket 0, L-1 \rrbracket$  and all  $p = (v_l, \dots, v_{L-1}) \in \mathcal{P}_l$ , and for all  $v_L \in V_L$ ,

$$\phi_{p,v_L}(\theta) = \begin{cases} \prod_{l'=0}^{L-1} w_{v_{l'} \rightarrow v_{l'+1}} & \text{if } l = 0 \\ b_{v_l} \prod_{l'=l}^{L-1} w_{v_{l'} \rightarrow v_{l'+1}} & \text{if } l \geq 1; \end{cases}$$

- for  $p = \beta$  and  $v_L \in V_L$ ,  $\phi_{\beta,v_L}(\theta) = b_{v_L}$ .

We now define the ‘activation operator’

$$\begin{aligned} \alpha : \mathbb{R}^{V_0} \times (\mathbb{R}^E \times \mathbb{R}^B) &\longrightarrow \mathbb{R}^{1 \times \mathcal{P}} \\ (x, \theta) &\longmapsto (\alpha_p(x, \theta))_{p \in \mathcal{P}} \end{aligned} \quad (2)$$

by:

- for all  $l \in \llbracket 0, L-1 \rrbracket$  and all  $p = (v_l, \dots, v_{L-1}) \in \mathcal{P}_l$ :

$$\alpha_p(x, \theta) = \begin{cases} x_{v_0} \prod_{l'=1}^{L-1} a_{v_{l'}}(x, \theta) & \text{if } l = 0 \\ \prod_{l'=l}^{L-1} a_{v_{l'}}(x, \theta) & \text{if } l \geq 1; \end{cases}$$

- for  $p = \beta$ ,  $\alpha_\beta(x, \theta) = 1$ .

We have the following decomposition of the function  $f_\theta$  implemented by the network.

**Proposition 1.** For all  $\theta \in \mathbb{R}^E \times \mathbb{R}^B$  and all  $x \in \mathbb{R}^{V_0}$ ,

$$f_\theta(x)^T = \alpha(x, \theta) \phi(\theta).$$

This result, which is proven in the supplement, is also stated in [37, Sec. 4] with slightly different notations. To describe the interest of Proposition 1, let us anticipate on Proposition 2, which states that  $\theta \mapsto \alpha(x, \theta)$  is piecewise constant. When  $x$  is fixed, on a piece where  $\alpha(x, \theta)$  is constant and therefore independent of  $\theta$ , Proposition 1 expresses the map  $\theta \mapsto f_\theta(x)$  as the composition of a fixed linear operator and a polynomial lifting operator. This allows to decompose the complex map  $\theta \mapsto f_\theta(x)$  in two simpler ‘bricks’.

Let us reformulate Proposition 1 with several inputs. We consider, for some  $n \in \mathbb{N}^*$ , some given inputs  $x^i \in \mathbb{R}^{V_0}$ , with  $i \in \llbracket 1, n \rrbracket$ . We denote by  $X \in \mathbb{R}^{n \times V_0}$  the matrix whose lines are the transpose  $(x^i)^T$  of the inputs. For all  $\theta \in \mathbb{R}^E \times \mathbb{R}^B$ , we denote by  $f_\theta(X) \in \mathbb{R}^{n \times V_L}$  the matrix whose lines are the transpose  $f_\theta(x^i)^T$  of the outputs, for all  $i \in \llbracket 1, n \rrbracket$ . We also denote by  $\alpha(X, \theta) \in \mathbb{R}^{n \times \mathcal{P}}$  the matrix whose lines are the line vectors  $\alpha(x^i, \theta)$ , for all  $i \in \llbracket 1, n \rrbracket$ . Using Proposition 1 for all the  $x^i$ , we have the relation

$$f_\theta(X) = \alpha(X, \theta) \phi(\theta). \quad (3)$$

We prove in the supplement the following proposition.

**Proposition 2.** For all  $n \in \mathbb{N}^*$ , for all  $X \in \mathbb{R}^{n \times V_0}$ , the mapping

$$\begin{aligned} \alpha_X : \mathbb{R}^E \times \mathbb{R}^B &\longrightarrow \mathbb{R}^{n \times \mathcal{P}} \\ \theta &\longmapsto \alpha(X, \theta) \end{aligned}$$

is piecewise-constant, with a finite number of pieces. Furthermore, the boundary of each piece has Lebesgue measure zero. We call  $\Delta_X$  the union of all these boundaries. The set  $\Delta_X$  is closed and has Lebesgue measure zero.

As discussed before, for a given  $X \in \mathbb{R}^{n \times V_0}$ , when studying the function  $\theta \mapsto f_\theta(X)$ , Proposition 2 alongside (3) shows that on a piece over which  $\alpha_X$  is constant,  $f_\theta(X)$  depends linearly on  $\phi(\theta)$ . Since  $\Delta_X$  is closed with measure zero, for almost all  $\theta \in \mathbb{R}^E \times \mathbb{R}^B$ , there exists a neighborhood of  $\theta$  over which  $\alpha_X$  is constant. As noted for instance by Stock and Gribonval [37, Sec. 2], for any  $\tilde{\theta}$  in such a neighborhood, we thus have

$$f_\theta(X) - f_{\tilde{\theta}}(X) = \alpha(X, \theta) (\phi(\theta) - \phi(\tilde{\theta})). \quad (4)$$

Hence, studying  $\phi$  will allow us to understand better how  $f_\theta(X)$  locally depends on  $\theta$ .

### 2.3 Invariant rescaling operations on $\theta$

Some well-known rescaling operations on the parameters  $\theta$  do not affect the value of  $\phi(\theta)$ . Before detailing them, let us define, for all  $t \in \mathbb{R}$ , the sign indicator  $\text{sign}(t)$  as 1, 0 or  $-1$  depending on whether  $t > 0$ ,  $t = 0$  or  $t < 0$  respectively. For any  $\theta \in \mathbb{R}^E \times \mathbb{R}^B$ , we then define

$$\text{sign}(\theta) = \left( (\text{sign}(w_{v \rightarrow v'})_{v \rightarrow v' \in E}, (\text{sign}(b_v))_{v \in B} \right) \in \{-1, 0, 1\}^E \times \{-1, 0, 1\}^B.$$

We can now describe the rescaling operations.

**Definition 3.** Let  $\theta \in \mathbb{R}^E \times \mathbb{R}^B$  and  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ .

- We say that  $\theta$  is equivalent to  $\tilde{\theta}$  modulo rescaling, and we write  $\theta \stackrel{R}{\sim} \tilde{\theta}$  iff there exists a family of vectors  $(\lambda^0, \dots, \lambda^L) \in (\mathbb{R}^*)^{V_0} \times \dots \times (\mathbb{R}^*)^{V_L}$ , with  $\lambda^0 = \mathbb{1}_{V_0}$  and  $\lambda^L = \mathbb{1}_{V_L}$ , such that, for all  $l \in \llbracket 1, L \rrbracket$ ,

$$\begin{cases} W_l = \text{Diag}(\lambda^l) \tilde{W}_l \text{Diag}(\lambda^{l-1})^{-1} \\ b_l = \text{Diag}(\lambda^l) \tilde{b}_l. \end{cases} \quad (5)$$

- We say that  $\theta$  is equivalent to  $\tilde{\theta}$  modulo positive rescaling, and we write  $\theta \sim \tilde{\theta}$  iff

$$\theta \stackrel{R}{\sim} \tilde{\theta} \quad \text{and} \quad \text{sign}(\theta) = \text{sign}(\tilde{\theta}).$$

For all  $l \in \llbracket 1, L \rrbracket$ , to satisfy (5) is equivalent to satisfy, for all  $(v_{l-1}, v_l) \in V_{l-1} \times V_l$ ,

$$\begin{cases} w_{v_{l-1} \rightarrow v_l} = \frac{\lambda_{v_l}^l}{\lambda_{v_{l-1}}^{l-1}} \tilde{w}_{v_{l-1} \rightarrow v_l} \\ b_{v_l} = \lambda_{v_l}^l \tilde{b}_{v_l} \end{cases} \quad (6)$$

The relations  $\stackrel{R}{\sim}$  and  $\sim$  are equivalence relations on the set of parameters  $\mathbb{R}^E \times \mathbb{R}^B$ . The equivalence modulo positive rescaling  $\sim$  is a well-known invariant for ReLU networks [36, 37, 4, 26, 39]. We have indeed the following property: if  $\theta \sim \tilde{\theta}$ , for all  $x \in \mathbb{R}^{V_0}$ ,

$$f_\theta(x) = f_{\tilde{\theta}}(x). \quad (7)$$

One of the interests of the operator  $\phi$  is that it captures this invariant, as described by Stock and Gribonval [37, Sec. 2.4]. Propositions 4 and 5 are similar to their results and are restated here and proven in the supplement for completeness. Indeed, combining the definition of  $\phi$  with (6), we have the following property.

**Proposition 4.** For all  $\theta, \tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ , we have

$$\theta \stackrel{R}{\sim} \tilde{\theta} \implies \phi(\theta) = \phi(\tilde{\theta}),$$

and thus in particular

$$\theta \sim \tilde{\theta} \implies \phi(\theta) = \phi(\tilde{\theta}).$$

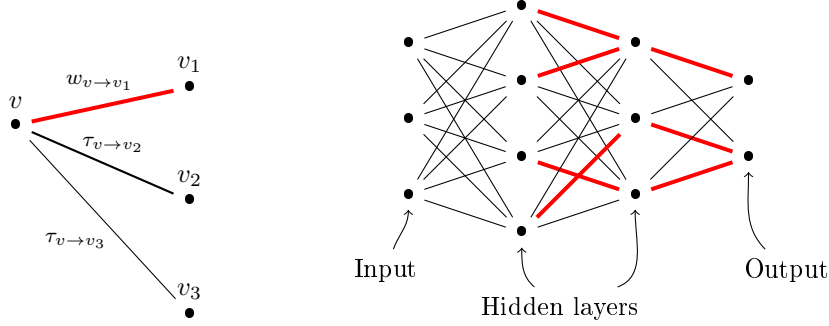


Figure 1: Left: The outward edges of a hidden neuron  $v$  and their weights. In this example,  $v_1 = s_{\max}^\theta(v)$ , so the weight of the edge in red,  $v \rightarrow v_1$ , has its value fixed as  $w_{v \rightarrow v_1}$ . The weights of the remaining edges,  $\tau_{v \rightarrow v_2}$  and  $\tau_{v \rightarrow v_3}$ , are free to vary. Right: In red, all the edges whose weights are fixed. The remaining edges, in black, constitute the set  $F_\theta$ .

The reciprocal of Proposition 4 holds provided we exclude some degenerate cases. Let us denote, for any  $l \in \llbracket 1, L-1 \rrbracket$  and any  $v \in V_l$ , by  $w_{\bullet \rightarrow v}$  the vector  $(w_{v' \rightarrow v})_{v' \in V_{l-1}} \in \mathbb{R}^{V_{l-1}}$  and by  $w_{v \rightarrow \bullet}$  the vector  $(w_{v \rightarrow v'})_{v' \in V_{l+1}} \in \mathbb{R}^{V_{l+1}}$ . We define the following set, which is close to the notion of ‘non admissible parameter’ in [37].

$$S = \{\theta \in \mathbb{R}^E \times \mathbb{R}^B, \exists v \in V_1 \cup \dots \cup V_{L-1}, w_{v \rightarrow \bullet} = 0 \text{ or } (w_{\bullet \rightarrow v}, b_v) = (0, 0)\}.$$

A parameterization  $\theta$  belongs to  $S$  iff there exists a hidden neuron  $v \in V_1 \cup \dots \cup V_{L-1}$  such that  $(w_{\bullet \rightarrow v}, b_v) = (0, 0)$  or  $w_{v \rightarrow \bullet} = 0$ . In the first case, all the inward weights and the bias of  $v$  are zero, so for any input the information flowing through neuron  $v$  is always zero. In the second case, all the outward weights of  $v$  are zero. In both cases, neuron  $v$  does not contribute to the output and could be removed from the network without changing the function  $f_\theta$ .

Since it is composed of a finite union of linear subspaces of codimension larger than 1, defined by linear equations, the set  $S$  is closed and has Lebesgue measure zero, so we can exclude the degenerate cases in  $S$  without loss of generality. Proposition 5 states that the reciprocal of Proposition 4 holds over  $(\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ .

**Proposition 5.** For all  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ , for all  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ ,

$$\phi(\theta) = \phi(\tilde{\theta}) \implies \theta \stackrel{R}{\sim} \tilde{\theta}.$$

### 3 The smooth manifold $\Sigma_1^*$

We explained in the previous section that studying  $\phi$  allows to better understand how the output  $f_\theta(X)$  locally depends on  $\theta$ . The image of  $\phi$  is of particular interest in this study and is the subject of this section. We define

$$\Sigma_1^* = \{\phi(\theta), \theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S\}.$$

The main result of this section, Theorem 6, states that  $\Sigma_1^*$  is a smooth manifold. This is a key element of the proofs of Theorems 7, 8 and 9.

As explained in Section 2.3, the positive rescaling operations on  $\theta$  described by the relation  $\sim$  do not affect the value of  $\phi(\theta)$  or of  $f_\theta(x)$  for any input  $x$ . This creates some degrees of freedom in the parameterization of a network without changing its output. To suppress these degrees of freedom, we propose to reduce the number of parameters by fixing locally the values of some weights as constants. More precisely, for a given  $\theta$ , for each neuron  $v$  in a hidden layer, we choose the outward edge  $v \rightarrow v'$  whose weight  $w_{v \rightarrow v'}$  has largest (absolute) value, and we consider its value to be fixed from now on (if there are several such edges, we choose one arbitrarily). We denote by  $s_{\max}^\theta(v)$  such a neuron  $v'$ . For each neuron  $v$  in a hidden layer  $V_l$ , there is exactly one neuron  $s_{\max}^\theta(v)$  in the layer  $V_{l+1}$ , and one corresponding edge  $v \rightarrow s_{\max}^\theta(v)$  whose weight is fixed. See Figure 1 for an illustration.

We denote by  $F_\theta \subset E$  the set of remaining edges, which is formally defined as<sup>1</sup>

$$F_\theta = E \setminus \left( \bigcup_{l=1}^{L-1} \left\{ (v, s_{\max}^\theta(v)), v \in V_l \right\} \right), \quad (8)$$

and we take as new parameters the weights in  $F_\theta$  and the biases  $B$ . This is formalized by the following application, for  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ ,

$$\begin{aligned} \rho_\theta : \mathbb{R}^{F_\theta} \times \mathbb{R}^B &\longrightarrow \mathbb{R}^E \times \mathbb{R}^B \\ \tau &\longmapsto \tilde{\theta} \quad \text{such that} \quad \begin{cases} \forall (v, v') \in F_\theta, & \tilde{w}_{v \rightarrow v'} = \tau_{v \rightarrow v'} \\ \forall (v, v') \in E \setminus F_\theta, & \tilde{w}_{v \rightarrow v'} = w_{v \rightarrow v'} \\ \forall v \in B, & \tilde{b}_v = \tau_v. \end{cases} \end{aligned} \quad (9)$$

In particular, if we define  $\tau_\theta \in \mathbb{R}^{F_\theta} \times \mathbb{R}^B$  by  $(\tau_\theta)_{v \rightarrow v'} = w_{v \rightarrow v'}$  and  $(\tau_\theta)_v = b_v$ , we have  $\rho_\theta(\tau_\theta) = \theta$ . The function  $\rho_\theta$  is affine and injective. We define

$$U_\theta = \rho_\theta^{-1} \left( (\mathbb{R}^E \times \mathbb{R}^B) \setminus S \right), \quad (10)$$

which is an open set of  $\mathbb{R}^{F_\theta} \times \mathbb{R}^B$ . We define, for all  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ , the local lifting operator

$$\begin{aligned} \psi^\theta : U_\theta &\longrightarrow \mathbb{R}^{\mathcal{P} \times V_L} \\ \tau &\longmapsto \phi \circ \rho_\theta(\tau). \end{aligned} \quad (11)$$

One can show that  $\psi^\theta$  is  $C^\infty$  and that it is a homeomorphism from  $U_\theta$  onto its image (see the proofs in the supplement), which we denote  $V_\theta$  and is thus an open subset of  $\Sigma_1^*$  (with the topology induced on  $\Sigma_1^*$  by the standard topology on  $\mathbb{R}^{\mathcal{P} \times V_L}$ ). In particular, since  $\rho_\theta(\tau_\theta) = \theta$ , we have  $\phi(\theta) = \psi^\theta(\tau_\theta) \in V_\theta$ . We have the following fundamental result.

**Theorem 6.**  $\Sigma_1^*$  is a smooth manifold of  $\mathbb{R}^{\mathcal{P} \times V_L}$  of dimension

$$|F_\theta| + |B| = N_0 N_1 + N_1 N_2 + \cdots + N_{L-1} N_L + N_L,$$

and the family  $(V_\theta, (\psi^\theta)^{-1})_{\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S}$  is an atlas.

Theorem 6 is proven in the supplement. Besides being key in Section 4, Theorem 6 (both the smooth manifold nature of  $\Sigma_1^*$  and the explicit atlas  $(V_\theta, (\psi^\theta)^{-1})_{\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S}$ ) may also be considered of more general independent interest. To our knowledge, such a result has not been established elsewhere in the literature.

## 4 Main results: necessary and sufficient conditions for local identifiability

The main results of this paper rely on the decomposition (4) introduced in Section 2. To reformulate (4), let us introduce the linear operator  $A(X, \theta)$ , which simply corresponds to the matrix product with  $\alpha(X, \theta)$ :

$$\begin{aligned} A(X, \theta) : \mathbb{R}^{\mathcal{P} \times V_L} &\longrightarrow \mathbb{R}^{n \times V_L} \\ \eta &\longmapsto \alpha(X, \theta)\eta, \end{aligned}$$

where  $\alpha(X, \theta)\eta$  is the matrix product between  $\alpha(X, \theta) \in \mathbb{R}^{n \times \mathcal{P}}$  and  $\eta \in \mathbb{R}^{\mathcal{P} \times V_L}$ . The operator  $A(X, \theta)$  inherits the properties of  $\alpha(X, \theta)$ , in particular those stated in Proposition 2. Using  $A(X, \theta)$ , the relation (4) satisfied by  $\tilde{\theta}$  in the neighborhood of  $\theta$  becomes

$$f_\theta(X) - f_{\tilde{\theta}}(X) = A(X, \theta) \cdot (\phi(\theta) - \phi(\tilde{\theta})). \quad (12)$$

Let us also define the affine space

$$N(X, \theta) = \phi(\theta) + \text{Ker } A(X, \theta).$$

If a parameterization  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$  is such that  $f_{\tilde{\theta}}(X) = f_\theta(X)$  and (12) holds, then  $\phi(\theta) - \phi(\tilde{\theta}) \in \text{Ker } A(X, \theta)$ , so by definition  $\phi(\tilde{\theta}) \in N(X, \theta)$ . Since for  $\tilde{\theta}$  in the neighborhood of  $\theta$ , the image  $\phi(\tilde{\theta})$

<sup>1</sup>Note, in the definition of  $F_\theta$ , the index  $l$  starting at  $l = 1$  and not  $l = 0$ .



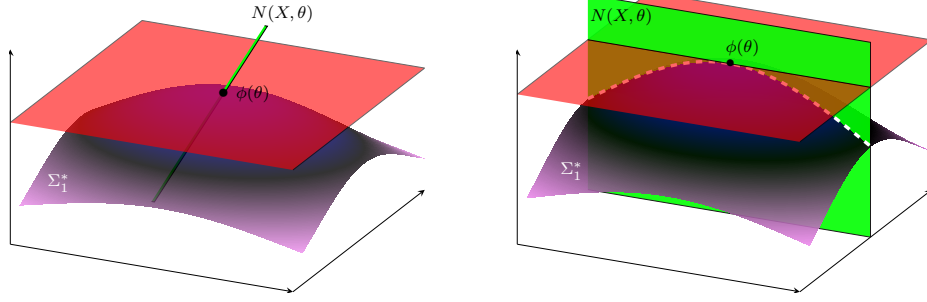


Figure 2: The local intersection between  $N(X, \theta)$  (in green) and  $\Sigma_1^*$  (color gradient). We also represent in red the tangent space to  $\Sigma_1^*$  at  $\phi(\theta)$ . Left: The identifiable case. The intersection is reduced to  $\{\phi(\theta)\}$ . Right: The non identifiable case. The intersection, represented with a dashed white line, is not reduced to  $\{\phi(\theta)\}$ .

belongs to  $\Sigma_1^*$ , this shows that local identifiability is closely related to the nature of the intersection between the smooth manifold  $\Sigma_1^*$  and the affine subspace  $N(X, \theta)$ .

Indeed, let us denote by  $B_\infty(\phi(\theta), \epsilon') = \{\eta \in \mathbb{R}^{\mathcal{P} \times V_L}, \|\phi(\theta) - \eta\|_\infty < \epsilon'\}$  the ball of center  $\phi(\theta)$  and of radius  $\epsilon' > 0$ . We have the following geometrical necessary and sufficient condition of local identifiability, which states that local identifiability of  $\theta$  holds if and only if the intersection between  $\Sigma_1^*$  and  $N(X, \theta)$  is locally reduced to the single point  $\{\phi(\theta)\}$ .

**Theorem 7.** For any  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus (S \cup \Delta_X)$ , the two following statements are equivalent.

i) There exists  $\epsilon > 0$  such that for any  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ , if  $\|\theta - \tilde{\theta}\|_\infty < \epsilon$ , then

$$f_\theta(X) = f_{\tilde{\theta}}(X) \implies \theta \sim \tilde{\theta}.$$

ii) There exists  $\epsilon' > 0$  such that

$$B_\infty(\phi(\theta), \epsilon') \cap \Sigma_1^* \cap N(X, \theta) = \{\phi(\theta)\}.$$

Theorem 7 is proven in the supplement, and is illustrated in Figure 2. This geometrical condition is crucial for showing the next two results which give testable conditions of identifiability. Theorems 8 and 9 rely on the rank of  $A(X, \theta)$  and of another linear operator  $\Gamma(X, \theta)$ , which we now define. Since, as we said, the function  $\psi^\theta$  is  $C^\infty$ , let us denote by  $D\psi^\theta(\tau) : \mathbb{R}^{F_\theta} \times \mathbb{R}^B \rightarrow \mathbb{R}^{\mathcal{P} \times V_L}$  its differential at the point  $\tau$ , for any  $\tau \in U_\theta$ . We define the linear operator  $\Gamma(X, \theta) : \mathbb{R}^{F_\theta} \times \mathbb{R}^B \rightarrow \mathbb{R}^{n \times V_L}$  by

$$\Gamma(X, \theta) = A(X, \theta) \circ D\psi^\theta(\tau_\theta). \quad (13)$$

We denote  $R_A = \text{rank}(A(X, \theta))$  and  $R_\Gamma = \text{rank}(\Gamma(X, \theta))$ . Since  $\Gamma(X, \theta)$  is defined on  $\mathbb{R}^{F_\theta} \times \mathbb{R}^B$ , we have

$$0 \leq R_\Gamma \leq |F_\theta| + |B|,$$

and the expression (13) shows that we also have

$$0 \leq R_\Gamma \leq R_A \leq |\mathcal{P}|N_L.$$

We can now define the two following conditions.

**Condition  $C_N$ .** Condition  $C_N$  is satisfied by  $(\theta, X)$  iff  $R_\Gamma < R_A$  or  $R_\Gamma = |F_\theta| + |B|$ .

**Condition  $C_S$ .** Condition  $C_S$  is satisfied by  $(\theta, X)$  iff  $R_\Gamma = |F_\theta| + |B|$ .

The following result states that  $C_N$  is a necessary condition for local and therefore global identifiability.

**Theorem 8** (Necessary condition of identifiability). Let  $X \in \mathbb{R}^{n \times V_0}$  and  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus (S \cup \Delta_X)$ . If  $C_N$  is not satisfied, then  $\theta$  is not locally identifiable, that is, for all  $\epsilon > 0$  there exists  $\tilde{\theta} \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus (S \cup \Delta_X)$  such that  $\|\theta - \tilde{\theta}\|_\infty < \epsilon$  and  $\theta \not\sim \tilde{\theta}$  but

$$f_\theta(X) = f_{\tilde{\theta}}(X).$$

Thus, in particular, if  $C_N$  is not satisfied, then  $\theta$  is not identifiable.

The following result states that  $C_S$  is a sufficient condition of local identifiability.

**Theorem 9** (Sufficient condition of local identifiability). *Let  $X \in \mathbb{R}^{n \times V_0}$  and  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus (S \cup \Delta_X)$ . If  $C_S$  is satisfied, then  $\theta$  is locally identifiable, that is there exists  $\epsilon > 0$  such that for all  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ , if  $\|\theta - \tilde{\theta}\|_\infty < \epsilon$ ,*

$$f_\theta(X) = f_{\tilde{\theta}}(X) \implies \theta \sim \tilde{\theta}.$$

Both theorems are proven in the supplement. To discuss these two results, let us first point out that  $C_N$  and  $C_S$  are close from one another. We argue in fact that they are sharp in the sense that the case separating them,  $R_\Gamma < R_A$ , corresponds to an existing alignment between the image of  $D\psi^\theta(\tau_\theta)$  and  $\text{Ker } A(X, \theta)$ , which should be unlikely as only  $\text{Ker } A(X, \theta)$  is sample dependent. Second, in order to have  $R_\Gamma = |F_\theta| + |B|$ , one needs to have  $nN_L \geq |F_\theta| + |B|$ . This means that the number of scalar measurements (number of samples  $n$  times output dimension  $N_L$ ) is larger than or equal to the number of parameters (up to local reparameterization).

## 5 Checking the conditions numerically

The key benefit of the conditions  $C_N$  and  $C_S$ , compared to the existing literature, is that they can be numerically tested for any fixed finite sample. They need the computation of the rank of two linear operators, namely  $\Gamma(X, \theta)$  and  $A(X, \theta)$ . The operator  $\Gamma(X, \theta)$  satisfies the following:

**Proposition 10.** *Let  $X \in \mathbb{R}^{n \times V_0}$  and  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus (S \cup \Delta_X)$ . The function*

$$\begin{aligned} U_\theta &\longrightarrow \mathbb{R}^{n \times V_L} \\ \tau &\longmapsto f_{\rho_\theta(\tau)}(X) \end{aligned}$$

*is differentiable in a neighborhood of  $\tau_\theta$ , and we denote by  $D_\tau f_\theta(X)$  its differential at  $\tau_\theta$ . We have*

$$D_\tau f_\theta(X) = \Gamma(X, \theta). \quad (14)$$

The proof of Proposition 10 is in the supplement. Since the reparameterization with  $\rho_\theta$  simply consists in fixing the weights of the edges  $v \rightarrow s_{\max}^\theta(v)$  to the value  $w_{v \rightarrow s_{\max}^\theta(v)}$ , (59) shows that the coefficients of  $\Gamma(X, \theta)$  can be computed by a classic backpropagation algorithm  $N_L$  times for each input  $x^i$ , simply omitting the derivatives with respect to the edges of the form  $v \rightarrow s_{\max}^\theta(v)$ . An explicit expression of the coefficients of  $\Gamma(X, \theta)$  is given in the supplement.

To be satisfied,  $C_S$  needs the dimensions of  $\Gamma(X, \theta)$  to satisfy  $nN_L \geq |F_\theta| + |B|$ . One then needs to compute the rank  $R_\Gamma$  of  $\Gamma(X, \theta)$ , which means computing the rank of a  $nN_L \times (|F_\theta| + |B|)$  matrix. Existing algorithms allow to do this with a complexity  $O(nN_L(|F_\theta| + |B|)^\omega)$  (up to polylog terms), where  $\omega$  is the matrix multiplication exponent and satisfies  $\omega < 2.38$  [8].

When it comes to  $C_N$ , one needs in addition to know the rank  $R_A$  of  $A(X, \theta)$ , which, as Proposition 11 states, requires to compute the rank of  $\alpha(X, \theta)$ .

**Proposition 11.** *Let  $X \in \mathbb{R}^{n \times V_0}$  and  $\theta \in \mathbb{R}^E \times \mathbb{R}^B$ . We have*

$$R_A = N_L \text{rank}(\alpha(X, \theta)).$$

The dimensions of  $\alpha(X, \theta)$  are sensibly larger, with  $|\mathcal{P}|$  columns and  $n$  lines, and typically  $|\mathcal{P}| \gg n$ . However it may have some sparsity properties, as its entries consist in products of activation indicators (with possibly one input  $x_{v_0}^i$ ), any one of them being zero causing many entries to vanish. The question of the efficient computation of  $R_A$  still needs to be explored and is left as open for future work.

## 6 Conclusion

This paper is the first to characterize local identifiability for deep ReLU networks for any given finite sample, with testable conditions. The practical use of these conditions deserves follow-up research, and so does an extension of our approach to inverse stability.

## Acknowledgments and Disclosure of Funding

The authors would like to thank Pierre Stock and Rémi Gribonval for the fruitful discussions around this work, notably regarding the construction of  $\phi$  and its link to the question of local identifiability.

This work has benefited from the AI Interdisciplinary Institute ANITI. ANITI is funded by the French “Investing for the Future – PIA3” program under the Grant agreement n°ANR-19-PI3A-0004.

The authors gratefully acknowledge the support of the DEEL project.<sup>2</sup>

## References

- [1] Rilwan A Adewoyin, Peter Dueben, Peter Watson, Yulan He, and Ritabrata Dutta. Tru-net: a deep learning approach to high resolution prediction of rainfall. *Machine Learning*, 110(8): 2035–2062, 2021.
- [2] Francesca Albertini, Eduardo D Sontag, and Vincent Maillot. Uniqueness of weights for neural networks. *Artificial Neural Networks for Speech and Vision*, pages 115–125, 1993.
- [3] Sanjeev Arora, Aditya Bhaskara, Rong Ge, and Tengyu Ma. Provable bounds for learning some deep representations. In Eric P. Xing and Tony Jebara, editors, *Proceedings of the 31st International Conference on Machine Learning*, volume 32 of *Proceedings of Machine Learning Research*, pages 584–592, Beijing, China, 22–24 Jun 2014. PMLR.
- [4] Joachim Bona-Pellissier, François Bachoc, and François Malgouyres. Parameter identifiability of a deep feedforward ReLU neural network. *arXiv preprint arXiv:2112.12982*, 2021.
- [5] Alon Brutzkus and Amir Globerson. Globally optimal gradient descent for a ConvNet with Gaussian inputs. In *Proceedings of the 34th International Conference on Machine Learning—Volume 70*, pages 605–614, 2017.
- [6] Nicholas Carlini, Chang Liu, Úlfar Erlingsson, Jernej Kos, and Dawn Song. The secret sharer: Evaluating and testing unintended memorization in neural networks. In *28th {USENIX} Security Symposium ({USENIX} Security 19)*, pages 267–284, 2019.
- [7] Nicholas Carlini, Matthew Jagielski, and Ilya Mironov. Cryptanalytic extraction of neural network models. In *Annual International Cryptology Conference*, pages 189–218. Springer, 2020.
- [8] Ho Yee Cheung, Tsz Chiu Kwok, and Lap Chi Lau. Fast matrix rank algorithms and applications. *Journal of the ACM (JACM)*, 60(5):1–25, 2013.
- [9] Dennis Maximilian Elbrächter, Julius Berner, and Philipp Grohs. How degenerate is the parametrization of neural networks with the ReLU activation function? In *Advances in Neural Information Processing Systems*, volume 32, 2019.
- [10] Charles Fefferman. Reconstructing a neural net from its output. *Revista Matemática Iberoamericana*, 10(3):507–555, 1994.
- [11] Matt Fredrikson, Somesh Jha, and Thomas Ristenpart. Model inversion attacks that exploit confidence information and basic countermeasures. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, pages 1322–1333, 2015.
- [12] Haoyu Fu, Yuejie Chi, and Yingbin Liang. Guaranteed recovery of one-hidden-layer neural networks via cross entropy. *IEEE Transactions on Signal Processing*, 68:3225–3235, 2020.
- [13] Rong Ge, Jason D Lee, and Tengyu Ma. Learning one-hidden-layer neural networks with landscape design. In *6th International Conference on Learning Representations, ICLR 2018*, 2018.
- [14] Awni Hannun, Carl Case, Jared Casper, Bryan Catanzaro, Greg Diamos, Erich Elsen, Ryan Prenger, Sanjeev Satheesh, Shubho Sengupta, Adam Coates, et al. Deep speech: Scaling up end-to-end speech recognition. *arXiv preprint arXiv:1412.5567*, 2014.

---

<sup>2</sup><https://www.deel.ai/>

- [15] Geoffrey Hinton, Li Deng, Dong Yu, George E. Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N. Sainath, and Brian Kingsbury. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6):82–97, 2012.
- [16] Paul C Kainen, Věra Kůrková, Vladik Kreinovich, and Ongard Sirisaengtaksin. Uniqueness of network parametrization and faster learning. *Neural, Parallel & Scientific Computations*, 2(4): 459–466, 1994.
- [17] Nal Kalchbrenner and Phil Blunsom. Recurrent continuous translation models. In *Proceedings of the 2013 conference on empirical methods in natural language processing*, pages 1700–1709, 2013.
- [18] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012.
- [19] Alexey Kurakin, Ian Goodfellow, and Samy Bengio. Adversarial examples in the physical world. *International Conference on Learning Representations*, 2017.
- [20] Věra Kůrková and Paul C Kainen. Functionally equivalent feedforward neural networks. *Neural Computation*, 6(3):543–558, 1994.
- [21] François Malgouyres and Joseph Landsberg. On the identifiability and stable recovery of deep/multi-layer structured matrix factorization. In *IEEE, Info. Theory Workshop*, Sept. 2016.
- [22] François Malgouyres and Joseph Landsberg. Multilinear compressive sensing and an application to convolutional linear networks. *SIAM Journal on Mathematics of Data Science*, 1(3):446–475, 2019.
- [23] Francois Malgouyres. On the stable recovery of deep structured linear networks under sparsity constraints. In *Mathematical and Scientific Machine Learning*, pages 107–127. PMLR, 2020.
- [24] Tomas Mikolov, Martin Karafiát, Lukas Burget, Jan Cernocký, and Sanjeev Khudanpur. Recurrent neural network based language model. In *Interspeech*, volume 2, pages 1045–1048, 2010.
- [25] Tomáš Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. In Yoshua Bengio and Yann LeCun, editors, *1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings*, 2013.
- [26] Behnam Neyshabur, Russ R Salakhutdinov, and Nati Srebro. Path-SGD: Path-normalized optimization in deep neural networks. *Advances in neural information processing systems*, 28, 2015.
- [27] Philipp Petersen, Mones Raslan, and Felix Voigtlaender. Topological properties of the set of functions generated by neural networks of fixed size. *Foundations of Computational Mathematics*, 21:375–444, 2021.
- [28] Henning Petzka, Martin Trimmel, and Cristian Sminchisescu. Notes on the symmetries of 2-layer ReLU-networks. In *Proceedings of the Northern Lights Deep Learning Workshop*, volume 1, pages 6–6, 2020.
- [29] Mary Phuong and Christoph H. Lampert. Functional vs. parametric equivalence of ReLU networks. In *International Conference on Learning Representations*, 2020.
- [30] José Pedro Pinto, André Pimenta, and Paulo Novais. Deep learning and multivariate time series for cheat detection in video games. *Machine Learning*, 110(11):3037–3057, 2021.
- [31] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.

- [32] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28:91–99, 2015.
- [33] David Rolnick and Konrad Kording. Reverse-engineering deep ReLU networks. In Hal Daumé III and Aarti Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 8178–8187, 13–18 Jul 2020.
- [34] Haşim Sak, Andrew Senior, and Françoise Beaufays. Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In *Fifteenth Annual Conference of the International Speech Communication Association*, 2014.
- [35] Hanie Sedghi and Anima Anandkumar. Provable methods for training neural networks with sparse connectivity. In *Deep Learning and representation learning workshop: NIPS*, 2014.
- [36] Pierre Stock. *Efficiency and Redundancy in Deep Learning Models : Theoretical Considerations and Practical Applications*. PhD thesis, Université de Lyon, April 2021. URL <https://tel.archives-ouvertes.fr/tel-03208517>.
- [37] Pierre Stock and Rémi Gribonval. An Embedding of ReLU Networks and an Analysis of their Identifiability. *Constructive Approximation*, 2022. URL <https://hal.archives-ouvertes.fr/hal-03292203>.
- [38] Héctor J Sussmann. Uniqueness of the weights for minimal feedforward nets with a given input-output map. *Neural networks*, 5(4):589–593, 1992.
- [39] Mingyang Yi, Qi Meng, Wei Chen, Zhi-ming Ma, and Tie-Yan Liu. Positively scale-invariant flatness of ReLU neural networks. *arXiv preprint arXiv:1903.02237*, 2019.
- [40] Jialong Zhang, Zhongshu Gu, Jiyong Jang, Hui Wu, Marc Ph Stoecklin, Heqing Huang, and Ian Molloy. Protecting intellectual property of deep neural networks with watermarking. In *Proceedings of the 2018 on Asia Conference on Computer and Communications Security*, pages 159–172, 2018.
- [41] Shuai Zhang, Meng Wang, Jinjun Xiong, Sijia Liu, and Pin-Yu Chen. Improved linear convergence of training CNNs with generalizability guarantees: A one-hidden-layer case. *IEEE Transactions on Neural Networks and Learning Systems*, 32(6):2622–2635, 2020.
- [42] Xiao Zhang, Yaodong Yu, Lingxiao Wang, and Quanquan Gu. Learning one-hidden-layer ReLU networks via gradient descent. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 1524–1534. PMLR, 2019.
- [43] Kai Zhong, Zhao Song, Prateek Jain, Peter L Bartlett, and Inderjit S Dhillon. Recovery guarantees for one-hidden-layer neural networks. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 4140–4149, 2017.
- [44] Mo Zhou, Rong Ge, and Chi Jin. A local convergence theory for mildly over-parameterized two-layer neural network. *arXiv preprint arXiv:2102.02410*, 2021.

## Checklist

1. For all authors...
  - (a) Do the main claims made in the abstract and introduction accurately reflect the paper’s contributions and scope? **[Yes]** The outline in Section 1.4 (“Overview of the article”) provides pointers to where the claimed contributions of the paper are provided.
  - (b) Did you describe the limitations of your work? **[Yes]** Section 5 acknowledges the open problem of an efficient computation of the rank of  $\alpha(X, \theta)$  and Section 6 describes other remaining open questions.

- (c) Did you discuss any potential negative societal impacts of your work? [N/A] This is a theoretical/foundation work that adds to the theory and methodology of deep learning. As for any such contributions, the positive or negative societal impact will depend on the application case. We do not promote any harmful use of this theory, but we expand on the existing knowledge.
  - (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes] See previous question.
2. If you are including theoretical results...
- (a) Did you state the full set of assumptions of all theoretical results? [Yes] All our results explicitly refer to their required assumptions. Some general assumptions that hold throughout the paper are also stated at the beginning.
  - (b) Did you include complete proofs of all theoretical results? [Yes] All the proofs are provided in the supplement.
3. If you ran experiments...
- (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [N/A] We did not run experiments.
  - (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [N/A]
  - (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [N/A]
  - (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [N/A]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
- (a) If your work uses existing assets, did you cite the creators? [N/A] We did not use existing assets (code, data or models) nor cure/release new assets (code, data or models).
  - (b) Did you mention the license of the assets? [N/A]
  - (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]
  - (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
  - (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
- (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A] We did not use crowdsourcing nor conducted research with human subjects.
  - (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
  - (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

## A Notations

In this section, we define notations, many of which are standard, that are useful in the proofs.

We denote by  $\mathbb{N}$  the set of all natural numbers, including 0, and by  $\mathbb{N}^*$  the set  $\mathbb{N}$  without 0. We denote by  $\mathbb{Z}$  the set of all integers. For any  $a, b \in \mathbb{Z}$ , we denote by  $\llbracket a, b \rrbracket$  the set of all integers  $k \in \mathbb{Z}$  satisfying  $a \leq k \leq b$ . For any finite set  $A$ , we denote by  $|A|$  the cardinal of  $A$ .

For  $n, N \in \mathbb{N}^*$ , we denote by  $\mathbb{R}^N$  the  $N$ -dimensional real vector space and by  $\mathbb{R}^{n \times N}$  the vector space of real matrices with  $n$  lines and  $N$  columns. For a vector  $x = (x_1, \dots, x_N)^T \in \mathbb{R}^N$ , we use the norm  $\|x\|_\infty = \max_{i \in \llbracket 1, N \rrbracket} |x_i|$ . For  $x \in \mathbb{R}^N$  and  $r > 0$ , we denote  $B_\infty(x, r) = \{y \in \mathbb{R}^N, \|y - x\|_\infty < r\}$ .

For any vector  $x = (x_1, \dots, x_N)^T \in \mathbb{R}^N$ , we define  $\text{sign}(x) = (\text{sign}(x_1), \dots, \text{sign}(x_N))^T \in \{-1, 0, 1\}^N$  as the vector whose  $i^{\text{th}}$  component is equal to

$$\text{sign}(x_i) = \begin{cases} 1 & \text{if } x_i > 0 \\ 0 & \text{if } x_i = 0 \\ -1 & \text{if } x_i < 0. \end{cases}$$

For any matrix  $M \in \mathbb{R}^{n \times N}$ , for all  $i \in \llbracket 1, n \rrbracket$ , we denote by  $M_{i,\cdot}$  the  $i^{\text{th}}$  line of  $M$ . The vector  $M_{i,\cdot}$  is a line vector whose  $j^{\text{th}}$  component is  $M_{i,j}$ . Similarly, for  $j \in \llbracket 1, N \rrbracket$ , we denote by  $M_{\cdot,j}$  the  $j^{\text{th}}$  column of  $M$ , which is the column vector whose  $i^{\text{th}}$  component is  $M_{i,j}$ . For any matrix  $M \in \mathbb{R}^{n \times N}$ , we denote by  $M^T \in \mathbb{R}^{N \times n}$  the transpose matrix of  $M$ .

We denote by  $\text{Id}_N$  the  $N \times N$  identity matrix and by  $\mathbb{1}_N$  the vector  $(1, 1, \dots, 1)^T \in \mathbb{R}^N$ . If  $\lambda \in \mathbb{R}^N$  is a vector of size  $N$ , for some  $N \in \mathbb{N}^*$ , we denote by  $\text{Diag}(\lambda)$  the  $N \times N$  matrix defined by:

$$\text{Diag}(\lambda)_{i,j} = \begin{cases} \lambda_i & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

If  $X$  and  $Y$  are two sets and  $h : X \rightarrow Y$  is a function, for a subset  $A \subset Y$ , we denote by  $h^{-1}(A)$  the preimage of  $A$  under  $f$ , that is

$$h^{-1}(A) = \{x \in X, h(x) \in A\}.$$

Note that this does not require the function  $h$  to be injective.

For any  $n, N \in \mathbb{N}^*$  and any differentiable function  $f : \mathbb{R}^n \rightarrow \mathbb{R}^N$ , for all  $x \in \mathbb{R}^n$ , we denote by  $Df(x)$  its differential at the point  $x$ , i.e. the linear application  $Df(x) : \mathbb{R}^n \rightarrow \mathbb{R}^N$  satisfying, for all  $h \in \mathbb{R}^n$ ,

$$f(x + h) = f(x) + Df(x) \cdot h + o(h).$$

If we denote by  $x_j$  and  $h_j$  the components of  $x$  and  $h$ , for  $j \in \llbracket 1, n \rrbracket$ , we have

$$Df(x) \cdot h = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x) h_j,$$

where for all  $j$ ,  $\frac{\partial f}{\partial x_j}(x) \in \mathbb{R}^N$ . If  $f : \mathbb{R}^n \rightarrow \mathbb{R}^N$  is a linear application, we denote by  $\text{Ker } f$  the set  $\{x \in \mathbb{R}^n, f(x) = 0\}$ , which is a linear subset of  $\mathbb{R}^n$ .

## B The lifting operator $\phi$

Let us introduce the notion of ‘path’, extending the definition in Section 2.2. A path is a sequence of neurons  $(v_k, v_{k+1}, \dots, v_l) \in V_k \times V_{k+1} \times \dots \times V_l$ , for integers  $k, l$  satisfying  $0 \leq k \leq l \leq L$ . In particular, for all  $l \in \llbracket 0, L - 1 \rrbracket$ , the set  $\mathcal{P}_l$  defined in Section 2.2 contains all the paths starting from layer  $l$  and ending in layer  $L - 1$ . We recall

$$\mathcal{P} = \left( \bigcup_{l=0}^{L-1} \mathcal{P}_l \right) \cup \{\beta\}.$$

If  $k, l, m \in \mathbb{N}$  are three integers satisfying  $0 \leq k < l \leq m \leq L$ , and  $p = (v_k, \dots, v_{l-1}) \in V_k \times \dots \times V_{l-1}$  and  $p' = (v_l, \dots, v_m) \in V_l \times \dots \times V_m$  are two paths such that  $p$  ends in the layer preceding the starting layer of  $p'$ , we define the union of the paths by

$$p \cup p' = (v_k, \dots, v_{l-1}, v_l, \dots, v_m) \in V_k \times \dots \times V_m.$$

Before proving Proposition 1, let us compare briefly our construction to [37]. The lifting operator  $\phi$  introduced in Section 2.2 is similar to the operator  $\Phi$  in [37], except that  $\Phi$  does not take a matrix form. The operator  $\alpha(x, \theta)$  introduced in Section 2.2 corresponds partly to the object  $\bar{\alpha}(\theta, x)$  in [37]. One of the differences is that  $\bar{\alpha}(\theta, x)$  does not include any product with  $x_{v_0}$  in its entries, as does  $\alpha(x, \theta)$ . Finally, a similar statement to Proposition 1 and a similar proof can be found in [37]. However, one of the present contributions is to simplify the construction.

Let us now prove Proposition 1, which we restate here.

**Proposition 12.** For all  $\theta \in \mathbb{R}^E \times \mathbb{R}^B$  and all  $x \in \mathbb{R}^{V_0}$ ,

$$f_\theta(x)^T = \alpha(x, \theta)\phi(\theta).$$

*Proof.* Let us prove first the following expression, for all  $v_L \in V_L$ :

$$\begin{aligned} f_\theta(x)_{v_L} = & \left( \sum_{v_0 \in V_0} x_{v_0} w_{v_0 \rightarrow v_1} \prod_{l=1}^{L-1} a_{v_l}(x, \theta) w_{v_l \rightarrow v_{l+1}} \right) \\ & \vdots \\ & v_{L-1} \in V_{L-1} \\ & + \left( \sum_{l=1}^{L-1} \sum_{v_l \in V_l} b_{v_l} \prod_{l'=l}^{L-1} a_{v_{l'}}(x, \theta) w_{v_{l'} \rightarrow v_{l'+1}} \right) + b_{v_L}. \end{aligned} \quad (15)$$

We prove this by induction on the number  $L$  of layers of the network.

Initialization ( $L = 2$ ). Let  $v_2 \in V_2$ .

$$\begin{aligned} f_\theta(x)_{v_2} &= (W_2)_{v_2, :} \sigma(W_1 x + b_1) + b_{v_2} \\ &= \left( \sum_{v_1 \in V_1} w_{v_1 \rightarrow v_2} [\sigma(W_1 x + b_1)]_{v_1} \right) + b_{v_2} \\ &= \left( \sum_{v_1 \in V_1} w_{v_1 \rightarrow v_2} \sigma((W_1)_{v_1, :} x + b_{v_1}) \right) + b_{v_2} \\ &= \left( \sum_{v_1 \in V_1} w_{v_1 \rightarrow v_2} a_{v_1}(x, \theta) \left( \sum_{v_0 \in V_0} w_{v_0 \rightarrow v_1} x_{v_0} + b_{v_1} \right) \right) + b_{v_2} \\ &= \left( \sum_{\substack{v_0 \in V_0 \\ v_1 \in V_1}} w_{v_1 \rightarrow v_2} a_{v_1}(x, \theta) w_{v_0 \rightarrow v_1} x_{v_0} \right) + \left( \sum_{v_1 \in V_1} w_{v_1 \rightarrow v_2} a_{v_1}(x, \theta) b_{v_1} \right) + b_{v_2} \\ &= \left( \sum_{\substack{v_0 \in V_0 \\ v_1 \in V_1}} x_{v_0} w_{v_0 \rightarrow v_1} a_{v_1}(x, \theta) w_{v_1 \rightarrow v_2} \right) + \left( \sum_{v_1 \in V_1} b_{v_1} a_{v_1}(x, \theta) w_{v_1 \rightarrow v_2} \right) + b_{v_2} \end{aligned}$$

which proves (15), when  $L = 2$ .

Now let  $L \geq 3$  and suppose (15) holds for all ReLU networks with  $L - 1$  layers. Let us consider a network with  $L$  layers.



Let us denote by  $g_\theta(x)$  the output of the  $L - 1$  first layers of the network pre-activation (before applying the ReLUs of the layer  $L - 1$ ). The function  $g_\theta$  is that of a ReLU network with  $L - 1$  layers, and we have

$$f_\theta(x) = W_L \sigma(g_\theta(x)) + b_L.$$

Let  $v_L \in V_L$ . We thus have

$$f_\theta(x)_{v_L} = \sum_{v_{L-1} \in V_{L-1}} w_{v_{L-1} \rightarrow v_L} \sigma(g_\theta(x)_{v_{L-1}}) + b_{v_L}. \quad (16)$$

By the induction hypothesis, for all  $v_{L-1} \in V_{L-1}$ ,  $g_\theta(x)_{v_{L-1}}$  can be expressed with (15). Considering that  $\sigma(g_\theta(x)_{v_{L-1}}) = a_{v_{L-1}}(x, \theta) g_\theta(x)_{v_{L-1}}$  and replacing  $g_\theta(x)_{v_{L-1}}$  by its expression using (15), (16) becomes

$$\begin{aligned} f_\theta(x)_{v_L} &= \sum_{v_{L-1} \in V_{L-1}} w_{v_{L-1} \rightarrow v_L} a_{v_{L-1}}(x, \theta) \left[ \left( \sum_{v_0 \in V_0} x_{v_0} w_{v_0 \rightarrow v_1} \prod_{l=1}^{L-2} a_{v_l}(x, \theta) w_{v_l \rightarrow v_{l+1}} \right) \right. \\ &\quad \vdots \\ &\quad \left. + \left( \sum_{l=1}^{L-2} \sum_{v_l \in V_l} b_{v_l} \prod_{l'=l}^{L-2} a_{v_{l'}}(x, \theta) w_{v_{l'} \rightarrow v_{l'+1}} \right) + b_{v_{L-1}} \right] + b_{v_L} \\ &\quad \vdots \\ &\quad \left. + \left( \sum_{v_{L-2} \in V_{L-2}} w_{v_{L-2} \rightarrow v_{L-1}} a_{v_{L-1}}(x, \theta) x_{v_0} w_{v_0 \rightarrow v_1} \prod_{l=1}^{L-2} a_{v_l}(x, \theta) w_{v_l \rightarrow v_{l+1}} \right) \right. \\ &\quad \vdots \\ &\quad \left. + \left( \sum_{l=1}^{L-2} \sum_{v_l \in V_l} w_{v_{L-1} \rightarrow v_L} a_{v_{L-1}}(x, \theta) b_{v_l} \prod_{l'=l}^{L-2} a_{v_{l'}}(x, \theta) w_{v_{l'} \rightarrow v_{l'+1}} \right) \right. \\ &\quad \vdots \\ &\quad \left. + \left( \sum_{v_{L-1} \in V_{L-1}} w_{v_{L-1} \rightarrow v_L} a_{v_{L-1}}(x, \theta) b_{v_{L-1}} \right) + b_{v_L} \right. \\ &= \left( \sum_{v_0 \in V_0} x_{v_0} w_{v_0 \rightarrow v_1} \prod_{l=1}^{L-1} a_{v_l}(x, \theta) w_{v_l \rightarrow v_{l+1}} \right) \\ &\quad \vdots \\ &\quad \left. + \left( \sum_{l=1}^{L-1} \sum_{v_l \in V_l} b_{v_l} \prod_{l'=l}^{L-1} a_{v_{l'}}(x, \theta) w_{v_{l'} \rightarrow v_{l'+1}} \right) + b_{v_L}, \right. \\ &\quad \vdots \\ &\quad \left. + \left( \sum_{v_{L-1} \in V_{L-1}} w_{v_{L-1} \rightarrow v_L} a_{v_{L-1}}(x, \theta) b_{v_{L-1}} \right) + b_{v_L}, \right. \\ &= \left( \sum_{v_0 \in V_0} x_{v_0} w_{v_0 \rightarrow v_1} \prod_{l=1}^{L-1} a_{v_l}(x, \theta) w_{v_l \rightarrow v_{l+1}} \right) \\ &\quad \vdots \\ &\quad \left. + \left( \sum_{l=1}^{L-1} \sum_{v_l \in V_l} b_{v_l} \prod_{l'=l}^{L-1} a_{v_{l'}}(x, \theta) w_{v_{l'} \rightarrow v_{l'+1}} \right) + b_{v_L}, \right. \\ &\quad \vdots \\ &\quad \left. + \left( \sum_{v_{L-1} \in V_{L-1}} w_{v_{L-1} \rightarrow v_L} a_{v_{L-1}}(x, \theta) b_{v_{L-1}} \right) + b_{v_L}, \right. \end{aligned}$$

which proves (15) holds for ReLU networks with  $L$  layers. This ends the induction, and we conclude that (15) holds for all ReLU networks.

We can now use this expression to prove Proposition 12. The first sum in (15) is taken over all the paths  $p = (v_0, \dots, v_{L-1}) \in \mathcal{P}_0$ , and each summand can be written as

$$x_{v_0} w_{v_0 \rightarrow v_1} \prod_{l=1}^{L-1} a_{v_l}(x, \theta) w_{v_l \rightarrow v_{l+1}} = \left( x_{v_0} \prod_{l=1}^{L-1} a_{v_l}(x, \theta) \right) \left( \prod_{l=0}^{L-1} w_{v_l \rightarrow v_{l+1}} \right) = \alpha_p(x, \theta) \phi_{p, v_L}(\theta).$$

For all  $l \in \llbracket 1, L-1 \rrbracket$ , the inner sum of the double sum in (15) is taken over all the paths  $p = (v_l, \dots, v_{L-1}) \in \mathcal{P}_l$ , and each summand can be written as

$$b_{v_l} \prod_{l'=l}^{L-1} a_{v_{l'}}(x, \theta) w_{v_{l'} \rightarrow v_{l'+1}} = \left( \prod_{l'=l}^{L-1} a_{v_{l'}}(x, \theta) \right) \left( b_{v_l} \prod_{l'=l}^{L-1} w_{v_{l'} \rightarrow v_{l'+1}} \right) = \alpha_p(x, \theta) \phi_{p, v_L}(\theta).$$

And finally, we can also write

$$b_{v_L} = \alpha_\beta(x, \theta) \phi_{\beta, v_L}(\theta).$$

Joining all these sums and denoting  $\phi_{\cdot, v_L}(\theta) = (\phi_{p, v_L}(\theta))_{p \in \mathcal{P}} \in \mathbb{R}^{\mathcal{P}}$ , we have

$$f_\theta(x)_{v_L} = \sum_{p \in \mathcal{P}} \alpha_p(x, \theta) \phi_{p, v_L}(\theta) = \alpha(x, \theta) \phi_{\cdot, v_L}(\theta),$$

so in other words,

$$f_\theta(x)^T = \alpha(x, \theta) \phi(\theta).$$

□

We restate here and prove Proposition 2.

**Proposition 13.** *For all  $n \in \mathbb{N}^*$ , for all  $X \in \mathbb{R}^{n \times V_0}$ , the mapping*

$$\begin{aligned} \alpha_X : \mathbb{R}^E \times \mathbb{R}^B &\longrightarrow \mathbb{R}^{n \times \mathcal{P}} \\ \theta &\longmapsto \alpha(X, \theta) \end{aligned}$$

*appearing in (3) is piecewise-constant, with a finite number of pieces. Furthermore, the boundary of each piece has Lebesgue measure zero. We call  $\Delta_X$  the union of all the boundaries. The set  $\Delta_X$  is closed and has Lebesgue measure zero.*

*Proof.* Let us first notice that for any  $i \in \llbracket 1, n \rrbracket$ , for any  $l \in \llbracket 1, L-1 \rrbracket$ ,

$$(a_v(x^i, \theta))_{v \in V_1 \cup \dots \cup V_{l-1}} \in \{0, 1\}^{V_1 \cup \dots \cup V_{l-1}}$$

takes at most  $2^{N_1 + \dots + N_{l-1}}$  distinct values, so the mapping  $\theta \mapsto (a_v(x^i, \theta))_{v \in V_1 \cup \dots \cup V_{l-1}}$  is piecewise constant, with a finite number of pieces.

Let  $i \in \llbracket 1, n \rrbracket$ . Let  $l \in \llbracket 1, L-1 \rrbracket$  and  $v \in V_l$ . Recall the definition of  $f_{l-1}$ , as given in Section 2.1. The function  $\theta \rightarrow a_v(x^i, \theta)$  takes only two values, 1 or 0, and its values are determined by the sign of

$$\sum_{v' \in V_{l-1}} w_{v' \rightarrow v} f_{l-1}(x^i)_{v'} + b_v. \quad (17)$$

For all  $v' \in V_{l-1}$ , the value of  $f_{l-1}(x^i)_{v'}$  depends on  $\theta$ . On a piece  $P \subset \mathbb{R}^E \times \mathbb{R}^B$  such that  $(a_{v''}(x^i, \theta))_{v'' \in V_1 \cup \dots \cup V_{l-1}}$  is constant, this dependence is polynomial. Thus, on  $P$ , the value of (17) is a polynomial function of  $\theta$ , and since the coefficient applied to  $b_v$  is equal to 1, the corresponding polynomial is non constant. Since the values of  $a_v(x^i, \theta)$  are determined by the sign of (17), inside  $P$ , the boundary between  $\{\theta \in \mathbb{R}^E \times \mathbb{R}^B, a_v(x^i, \theta) = 0\}$  and  $\{\theta \in \mathbb{R}^E \times \mathbb{R}^B, a_v(x^i, \theta) = 1\}$  is included in the set of  $\theta$  for which (17) equals 0. This piece of boundary is thus contained in a level set of a non constant polynomial, whose Lebesgue measure is zero.

Since there is a finite number of pieces  $P$ , the Lebesgue measure of the boundary between  $\{\theta \in \mathbb{R}^E \times \mathbb{R}^B, a_v(x^i, \theta) = 0\}$  and  $\{\theta \in \mathbb{R}^E \times \mathbb{R}^B, a_v(x^i, \theta) = 1\}$ , which is contained in the union of the boundaries on all the pieces  $P$ , is thus equal to 0.

Since this is true for all  $l \in \llbracket 1, L-1 \rrbracket$  and all  $v \in V_l$ , the boundary of a piece over which  $(a_v(x^i, \theta))_{v \in V_1 \cup \dots \cup V_{L-1}}$  is constant also has Lebesgue measure zero.

Now since, for all  $x^i$ , the value of  $\alpha(x^i, \theta)$  only depends on  $(a_v(x^i, \theta))_{v \in V_1 \cup \dots \cup V_{L-1}}$  and since  $\alpha_X(\theta)$  is a matrix whose lines are the vectors  $\alpha(x^i, \theta)$ , we can conclude that

$\alpha_X : \mathbb{R}^E \times \mathbb{R}^B \longrightarrow \mathbb{R}^{n \times \mathcal{P}}$   
 $\theta \longmapsto \alpha(X, \theta)$  is piecewise-constant, with a finite number of pieces, and that the boundary of each piece has Lebesgue measure zero.

A boundary is, by definition, closed. Finally, a finite union of closed sets with Lebesgue measure 0, as  $\Delta_X$  is, is closed and has Lebesgue measure 0. □

For convenience, we introduce the two following notations. Let  $l \in \llbracket 0, L \rrbracket$ . For any  $l' \in \llbracket 0, l \rrbracket$  and any path  $p_i = (v_{l'}, \dots, v_l) \in V_{l'} \times \dots \times V_l$ , we denote

$$\theta_{p_i} = \begin{cases} \prod_{k=0}^{l'-1} w_{v_k \rightarrow v_{k+1}} & \text{if } l' = 0 \\ b_{l'} \prod_{k=l'}^{l-1} w_{v_k \rightarrow v_{k+1}} & \text{if } l' \geq 1, \end{cases} \quad (18)$$

where as a classic convention, an empty product is equal to 1. In particular, if  $l = 0$ , for any  $p_i = (v_0) \in V_0$ , we have  $\theta_{p_i} = 1$ . For any path  $p_o = (v_l, \dots, v_L) \in V_l \times \dots \times V_L$ , we denote

$$\theta_{p_o} = \prod_{k=l}^{L-1} w_{v_k \rightarrow v_{k+1}}, \quad (19)$$

with again the convention that an empty product is equal to 1, so if  $l = L$ ,  $\theta_{p_o} = 1$ .

Some attention must be paid to the fact that for any  $l' \in \llbracket 1, L \rrbracket$ , if we take  $p_i$  in the case  $l = L$  and  $p_o$  in the case  $l = l'$ , it is possible to have

$$p_i = (v_{l'}, \dots, v_L) = p_o,$$

but in that case we DO NOT have  $\theta_{p_i} = \theta_{p_o}$ , since  $\theta_{p_i} = b_{l'} \prod_{k=l'}^{L-1} w_{v_k \rightarrow v_{k+1}}$  and  $\theta_{p_o} = \prod_{k=l'}^{L-1} w_{v_k \rightarrow v_{k+1}}$ . We will always denote the paths  $p_i$  and  $p_o$  with an  $i$  (as in ‘input’) or an  $o$  (as in ‘output’) to clarify which definition is used.

When considering another parameterization  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ , we denote by  $\tilde{\theta}_{p_i}$  and  $\tilde{\theta}_{p_o}$  the corresponding objects.

We establish different characterizations of the set  $S$  defined in Section 2.3 that will be useful in the proofs. As mentioned in Section 2.3, the subset of parameters  $(\mathbb{R}^E \times \mathbb{R}^B) \setminus S$  is close to the notion of ‘admissible’ parameter in [37], but is slightly larger since the condition  $w_{\bullet \rightarrow v} \neq 0$  is replaced by  $(w_{\bullet \rightarrow v}, b_v) \neq (0, 0)$ , for each hidden neuron  $v$ .

**Proposition 14.** *Let  $\theta \in \mathbb{R}^E \times \mathbb{R}^B$ . The following statements are equivalent.*

- i)  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ .
- ii) For all  $l \in \llbracket 1, L-1 \rrbracket$  and all  $v_l \in V_l$ , there exist  $l' \in \llbracket 0, l \rrbracket$ , a path  $p_i = (v_{l'}, \dots, v_l) \in V_{l'} \times \dots \times V_l$  and a path  $p_o = (v_l, \dots, v_L) \in V_l \times \dots \times V_L$  such that
 
$$\theta_{p_i} \neq 0 \quad \text{and} \quad \theta_{p_o} \neq 0.$$
- iii) For all  $l \in \llbracket 1, L-1 \rrbracket$  and all  $v_l \in V_l$ , there exist  $l' \in \llbracket 0, l \rrbracket$ , a path  $p = (v_{l'}, \dots, v_l, \dots, v_{L-1}) \in \mathcal{P}_{l'}$  and  $v_L \in V_L$  such that
 
$$\phi_{p, v_L}(\theta) \neq 0.$$

*Proof.* Let us show successively that i)  $\Rightarrow$  ii), ii)  $\Rightarrow$  iii) and iii)  $\Rightarrow$  i).

i)  $\Rightarrow$  ii) Let  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ . Let us show ii) holds.

Let  $l \in \llbracket 1, L \rrbracket$  and  $v_l \in V_l$ . To form a path  $p_i$  satisfying the condition, we follow the procedure:

```

 $p_i \leftarrow (v_l)$ 
 $k \leftarrow l$ 
while  $k \geq 1$  and  $b_k = 0$  do
   $\exists v_{k-1} \in V_{k-1}, w_{v_{k-1} \rightarrow v_k} \neq 0$ 
   $p_i \leftarrow (v_{k-1}, p_i)$ 
   $k \leftarrow k - 1$ 
end while
 $l' \leftarrow k$ 

```

The existence of  $v_{k-1}$  in the loop is guaranteed by the fact that  $\theta \notin S$  and  $b_k = 0$  in the condition of the while loop. In the end, we obtain a path  $p_i = (v_{l'}, \dots, v_l)$  with either  $l' > 0$  and  $b_{l'} \neq 0$ , or  $l' = 0$ . In both cases, we have by construction

$$\theta_{p_i} \neq 0.$$

We do similarly the other way to form a path  $p_o = (v_l, \dots, v_L)$ . We follow the procedure:

$p_o \leftarrow (v_l)$   
 $k \leftarrow l$   
**while**  $k \leq L - 1$  **do**  
 $\exists v_{k+1} \in V_{k+1}, w_{v_k \rightarrow v_{k+1}} \neq 0$   
 $p_o \leftarrow (p_o, v_{k+1})$   
 $k \leftarrow k + 1$   
**end while**

The existence of  $v_{k+1}$  in the loop is guaranteed by the fact that  $\theta \notin S$ . In the end, we obtain a path  $p_o = (v_l, \dots, v_L)$  satisfying by construction

$$\theta_{p_o} \neq 0.$$

ii)  $\rightarrow$  iii) Let  $l \in \llbracket 1, L - 1 \rrbracket$  and  $v_l \in V_l$ . There exist  $l' \in \llbracket 0, l \rrbracket$ , a path  $p_i = (v_{l'}, \dots, v_l) \in V_{l'} \times \dots \times V_l$  and a path  $p_o = (v_l, \dots, v_L) \in V_l \times \dots \times V_L$  such that

$$\theta_{p_i} \neq 0 \quad \text{and} \quad \theta_{p_o} \neq 0.$$

Denoting  $p = (v_{l'}, \dots, v_l, \dots, v_{L-1})$ , we have

$$\phi_{p, v_L}(\theta) = \theta_{p_i} \theta_{p_o} \neq 0.$$

iii)  $\rightarrow$  i) Let us show the contrapositive: let  $\theta \in S$ , and let us show the statement iii) is not true. Indeed, if  $\theta \in S$ , there exist  $l \in \llbracket 1, L - 1 \rrbracket$  and  $v_l \in V_l$  such that  $(w_{\bullet \rightarrow v_l}, b_{v_l}) = (0, 0)$  or  $w_{v_l \rightarrow \bullet} = 0$ . Consider a path  $p = (v_{l'}, \dots, v_l, \dots, v_{L-1})$  and  $v_L \in V_L$ . We have

$$\phi_{p, v_L}(\theta) = \begin{cases} b_{v_{l'}} w_{v_{l'} \rightarrow v_{l'+1}} \dots w_{v_{l-1} \rightarrow v_l} w_{v_l \rightarrow v_{l+1}} \dots w_{v_{L-1} \rightarrow v_L} & \text{if } l' \geq 1 \\ w_{v_0 \rightarrow v_1} \dots w_{v_{l-1} \rightarrow v_l} w_{v_l \rightarrow v_{l+1}} \dots w_{v_{L-1} \rightarrow v_L} & \text{if } l' = 0. \end{cases}$$

If  $(w_{\bullet \rightarrow v_l}, b_{v_l}) = (0, 0)$ , either  $l' = l$  and  $b_{v_{l'}} = 0$  so  $\phi_{p, v_L}(\theta) = 0$ , or  $l' < l$  and since  $w_{v_{l-1} \rightarrow v_l} = 0$ , we have  $\phi_{p, v_L}(\theta) = 0$ .

If  $w_{v_l \rightarrow \bullet} = 0$ ,  $w_{v_l \rightarrow v_{l+1}} = 0$  so  $\phi_{p, v_L}(\theta) = 0$ . Thus iii) is not satisfied.  $\square$

We restate and prove Proposition 4.

**Proposition 15.** For all  $\theta, \tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ , we have

$$\theta \stackrel{R}{\sim} \tilde{\theta} \implies \phi(\theta) = \phi(\tilde{\theta}),$$

and thus in particular

$$\theta \sim \tilde{\theta} \implies \phi(\theta) = \phi(\tilde{\theta}).$$

*Proof.* Let  $\theta, \tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$  such that  $\theta \stackrel{R}{\sim} \tilde{\theta}$ . There exists a family  $(\lambda^0, \dots, \lambda^L) \in (\mathbb{R}^*)^{V_0} \times \dots \times (\mathbb{R}^*)^{V_L}$ , with  $\lambda^0 = \mathbb{1}_{V_0}$  and  $\lambda^L = \mathbb{1}_{V_L}$ , such that for all  $l \in \llbracket 1, L \rrbracket$ , for all  $(v_{l-1}, v_l) \in V_{l-1} \times V_l$ , (6) holds. We consider first a path  $p = (v_0, \dots, v_{L-1}) \in \mathcal{P}_0$  and  $v_L \in V_L$ . Using (6) and the fact that  $\lambda_{v_0}^0 = \lambda_{v_L}^L = 1$ , we have

$$\phi_{p, v_L}(\theta) = \prod_{l=1}^L w_{v_{l-1} \rightarrow v_l} = \prod_{l=1}^L \frac{\lambda_{v_l}^l}{\lambda_{v_{l-1}}^{l-1}} \tilde{w}_{v_{l-1} \rightarrow v_l} = \frac{\lambda_{v_L}^L}{\lambda_{v_0}^0} \prod_{l=1}^L \tilde{w}_{v_{l-1} \rightarrow v_l} = \phi_{p, v_L}(\tilde{\theta}).$$

Similarly, for  $l \in \llbracket 1, L - 1 \rrbracket$  and a path  $p = (v_l, \dots, v_{L-1}) \in \mathcal{P}_l$ , and for all  $v_L \in V_L$ , we have, using (6) and the fact that  $\lambda_{v_L}^L = 1$ ,

$$\begin{aligned} \phi_{p, v_L}(\theta) &= b_{v_l} \prod_{l'=l+1}^L w_{v_{l'-1} \rightarrow v_{l'}} = \lambda_{v_l}^l \tilde{b}_{v_l} \prod_{l'=l+1}^L \frac{\lambda_{v_{l'}}^{l'}}{\lambda_{v_{l'-1}}^{l'-1}} \tilde{w}_{v_{l'-1} \rightarrow v_{l'}} = \lambda_{v_L}^L \tilde{b}_{v_l} \prod_{l'=l+1}^L \tilde{w}_{v_{l'-1} \rightarrow v_{l'}} \\ &= \phi_{p, v_L}(\tilde{\theta}). \end{aligned}$$

Finally, for  $p = \beta$  and  $v_L \in V_L$ , we have

$$\phi_{p, v_L}(\theta) = b_{v_L} = \lambda_{v_L}^L \tilde{b}_{v_L} = \tilde{b}_{v_L} = \phi_{p, v_L}(\tilde{\theta}).$$

This shows  $\phi(\theta) = \phi(\tilde{\theta})$ .

For the second implication, we simply use the fact that if  $\theta \sim \tilde{\theta}$ , in particular,  $\theta \stackrel{R}{\sim} \tilde{\theta}$ .  $\square$

**Corollary 16.** The set  $(\mathbb{R}^E \times \mathbb{R}^B) \setminus S$  is stable by rescaling equivalence: if  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ , and  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$  satisfies  $\theta \stackrel{R}{\sim} \tilde{\theta}$ , then  $\tilde{\theta} \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ .

*Proof.* Let  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$  and  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$  such that  $\theta \stackrel{R}{\sim} \tilde{\theta}$ . Proposition 15 shows that  $\phi(\tilde{\theta}) = \phi(\theta)$ .

Let  $l \in \llbracket 1, L \rrbracket$  and  $v \in V_l$ . Since  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ , according to Proposition 14 there exists  $l' \in \llbracket 0, l \rrbracket$ , a path  $p = (v_{l'}, \dots, v_l, \dots, v_{L-1})$  and  $v_L \in V_L$  such that  $\phi_{p, v_L}(\theta) \neq 0$ . We have

$$\phi_{p, v_L}(\tilde{\theta}) = \phi_{p, v_L}(\theta) \neq 0,$$

and since this is true for any  $l \in \llbracket 1, L \rrbracket$  and  $v \in V_l$ , Proposition 14 shows that  $\tilde{\theta} \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ .  $\square$

We restate and prove Proposition 5.

**Proposition 17.** For all  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ , for all  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ ,

$$\phi(\theta) = \phi(\tilde{\theta}) \implies \theta \stackrel{R}{\sim} \tilde{\theta}.$$

*Proof.* Let us choose  $(\lambda^0, \dots, \lambda^L) \in (\mathbb{R}^*)^{V_0} \times \dots \times (\mathbb{R}^*)^{V_L}$  as follows. For all  $l \in \llbracket 1, L-1 \rrbracket$  and all  $v_l \in V_l$ , since  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ , Proposition 14 shows that there exists a path  $p_o(v_l) = (v_l, \dots, v_L) \in V_l \times \dots \times V_L$  such that  $\theta_{p_o(v_l)} \neq 0$ . Let us define  $\lambda^0 = \mathbb{1}_{V_0}$ ,  $\lambda^L = \mathbb{1}_{V_L}$  and for all  $l \in \llbracket 1, L-1 \rrbracket$ ,

$$\lambda_{v_l}^l = \frac{\tilde{\theta}_{p_o(v_l)}}{\theta_{p_o(v_l)}}.$$

The value of  $\lambda_{v_l}^l$  a priori depends on the choice of the path  $p_o(v_l)$ , but the first of the two following facts, that we are going to prove, shows it only depends on  $v_l$ , since in (20),  $p_i$  does not depend on  $p_o(v_l)$ .

- For all  $l \in \llbracket 0, L \rrbracket$ , for all  $v_l \in V_l$ , for any  $l' \in \llbracket 0, l \rrbracket$  and any  $p_i = (v_{l'}, \dots, v_l) \in V_{l'} \times \dots \times V_l$ ,

$$\theta_{p_i} = \lambda_{v_l}^l \tilde{\theta}_{p_i}. \quad (20)$$

- For all  $l \in \llbracket 0, L \rrbracket$ , for all  $v_l \in V_l$ ,  $\lambda_{v_l}^l \neq 0$ .

Indeed, let  $l \in \llbracket 0, L \rrbracket$  and let us consider  $l' \in \llbracket 0, l \rrbracket$  and a path  $p_i = (v_{l'}, \dots, v_l) \in V_{l'} \times \dots \times V_l$ . Let  $v_{l+1}, \dots, v_L \in V_{l+1} \times \dots \times V_L$  such that  $p_o(v_l) = (v_l, v_{l+1}, \dots, v_L)$ . Let  $p = (v_{l'}, \dots, v_l, \dots, v_{L-1}) \in \mathcal{P}_{l'}$  so that  $p_i \cup p_o(v_l) = p \cup (v_L)$ . We have by hypothesis

$$\theta_{p_i} \theta_{p_o(v_l)} = \phi_{p, v_L}(\theta) = \phi_{p, v_L}(\tilde{\theta}) = \tilde{\theta}_{p_i} \tilde{\theta}_{p_o(v_l)},$$

thus

$$\theta_{p_i} = \frac{\tilde{\theta}_{p_o(v_l)}}{\theta_{p_o(v_l)}} \tilde{\theta}_{p_i} = \lambda_{v_l}^l \tilde{\theta}_{p_i},$$

which proves the first point. To prove the second point, we simply use Proposition 14 to consider a path  $p_i$  such that  $\theta_{p_i} \neq 0$ , and (20) shows that  $\lambda_{v_l}^l \neq 0$ .

Let us now prove the rescaling equivalence. Let  $l \in \llbracket 1, L \rrbracket$ , and let  $(v_{l-1}, v_l) \in V_{l-1} \times V_l$ . Let us consider, thanks to Proposition 14,  $l' \in \llbracket 0, l-1 \rrbracket$  and a path  $p_i = (v_{l'}, \dots, v_{l-1}) \in V_{l'} \times \dots \times V_{l-1}$  such that  $\theta_{p_i} \neq 0$ . The relation (20) shows we also have  $\tilde{\theta}_{p_i} \neq 0$ . Let  $p'_i = p_i \cup (v_l)$ . Using (20) with  $\theta_{p'_i}$  we have

$$\theta_{p_i} w_{v_{l-1} \rightarrow v_l} = \theta_{p'_i} = \lambda_{v_l}^l \tilde{\theta}_{p'_i} = \lambda_{v_l}^l \tilde{\theta}_{p_i} \tilde{w}_{v_{l-1} \rightarrow v_l}.$$

At the same time, using (20) with  $\theta_{p_i}$  we have,

$$\theta_{p_i} w_{v_{l-1} \rightarrow v_l} = \lambda_{v_{l-1}}^{l-1} \tilde{\theta}_{p_i} w_{v_{l-1} \rightarrow v_l},$$

so combining both equalities, we have

$$\lambda_{v_l}^l \tilde{\theta}_{p_i} \tilde{w}_{v_{l-1} \rightarrow v_l} = \lambda_{v_{l-1}}^{l-1} \tilde{\theta}_{p_i} w_{v_{l-1} \rightarrow v_l}.$$

Using the fact that  $\tilde{\theta}_{p_i} \neq 0$  and  $\lambda_{v_{l-1}}^{l-1} \neq 0$ , we finally obtain, for all  $l \in \llbracket 1, L \rrbracket$  and all  $(v_{l-1}, v_l) \in V_{l-1} \times V_l$ :

$$w_{v_{l-1} \rightarrow v_l} = \frac{\lambda_{v_l}^l}{\lambda_{v_{l-1}}^{l-1}} \tilde{w}_{v_{l-1} \rightarrow v_l}.$$

For all  $l \in \llbracket 1, L \rrbracket$  and all  $v_l \in V_l$ , using (20) with  $p_i = (v_l)$ , we obtain

$$b_{v_l} = \lambda_{v_l}^l \tilde{b}_{v_l}.$$

This shows that (6) is satisfied for all  $(v_{l-1}, v_l) \in V_{l-1} \times V_l$ , and thus  $\theta \stackrel{R}{\sim} \tilde{\theta}$ .  $\square$

The following proposition is useful in the proof of Theorem 25 and allows to improve identifiability modulo rescaling into identifiability modulo positive rescaling.

**Proposition 18.** *For all  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ , there exists  $\epsilon > 0$  such that for all  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ ,*

$$\|\theta - \tilde{\theta}\|_\infty < \epsilon \text{ and } \theta \stackrel{R}{\sim} \tilde{\theta} \implies \theta \sim \tilde{\theta}.$$

*Proof.* Let  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ . We define

$$\epsilon = \min \left( \{|w_{v \rightarrow v'}|, v \rightarrow v' \in E \text{ and } w_{v \rightarrow v'} \neq 0\} \cup \{|b_v|, v \in B \text{ and } b_v \neq 0\} \right).$$

Let  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$  such that  $\|\theta - \tilde{\theta}\|_\infty < \epsilon$  and  $\theta \stackrel{R}{\sim} \tilde{\theta}$ . To prove  $\theta \sim \tilde{\theta}$ , we simply have to prove  $\text{sign}(\theta) = \text{sign}(\tilde{\theta})$ . There exists  $(\lambda^0, \dots, \lambda^L) \in (\mathbb{R}^*)^{V_0} \times \dots \times (\mathbb{R}^*)^{V_L}$ , with  $\lambda^0 = \mathbb{1}_{V_0}$  and  $\lambda^L = \mathbb{1}_{V_L}$ , such that, for all  $l \in \llbracket 1, L \rrbracket$ , for all  $(v_{l-1}, v_l) \in V_{l-1} \times V_l$ , (6) holds. Let us show that  $\text{sign}(\theta) = \text{sign}(\tilde{\theta})$ .

Indeed, let  $l \in \llbracket 1, L \rrbracket$ , and let  $(v, v') \in V_{l-1} \times V_l$ . If  $w_{v \rightarrow v'} \neq 0$ , then since  $|w_{v \rightarrow v'} - \tilde{w}_{v \rightarrow v'}| < \epsilon$  and by definition  $\epsilon \leq |w_{v \rightarrow v'}|$ , we have  $\text{sign}(w_{v \rightarrow v'}) = \text{sign}(\tilde{w}_{v \rightarrow v'})$ . Otherwise, if  $w_{v \rightarrow v'} = 0$ , (6) shows that we have

$$\tilde{w}_{v \rightarrow v'} = \frac{\lambda_v^{l-1}}{\lambda_{v'}^l} w_{v \rightarrow v'} = 0,$$

so we still have  $\text{sign}(w_{v \rightarrow v'}) = \text{sign}(\tilde{w}_{v \rightarrow v'})$ .

Now let  $l \in \llbracket 1, L \rrbracket$  and let  $v \in V_l$ . Similarly, if  $b_v \neq 0$ , we have  $|b_v - \tilde{b}_v| < \epsilon \leq |b_v|$ , so  $\text{sign}(b_v) = \text{sign}(\tilde{b}_v)$ , and if  $b_v = 0$ , we have

$$\tilde{b}_v = \frac{b_v}{\lambda_v^l} = 0,$$

so again  $\text{sign}(b_v) = \text{sign}(\tilde{b}_v)$ .

This shows  $\text{sign}(\theta) = \text{sign}(\tilde{\theta})$ , so  $\theta \sim \tilde{\theta}$ .  $\square$

## C The smooth manifold structure of $\Sigma_1^*$

In this section, we prove Theorem 6, which is restated as Theorem 24. Before doing so, we establish intermediary results, some of which are evoked in Section 3.

Let us discuss the cardinal of  $F_\theta$  defined in Section 3. The set  $F_\theta$  is obtained by removing the edges of the form  $v \rightarrow s_{\max}^\theta(v)$  for  $v \in V_1 \cup \dots \cup V_{L-1}$ . Note that we do not remove the edges of the form  $v \rightarrow s_{\max}^\theta(v)$  for  $v \in V_0$ . For all  $l \in \llbracket 1, L-1 \rrbracket$ , there are precisely  $N_l$  edges of the form  $(v, s_{\max}^\theta(v))$  with  $v \in V_l$ , so

$$\begin{aligned} |F_\theta| &= |E| - (N_1 + \dots + N_{L-1}) \\ &= N_0 N_1 + \dots + N_{L-1} N_L - N_1 - \dots - N_{L-1}. \end{aligned}$$

As a consequence, since  $|B| = N_1 + \dots + N_L$ , we have in particular

$$\begin{aligned} |F_\theta| + |B| &= N_0 N_1 + \dots + N_{L-1} N_L - N_1 - \dots - N_{L-1} + N_1 + \dots + N_L \\ &= N_0 N_1 + \dots + N_{L-1} N_L + N_L. \end{aligned} \quad (21)$$

The following proposition is a first step towards Proposition 20, which states that  $\psi^\theta$  is a homeomorphism.

**Proposition 19.** *For all  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ , the function  $\psi^\theta : U_\theta \rightarrow \mathbb{R}^{\mathcal{P} \times V_L}$  is injective.*

*Proof.* Let  $\tau, \tilde{\tau} \in U_\theta$  such that  $\psi^\theta(\tau) = \psi^\theta(\tilde{\tau})$ . Let us show  $\tau = \tilde{\tau}$ . We have  $\phi(\rho_\theta(\tau)) = \phi(\rho_\theta(\tilde{\tau}))$  and by definition of  $U_\theta$ ,  $\rho_\theta(\tau) \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ , so by Proposition 17 we have the rescaling equivalence

$$\rho_\theta(\tau) \stackrel{R}{\sim} \rho_\theta(\tilde{\tau}).$$

By definition of the rescaling equivalence, in its formulation (6), there exists  $(\lambda^0, \dots, \lambda^L) \in (\mathbb{R}^*)^{V_0} \times \dots \times (\mathbb{R}^*)^{V_L}$ , with  $\lambda^0 = \mathbb{1}_{V_0}$  and  $\lambda^L = \mathbb{1}_{V_L}$ , such that, for all  $l \in \llbracket 1, L \rrbracket$ , for all  $(v_{l-1}, v_l) \in V_{l-1} \times V_l$ ,

$$\begin{cases} \rho_\theta(\tau)_{v_{l-1} \rightarrow v_l} = \frac{(\lambda^l)_{v_l}}{(\lambda^{l-1})_{v_{l-1}}} \rho_\theta(\tilde{\tau})_{v_{l-1} \rightarrow v_l} \\ b_{v_l} = \lambda^l_{v_l} \tilde{b}_{v_l}. \end{cases} \quad (22)$$

Let  $l \in \llbracket 2, L \rrbracket$  and let  $v_{l-1} \in V_{l-1}$ . Let  $v_l = s_{\max}^\theta(v_{l-1})$ . According to (22) we have

$$\rho_\theta(\tau)_{v_{l-1} \rightarrow v_l} = \frac{(\lambda^l)_{v_l}}{(\lambda^{l-1})_{v_{l-1}}} \rho_\theta(\tilde{\tau})_{v_{l-1} \rightarrow v_l}.$$

But since  $v_l = s_{\max}^\theta(v_{l-1})$  and  $v_{l-1} \in V_{l-1}$  with  $l-1 \in \llbracket 1, L-1 \rrbracket$ , we have  $v_{l-1} \rightarrow v_l \in E \setminus F_\theta$ , so by definition of  $\rho_\theta$  in (9),

$$\rho_\theta(\tau)_{v_{l-1} \rightarrow v_l} = w_{v_{l-1} \rightarrow v_l} = \rho_\theta(\tilde{\tau})_{v_{l-1} \rightarrow v_l} \neq 0,$$

so  $\frac{(\lambda^l)_{v_l}}{(\lambda^{l-1})_{v_{l-1}}} = 1$ .

We have shown that for all  $l \in \llbracket 2, L \rrbracket$ , for all  $v_{l-1} \in V_{l-1}$ , there exists  $v_l \in V_l$  such that

$$(\lambda^{l-1})_{v_{l-1}} = (\lambda^l)_{v_l}.$$

As a consequence, if  $l$  is such that  $\lambda^l = \mathbb{1}_{V_l}$ , then  $\lambda^{l-1} = \mathbb{1}_{V_{l-1}}$ .

Starting from  $\lambda^L = \mathbb{1}_{V_L}$ , this shows by induction that for all  $l \in \llbracket 1, L \rrbracket$ ,

$$\lambda^l = \mathbb{1}_{V_l}.$$

By hypothesis we also have  $\lambda^0 = \mathbb{1}_{V_0}$ . Using (22), this shows that

$$\rho_\theta(\tau) = \rho_\theta(\tilde{\tau}).$$

The injectivity of  $\rho_\theta$  allows us to conclude that

$$\tau = \tilde{\tau}.$$

□

The following proposition shows, as mentioned in Section 3, that  $\psi^\theta$  is a homeomorphism. This is a necessary step to prove that  $(V_\theta, (\psi^\theta)^{-1})_{\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S}$  is an atlas of  $\Sigma_1^*$ .

**Proposition 20.** *For all  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ ,  $\psi^\theta$  is a homeomorphism from  $U_\theta$  onto its image  $V_\theta$ .*

*Proof.* We already know from Proposition 19 that  $\psi^\theta$  is injective, so we need to prove that  $\psi^\theta$  is continuous and its inverse is continuous. The function  $\rho_\theta$  is affine and  $\phi$  is a polynomial function, so the function  $\psi^\theta = \phi \circ \rho_\theta$  is a polynomial function, and in particular it is continuous.

To prove that  $(\psi^\theta)^{-1}$  is continuous, we consider a sequence  $(\tau_n)$  taking values in  $U_\theta$  and  $\tau \in U_\theta$  such that  $\psi^\theta(\tau_n) \rightarrow \psi^\theta(\tau)$ , and we want to show that  $\tau_n \rightarrow \tau$ .

Let us first show that for all  $v \in B$ ,  $(\tau_n)_v \rightarrow \tau_v$ . Indeed, let  $l \in \llbracket 1, L \rrbracket$  and let  $v_l \in V_l$ , so that  $v_l$  is an arbitrary element of  $B$ . Let us define  $v_{l+1} = s_{\max}^\theta(v_l)$ , then  $v_{l+2} = s_{\max}^\theta(v_{l+1})$  and so on up to  $v_L = s_{\max}^\theta(v_{L-1})$ . Since for all  $l' \in \llbracket l, L-1 \rrbracket$ ,  $v_{l'+1} = s_{\max}^\theta(v_{l'})$ , by definition of  $F_\theta$  and  $\rho_\theta$  (see (8) and (9)), we have

$$\rho_\theta(\tau_n)_{v_{l'} \rightarrow v_{l'+1}} = w_{v_{l'} \rightarrow v_{l'+1}}, \quad (23)$$

and

$$\rho_\theta(\tau)_{v_{l'} \rightarrow v_{l'+1}} = w_{v_{l'} \rightarrow v_{l'+1}}. \quad (24)$$

In particular, since  $\theta \notin S$ , for all  $l' \in \llbracket l, L-1 \rrbracket$  we have  $w_{v_{l'} \rightarrow \bullet} \neq 0$ , so by definition of  $s_{\max}^\theta$ ,  $w_{v_{l'} \rightarrow v_{l'+1}} \neq 0$ . We thus have

$$w_{v_l \rightarrow v_{l+1}} \cdots w_{v_{L-1} \rightarrow v_L} \neq 0. \quad (25)$$

If we denote  $p = (v_l, \dots, v_{L-1})$ , we have, using the definition of  $\phi$  and (23),

$$\psi_{p, v_L}^\theta(\tau_n) = (\tau_n)_{v_l} w_{v_l \rightarrow v_{l+1}} \cdots w_{v_{L-1} \rightarrow v_L}$$

and using (24),

$$\psi_{p, v_L}^\theta(\tau) = (\tau)_{v_l} w_{v_l \rightarrow v_{l+1}} \cdots w_{v_{L-1} \rightarrow v_L}.$$

Using (25) and the fact that

$$\psi^\theta(\tau_n) \rightarrow \psi^\theta(\tau),$$

we conclude that

$$(\tau_n)_{v_l} \rightarrow \tau_{v_l}.$$

Let us now prove that for all  $(v, v') \in E$ ,  $(\tau_n)_{v \rightarrow v'} \rightarrow \tau_{v \rightarrow v'}$ . Let us show by induction on  $l \in \llbracket 1, L \rrbracket$  the following hypothesis

$$\forall l' \in \llbracket 1, l \rrbracket, \quad \forall (v, v') \in (V_{l'-1} \times V_{l'}) \cap F_\theta, \quad (\tau_n)_{v \rightarrow v'} \rightarrow \tau_{v \rightarrow v'}. \quad (H_l)$$

**Initialization.** Let  $(v_0, v_1) \in (V_0 \times V_1) \cap F_\theta$ . We define  $v_2 = s_{\max}^\theta(v_1)$ , then we define  $v_3 = s_{\max}^\theta(v_2)$ , and so on up to  $v_L = s_{\max}^\theta(v_{L-1})$ . Let  $p = (v_0, \dots, v_{L-1}) \in \mathcal{P}$ .

As above, using the definition of  $\rho_\theta$ ,  $F_\theta$  and  $\phi$ , we have

$$\psi_{p, v_L}^\theta(\tau_n) = (\tau_n)_{v_0 \rightarrow v_1} w_{v_1 \rightarrow v_2} \cdots w_{v_{L-1} \rightarrow v_L}$$

and

$$\psi_{p, v_L}^\theta(\tau) = (\tau)_{v_0 \rightarrow v_1} w_{v_1 \rightarrow v_2} \cdots w_{v_{L-1} \rightarrow v_L},$$

and since  $\theta \notin S$ , we also have, as above,

$$w_{v_1 \rightarrow v_2} \cdots w_{v_{L-1} \rightarrow v_L} \neq 0. \quad (26)$$

Since

$$\psi^\theta(\tau_n) \rightarrow \psi^\theta(\tau)$$

we conclude using (26) that

$$(\tau_n)_{v_0 \rightarrow v_1} \rightarrow \tau_{v_0 \rightarrow v_1}.$$

We have shown  $H_1$ .

**Induction step.** Let  $l \in \llbracket 2, L \rrbracket$  and let us assume that  $H_{l-1}$  holds.

Let  $(v_{l-1}, v_l) \in (V_{l-1} \times V_l) \cap F_\theta$ . We define  $v_{l+1} = s_{\max}^\theta(v_l)$ ,  $v_{l+2} = s_{\max}^\theta(v_{l+1})$ , and so on up to  $v_L = s_{\max}^\theta(v_{L-1})$ . Let us denote  $p_o = (v_l, \dots, v_L)$ . Recalling the notation defined in (19), we have

$$\rho_\theta(\tau_n)_{p_o} = w_{v_l \rightarrow v_{l+1}} \cdots w_{v_{L-1} \rightarrow v_L} = \rho_\theta(\tau)_{p_o} \neq 0. \quad (27)$$

At the same time, since  $\tau \in U_\theta$ , Proposition 14 shows there exist  $l' \in \llbracket 0, l-1 \rrbracket$  and a path  $p_i = (v_{l'}, \dots, v_{l-2}, v_{l-1})$  such that

$$\rho_\theta(\tau)_{p_i} \neq 0. \quad (28)$$

If  $l' \geq 1$ , we have shown in the first part of the proof that  $(\tau_n)_{v_{l'}} \rightarrow \tau_{v_{l'}}$ . Moreover, whatever the value of  $l'$  is, for  $k \in \llbracket l', l-2 \rrbracket$ , if  $(v_k, v_{k+1}) \in E \setminus F_\theta$ ,

$$\rho_\theta(\tau_n)_{v_k \rightarrow v_{k+1}} = w_{v_k \rightarrow v_{k+1}} = \rho_\theta(\tau)_{v_k \rightarrow v_{k+1}},$$



and if  $(v_k, v_{k+1}) \in F_\theta$ , according to  $H_{l-1}$ ,

$$\rho_\theta(\tau_n)_{v_k \rightarrow v_{k+1}} = (\tau_n)_{v_k \rightarrow v_{k+1}} \longrightarrow \tau_{v_k \rightarrow v_{k+1}} = \rho_\theta(\tau)_{v_k \rightarrow v_{k+1}}.$$

We therefore have

$$\rho_\theta(\tau_n)_{p_i} \longrightarrow \rho_\theta(\tau)_{p_i}, \quad (29)$$

and in particular, since  $\rho_\theta(\tau)_{p_i} \neq 0$ , there exists  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$ ,

$$\rho_\theta(\tau_n)_{p_i} \neq 0. \quad (30)$$

We can write

$$\psi_{p, v_L}^\theta(\tau_n) = \rho_\theta(\tau_n)_{p_i} (\tau_n)_{v_{l-1} \rightarrow v_l} \rho_\theta(\tau_n)_{p_o}$$

and

$$\psi_{p, v_L}^\theta(\tau) = \rho_\theta(\tau)_{p_i} (\tau)_{v_{l-1} \rightarrow v_l} \rho_\theta(\tau)_{p_o},$$

so using (27), (30) and (29), we have

$$(\tau_n)_{v_{l-1} \rightarrow v_l} = \frac{\psi_{p, v_L}^\theta(\tau_n)}{\rho_\theta(\tau_n)_{p_i} \rho_\theta(\tau_n)_{p_o}} \longrightarrow \frac{\psi_{p, v_L}^\theta(\tau)}{\rho_\theta(\tau)_{p_i} \rho_\theta(\tau)_{p_o}} = \tau_{v_{l-1} \rightarrow v_l}.$$

We have shown  $H_l$ , which concludes the induction step.

In particular,  $H_L$  is satisfied, and finally  $\tau_n \rightarrow \tau$ .

This shows that  $\psi^\theta$  is a homeomorphism.  $\square$

The following lemma is necessary for the proof of Proposition 22.

**Lemma 21.** Let  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ . Let  $(v, v') \in E$  (resp.  $v \in B$ ). If  $w_{v \rightarrow v'} \neq 0$  (resp.  $b_v \neq 0$ ), then there exists  $\epsilon > 0$  such that for all  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ , if  $\|\phi(\theta) - \phi(\tilde{\theta})\|_\infty < \epsilon$ , then  $\tilde{w}_{v \rightarrow v'} \neq 0$  (resp.  $\tilde{b}_v \neq 0$ ).

*Proof.* Let  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$  and  $(v, v') \in E$  such that  $w_{v \rightarrow v'} \neq 0$ . Denote  $l \in \llbracket 0, L-1 \rrbracket$  such that  $v \in V_l$ . If  $l = 0$ , we take  $p_i = (v)$  so that by convention  $\theta_{p_i} = 1 \neq 0$ , and if  $l \geq 1$ , we use Proposition 14 which states that there exists  $l' \in \llbracket 0, l-1 \rrbracket$  and a path  $p_i = (v_{l'}, \dots, v_{l-2}, v)$  such that  $\theta_{p_i} \neq 0$ . Similarly, if  $l = L-1$ , we take  $p_o = (v')$  so that by convention  $\theta_{p_o} = 1 \neq 0$  and if  $l < L-1$ , we use Proposition 14 which states that there exists a path  $p_o = (v', v_{l+1}, \dots, v_L)$  such that  $\theta_{p_o} \neq 0$ . If we denote

$$p = \begin{cases} (v, v', v_{l+2}, \dots, v_{L-1}) & \text{if } l = 0 \\ (v_{l'}, \dots, v_{l-1}, v, v') & \text{if } l = L-1 \\ (v_{l'}, \dots, v_{l-1}, v, v', v_{l+2}, \dots, v_{L-1}) & \text{otherwise,} \end{cases}$$

we have

$$\phi_{p, v_L}(\theta) = \theta_{p_i} w_{v \rightarrow v'} \theta_{p_o} \neq 0.$$

We define  $\epsilon = |\phi_{p, v_L}(\theta)| > 0$ . For all  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$  such that  $\|\phi(\tilde{\theta}) - \phi(\theta)\|_\infty < \epsilon$  we have

$$\phi_{p, v_L}(\tilde{\theta}) \neq 0.$$

Since  $\phi_{p, v_L}(\tilde{\theta}) = \tilde{\theta}_{p_i} \tilde{w}_{v \rightarrow v'} \tilde{\theta}_{p_o}$ , this implies in particular that

$$\tilde{w}_{v \rightarrow v'} \neq 0.$$

The proof is similar in the case  $v \in B$  and  $b_v \neq 0$ .  $\square$

The following proposition, which states that for any  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ ,  $V_\theta = \psi^\theta(U_\theta)$  is open with respect to the topology induced on  $\Sigma_1^*$  by the standard topology of  $\mathbb{R}^{\mathcal{P} \times V_L}$ , is necessary to show that  $(V_\theta, (\psi^\theta)^{-1})_{\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S}$  is an atlas of  $\Sigma_1^*$ .

**Proposition 22.** For any  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ , for any  $\tau \in U_\theta$ , there exists  $\epsilon > 0$  such that

$$\Sigma_1^* \cap B_\infty(\psi^\theta(\tau), \epsilon) \subset V_\theta.$$

*Proof.* Let us first construct  $\epsilon$  and then consider an element of the set on the left of the inclusion and prove it belongs to  $V_\theta$ . Let  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$  and  $\tau \in U_\theta$ . For all  $l \in \llbracket 1, L-1 \rrbracket$ , for all  $v \in V_l$ , by definition of  $F_\theta$  and  $\rho_\theta$ , we have  $\rho_\theta(\tau)_{v \rightarrow s_{\max}^\theta(v)} = w_{v \rightarrow s_{\max}^\theta(v)}$ , and since  $\theta \notin S$ , by definition of  $s_{\max}^\theta$ ,  $w_{v \rightarrow s_{\max}^\theta(v)} \neq 0$ , so according to Lemma 21 there exists  $\epsilon_v > 0$  such that for all  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ ,

$$\|\phi(\rho_\theta(\tau)) - \phi(\tilde{\theta})\|_\infty < \epsilon_v \implies \tilde{w}_{v \rightarrow s_{\max}^\theta(v)} \neq 0.$$

Let  $\epsilon = \min_{v \in V_1 \cup \dots \cup V_{L-1}} \epsilon_v$ .

Let us now show the inclusion: let  $\tilde{\theta} \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$  such that  $\|\phi(\rho_\theta(\tau)) - \phi(\tilde{\theta})\|_\infty < \epsilon$ , and let us show that  $\phi(\tilde{\theta}) \in V_\theta$ . Notice first that for all  $l \in \llbracket 1, L-1 \rrbracket$  and  $v \in V_l$ , by definition of  $\epsilon$ ,  $w_{v \rightarrow s_{\max}^\theta(v)} \neq 0$  and  $\tilde{w}_{v \rightarrow s_{\max}^\theta(v)} \neq 0$ . We are going to define  $\tilde{\tau} \in U_\theta$  such that  $\rho_\theta(\tilde{\tau}) \stackrel{R}{\sim} \tilde{\theta}$ , so that using Proposition 15,  $\psi^\theta(\tilde{\tau}) = \phi(\tilde{\theta})$ .

Let us define recursively a family  $(\lambda^0, \dots, \lambda^L) \in (\mathbb{R}^*)^{V_0} \times \dots \times (\mathbb{R}^*)^{V_L}$  as follows:

- we define  $\lambda^L = \mathbb{1}_{V_L}$ ;
- for all  $l \in \llbracket 1, L-1 \rrbracket$ , for all  $v \in V_l$ , we define

$$\lambda_v^l = \frac{\tilde{w}_{v \rightarrow s_{\max}^\theta(v)}}{w_{v \rightarrow s_{\max}^\theta(v)}} \lambda_{s_{\max}^\theta(v)}^{l+1}. \quad (31)$$

- we define finally  $\lambda^0 = \mathbb{1}_{V_0}$ .

Note that for all  $l \in \llbracket 0, L \rrbracket$  and for all  $v \in V_l$ ,  $\lambda_v^l \neq 0$ . Also note that for all  $l \in \llbracket 2, L \rrbracket$ , for all  $v \in V_{l-1}$ , reformulating (31) in a way that will be useful later, we have

$$\frac{\lambda_{s_{\max}^\theta(v)}^l}{\lambda_v^{l-1}} = \frac{w_{v \rightarrow s_{\max}^\theta(v)}}{\tilde{w}_{v \rightarrow s_{\max}^\theta(v)}}. \quad (32)$$

We then define  $\tilde{\tau} \in \mathbb{R}^{F_\theta} \times \mathbb{R}^B$  by:

- for all  $l \in \llbracket 1, L \rrbracket$ , for all  $(v, v') \in (V_{l-1} \times V_l) \cap F_\theta$ ,

$$\tilde{\tau}_{v \rightarrow v'} = \frac{\lambda_{v'}^l}{\lambda_v^{l-1}} \tilde{w}_{v \rightarrow v'}; \quad (33)$$

- for all  $l \in \llbracket 1, L \rrbracket$ , for all  $v \in V_l$ ,

$$\tilde{\tau}_v = \lambda_v^l \tilde{b}_v. \quad (34)$$

Let us show  $\rho_\theta(\tilde{\tau}) \stackrel{R}{\sim} \tilde{\theta}$ . Indeed, let  $l \in \llbracket 1, L \rrbracket$  and let  $(v, v') \in V_{l-1} \times V_l$ . If  $v \in V_0$  or  $v \in V_1 \cup \dots \cup V_{L-1}$  and  $v' \neq s_{\max}^\theta(v)$ , then by definition (8) of  $F_\theta$ , we have  $v \rightarrow v' \in F_\theta$ , so using (9) and (33) we have

$$\rho_\theta(\tilde{\tau})_{v \rightarrow v'} = \tilde{\tau}_{v \rightarrow v'} = \frac{\lambda_{v'}^l}{\lambda_v^{l-1}} \tilde{w}_{v \rightarrow v'}. \quad (35)$$

If  $v \in V_1 \cup \dots \cup V_{L-1}$  and  $v' = s_{\max}^\theta(v)$ , then by definition (8) of  $F_\theta$ , we have  $v \rightarrow v' \in E \setminus F_\theta$ , and since in that case,  $l \geq 2$ , using (9) and (32), we see that

$$\rho_\theta(\tilde{\tau})_{v \rightarrow v'} = w_{v \rightarrow v'} = \frac{\lambda_{v'}^l}{\lambda_v^{l-1}} \tilde{w}_{v \rightarrow v'}. \quad (36)$$

If  $v \in B$ , using (9) and (34), we have

$$\rho_\theta(\tilde{\tau})_v = \tilde{\tau}_v = \lambda_v^l \tilde{b}_v. \quad (37)$$

Equations (35), (36) and (37) prove that

$$\rho_\theta(\tilde{\tau}) \stackrel{R}{\sim} \tilde{\theta}.$$

Using Corollary 16, since  $\tilde{\theta} \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$  and  $\rho_\theta(\tilde{\tau}) \stackrel{R}{\sim} \tilde{\theta}$ , we also have  $\rho_\theta(\tilde{\tau}) \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ . Since, by definition,  $U_\theta = \rho_\theta^{-1}((\mathbb{R}^E \times \mathbb{R}^B) \setminus S)$ , we have  $\tilde{\tau} \in U_\theta$ . We have shown

$$\Sigma_1^* \cap B_\infty(\psi^\theta(\tau), \epsilon) \subset V_\theta.$$

□

The following proposition is necessary in order to show that  $(V_\theta, (\psi^\theta)^{-1})_{\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S}$  is an atlas of  $\Sigma_1^*$ .

**Proposition 23.** *For all  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ , the function  $\psi^\theta$  is  $C^\infty$  and its differential  $D\psi^\theta(\tau)$  is injective for all  $\tau \in U_\theta$ .*

*Proof.* Let  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ . First of all,  $\psi^\theta$  is a polynomial function as a composition of  $\phi$  and  $\rho_\theta$  which are both polynomial functions. So,  $\psi^\theta$  is  $C^\infty$ .

In order to show the injectivity of the differential  $D\psi^\theta(\tau)$  for all  $\tau \in U_\theta$ , let us compute the partial derivatives of  $\psi_{p, v_L}^\theta(\tau)$ . Let  $p \in \mathcal{P}$  and  $v_L \in V_L$ . Using the definition of  $\psi^\theta$  and  $\phi$ , three cases are possible.

Case 1. The path  $p$  is of the form  $(v_0, v_1, \dots, v_{L-1})$ . We have

$$\psi_{p, v_L}^\theta(\tau) = \rho_\theta(\tau)_{v_0 \rightarrow v_1} \cdots \rho_\theta(\tau)_{v_{L-1} \rightarrow v_L}.$$

Case 2. The path  $p$  is of the form  $(v_l, \dots, v_{L-1})$  with  $l \in \llbracket 1, L-1 \rrbracket$ . We have, for all  $\tau \in U_\theta$ ,

$$\psi_{p, v_L}^\theta(\tau) = \tau_{v_l} \rho_\theta(\tau)_{v_l \rightarrow v_{l+1}} \cdots \rho_\theta(\tau)_{v_{L-1} \rightarrow v_L}.$$

Case 3. For  $p = \beta$ , we have, for all  $\tau \in U_\theta$ ,

$$\psi_{p, v_L}^\theta(\tau) = \tau_{v_L}.$$

Let  $(v, v') \in F_\theta$ , and let us compute  $\frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_{v \rightarrow v'}}(\tau)$ .

Case 1. We have  $p = (v_0, \dots, v_{L-1}) \in \mathcal{P}_0$ . If  $\{v, v'\} \subset \{v_0, \dots, v_L\}$ , there exists  $l \in \llbracket 0, L-1 \rrbracket$  such that  $(v, v') = (v_l, v_{l+1})$ , in which case, since  $(v, v') \in F_\theta$ ,  $\rho_\theta(\tau)_{v_l \rightarrow v_{l+1}} = \tau_{v_l \rightarrow v_{l+1}}$  and

$$\frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_{v \rightarrow v'}}(\tau) = \prod_{\substack{k \in \llbracket 0, L-1 \rrbracket \\ k \neq l}} \rho_\theta(\tau)_{v_k \rightarrow v_{k+1}}. \quad (38)$$

Otherwise if  $\{v, v'\} \not\subset \{v_0, \dots, v_L\}$ ,

$$\frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_{v \rightarrow v'}}(\tau) = 0.$$

Case 2. We have  $p = (v_l, \dots, v_{L-1}) \in \mathcal{P}_l$ , for  $l \in \llbracket 1, L-1 \rrbracket$ . If  $\{v, v'\} \subset \{v_l, \dots, v_L\}$ , there exists  $l' \in \llbracket l, L-1 \rrbracket$  such that  $(v, v') = (v_{l'}, v_{l'+1})$ , in which case, since  $(v, v') \in F_\theta$ ,  $\rho_\theta(\tau)_{v_{l'} \rightarrow v_{l'+1}} = \tau_{v_{l'} \rightarrow v_{l'+1}}$  and

$$\frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_{v \rightarrow v'}}(\tau) = \tau_{v_l} \prod_{\substack{k \in \llbracket l, L-1 \rrbracket \\ k \neq l'}} \rho_\theta(\tau)_{v_k \rightarrow v_{k+1}}. \quad (39)$$

Otherwise if  $\{v, v'\} \not\subset \{v_l, \dots, v_L\}$ ,

$$\frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_{v \rightarrow v'}}(\tau) = 0.$$

Case 3. We have  $p = \beta$ . In that case, we have

$$\frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_{v \rightarrow v'}}(\tau) = 0.$$

Now let  $v \in B$ , and let us compute  $\frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_v}(\tau)$ .

Case 1. We have  $p = (v_0, \dots, v_{L-1}) \in \mathcal{P}_0$  and

$$\frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_v}(\tau) = 0.$$

Case 2. We have  $p = (v_l, \dots, v_{L-1}) \in \mathcal{P}_l$  for  $l \in \llbracket 1, L-1 \rrbracket$ . If  $v = v_l$ , then

$$\frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_v}(\tau) = \prod_{k \in \llbracket l, L-1 \rrbracket} \rho_\theta(\tau)_{v_k \rightarrow v_{k+1}}.$$

If  $v \neq v_l$ ,

$$\frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_v}(\tau) = 0.$$

Case 3. We have  $p = \beta$  and

$$\frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_v}(\tau) = \begin{cases} 1 & \text{if } v = v_L \\ 0 & \text{if } v \neq v_L. \end{cases}$$

Now that we know the partial derivatives, let us show  $D\psi^\theta(\tau)$  is injective for all  $\tau \in U_\theta$ . Let  $\tau \in U_\theta$  and let  $h \in \mathbb{R}^{F_\theta} \times \mathbb{R}^B$  such that

$$D\psi^\theta(\tau) \cdot h = 0.$$

We need to prove that  $h = 0$ .

Let us show first that for all  $v \in B$ ,  $h_v = 0$ . Let  $l \in \llbracket 1, L-1 \rrbracket$ , and let  $v_l \in V_l$  so that  $v_l$  is arbitrary in  $B \setminus V_L$ . Let us define  $v_{l+1} = s_{\max}^\theta(v_l)$ , then  $v_{l+2} = s_{\max}^\theta(v_{l+1})$ , and so on up to  $v_L = s_{\max}^\theta(v_{L-1})$ . Let us denote  $p = (v_l, \dots, v_{L-1})$ . We have

$$\psi_{p, v_L}^\theta(\tau) = \tau_{v_l} w_{v_l \rightarrow v_{l+1}} \cdots w_{v_{L-1} \rightarrow v_L},$$

so

$$[D\psi^\theta(\tau) \cdot h]_{p, v_L} = \frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_{v_l}}(\tau) h_{v_l} = w_{v_l \rightarrow v_{l+1}} \cdots w_{v_{L-1} \rightarrow v_L} h_{v_l}.$$

Since  $[D\psi^\theta(\tau) \cdot h]_{p, v_L} = 0$  and  $w_{v_l \rightarrow v_{l+1}} \cdots w_{v_{L-1} \rightarrow v_L} \neq 0$ , we conclude that  $h_{v_l} = 0$ . Now let  $v_L \in V_L$ . We consider  $p = \beta$  and we have

$$[D\psi^\theta(\tau) \cdot h]_{p, v_L} = h_{v_L}.$$

Since  $[D\psi^\theta(\tau) \cdot h]_{p, v_L} = 0$ , we also conclude in that case that  $h_{v_L} = 0$ .

Let us now show that for all  $(v, v') \in F_\theta$ ,  $h_{v \rightarrow v'} = 0$ . Let  $l \in \llbracket 1, L \rrbracket$  and let  $(v_{l-1}, v_l) \in (V_{l-1} \times V_l) \cap F_\theta$  so that  $(v_{l-1}, v_l)$  is arbitrary in  $F_\theta$ . If  $l = 1$ , we define  $p_i = (v_{l-1})$  and we have by convention  $\theta_{p_i} = 1 \neq 0$ . If  $l > 1$ , using Proposition 14 there exist  $l' \in \llbracket 0, l-1 \rrbracket$  and a path  $p_i = (v_{l'}, \dots, v_{l-1})$  such that  $\rho_\theta(\tau)_{p_i} \neq 0$ . If  $l < L$ , we define  $v_{l+1} = s_{\max}^\theta(v_l)$ , then  $v_{l+2} = s_{\max}^\theta(v_{l+1})$ , and so on up to  $v_L = s_{\max}^\theta(v_{L-1})$ , and we denote  $p = p_i \cup (v_{l-1}, v_l, \dots, v_{L-1})$ . If  $l = L$ , we denote  $p = p_i$ . Let us show the following expression.

$$[D\psi^\theta(\tau) \cdot h]_{p, v_L} = \sum_{\substack{k \in \llbracket l', l-1 \rrbracket \\ (v_k, v_{k+1}) \in F_\theta}} \frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_{v_k \rightarrow v_{k+1}}}(\tau) h_{v_k \rightarrow v_{k+1}} \quad (40)$$

Indeed, if  $l' \geq 1$ , we have

$$\psi_{p,v_L}^\theta(\tau) = \tau_{v_{l'}} \prod_{k=l'}^{l-1} \rho_\theta(\tau)_{v_k \rightarrow v_{k+1}} \prod_{k=l}^{L-1} w_{v_k \rightarrow v_{k+1}},$$

with the classical convention that if  $l = L$ , the product on the right is empty thus equal to 1. We thus have

$$\begin{aligned} [D\psi^\theta(\tau) \cdot h]_{p,v_L} &= \frac{\partial \psi_{p,v_L}^\theta}{\partial \tau_{v_{l'}}}(\tau) h_{v_{l'}} + \sum_{\substack{k \in \llbracket l', l-1 \rrbracket \\ (v_k, v_{k+1}) \in F_\theta}} \frac{\partial \psi_{p,v_L}^\theta}{\partial \tau_{v_k \rightarrow v_{k+1}}}(\tau) h_{v_k \rightarrow v_{k+1}} \\ &= \sum_{\substack{k \in \llbracket l', l-1 \rrbracket \\ (v_k, v_{k+1}) \in F_\theta}} \frac{\partial \psi_{p,v_L}^\theta}{\partial \tau_{v_k \rightarrow v_{k+1}}}(\tau) h_{v_k \rightarrow v_{k+1}}, \end{aligned}$$

since we have already shown that  $h_{v_{l'}} = 0$ .

If  $l' = 0$ , we have

$$\psi_{p,v_L}^\theta(\tau) = \prod_{k=0}^{l-1} \rho_\theta(\tau)_{v_k \rightarrow v_{k+1}} \prod_{k=l}^{L-1} w_{v_k \rightarrow v_{k+1}},$$

with the same convention that when  $l = L$  the product on the right is equal to 1, so again

$$[D\psi^\theta(\tau) \cdot h]_{p,v_L} = \sum_{\substack{k \in \llbracket 0, l-1 \rrbracket \\ (v_k, v_{k+1}) \in F_\theta}} \frac{\partial \psi_{p,v_L}^\theta}{\partial \tau_{v_k \rightarrow v_{k+1}}}(\tau) h_{v_k \rightarrow v_{k+1}}.$$

This concludes the proof of (40).

We can now show by induction the following statement, for  $l \in \llbracket 0, L \rrbracket$ .

$$\forall l' \in \llbracket 1, l \rrbracket, \forall (v, v') \in (V_{l'-1} \times V_{l'}) \cap F_\theta, h_{v \rightarrow v'} = 0. \quad (H_l)$$

Since  $\llbracket 1, 0 \rrbracket = \emptyset$ ,  $H_0$  is trivially true. Now let  $l \in \llbracket 1, L \rrbracket$  and suppose  $H_{l-1}$  is true. We consider  $(v_{l-1}, v_l) \in (V_{l-1} \times V_l) \cap F_\theta$ , and  $l' \in \llbracket 0, l \rrbracket$ ,  $p_i$  and  $p$  just as before. Since for all  $k \in \llbracket 0, l-2 \rrbracket$ , the induction hypothesis guarantees that  $h_{v_k \rightarrow v_{k+1}} = 0$ , (40) becomes

$$[D\psi^\theta(\tau) \cdot h]_{p,v_L} = \frac{\partial \psi_{p,v_L}^\theta}{\partial \tau_{v_{l-1} \rightarrow v_l}}(\tau) h_{v_{l-1} \rightarrow v_l}.$$

Using (38) and (39), we obtain

$$[D\psi^\theta(\tau) \cdot h]_{p,v_L} = \begin{cases} \rho_\theta(\tau)_{p_i} w_{v_l \rightarrow v_{l+1}} \cdots w_{v_{L-1} \rightarrow v_L} h_{v_{l-1} \rightarrow v_l} & \text{if } l < L \\ \rho_\theta(\tau)_{p_i} h_{v_{l-1} \rightarrow v_l} & \text{if } l = L. \end{cases}$$

Since  $\rho_\theta(\tau)_{p_i} \neq 0$ , and for  $l < L$ ,  $w_{v_l \rightarrow v_{l+1}} \cdots w_{v_{L-1} \rightarrow v_L} \neq 0$ , we conclude that  $h_{v_{l-1} \rightarrow v_l} = 0$  and that  $H_l$  holds.

This induction leads to the conclusion that  $h = 0$  and  $D\psi^\theta(\tau)$  is injective.  $\square$

We are now equipped to prove Theorem 6, which we restate here.

**Theorem 24.**  $\Sigma_*^1$  is a smooth manifold of  $\mathbb{R}^{\mathcal{P} \times V_L}$  of dimension

$$|F_\theta| + |B| = N_0 N_1 + N_1 N_2 + \cdots + N_{L-1} N_L + N_L,$$

and the family  $(V_\theta, (\psi^\theta)^{-1})_{\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S}$  is an atlas.

*Proof.* Our goal is to show that the family  $(V_\theta, (\psi^\theta)^{-1})_{\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S}$  is a smooth atlas, which will show that  $\Sigma_*^1$  is a smooth manifold.

We already know from Proposition 22 that for any  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ ,  $V_\theta$  is an open subset of  $\Sigma_*^1$  and from Proposition 20 that  $(\psi^\theta)^{-1}$  is a homeomorphism from  $V_\theta$  onto  $U_\theta$ . Since for any

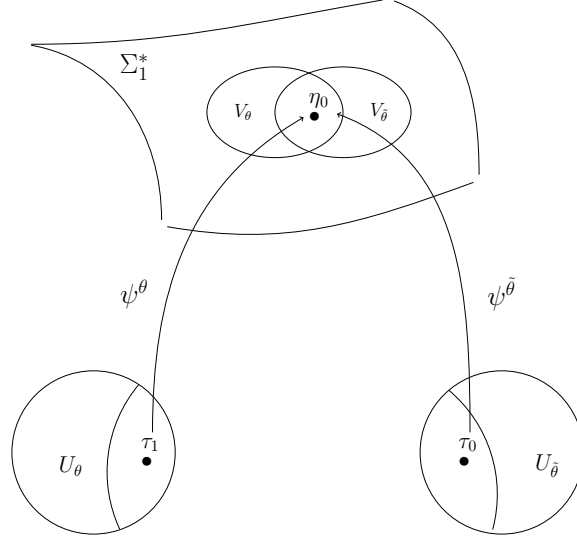


Figure 3: The points  $\eta_0, \tau_0, \tau_1$  and the inverse charts  $\psi^\theta$  and  $\psi^{\tilde{\theta}}$ .

$\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ ,  $\tau_\theta \in U_\theta$ , we have  $\phi(\theta) = \psi^\theta(\tau_\theta) \in V_\theta$  which shows that  $(V_\theta)_{\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S}$  covers  $\Sigma_1^*$ .

Let  $\theta, \tilde{\theta} \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S$ , let us show that the transition map

$$(\psi^\theta)^{-1} \circ \psi^{\tilde{\theta}} : (\psi^{\tilde{\theta}})^{-1}(V_\theta \cap V_{\tilde{\theta}}) \rightarrow (\psi^\theta)^{-1}(V_\theta \cap V_{\tilde{\theta}})$$

is smooth.

Let  $\tau_0 \in U_{\tilde{\theta}}$  such that  $\tau_0 \in (\psi^{\tilde{\theta}})^{-1}(V_\theta \cap V_{\tilde{\theta}})$ . We are going to show that the function  $(\psi^\theta)^{-1} \circ \psi^{\tilde{\theta}}$  is  $C^\infty$  in a neighborhood of  $\tau_0$ .

For ease of reading, let us denote  $\psi^{\tilde{\theta}}(\tau_0)$  by  $\eta_0$ . By definition,  $\eta_0 \in V_\theta \cap V_{\tilde{\theta}}$ . In particular, since  $\eta_0 \in V_\theta$ , we can define  $\tau_1 = (\psi^\theta)^{-1}(\eta_0)$ . See Figure 3 for a representation.

Let  $T = \text{Im } D\psi^\theta(\tau_1)$ , and let us consider a linear subspace  $G$  such that  $T \oplus G = \mathbb{R}^{\mathcal{P} \times V_L}$ . Let  $N_C = |\mathcal{P}|N_L - |F_\theta| - |B| = \dim(G)$ . Let  $i : \mathbb{R}^{N_C} \rightarrow G$  be linear and invertible. Let us consider the function

$$\begin{aligned} \varphi_\theta : U_\theta \times \mathbb{R}^{N_C} &\longrightarrow \mathbb{R}^{\mathcal{P} \times V_L} \\ (\tau, x) &\longmapsto \psi^\theta(\tau) + i(x). \end{aligned}$$

We are going to show that there exist an open neighborhood  $\tilde{U}$  of  $(\tau_1, 0)$  in  $(\mathbb{R}^{F_\theta} \times \mathbb{R}^B) \times \mathbb{R}^{N_C}$  and an open neighborhood  $\tilde{V}$  of  $\eta_0$  in  $\mathbb{R}^{\mathcal{P} \times V_L}$  such that  $\varphi_\theta$  is a  $C^\infty$  diffeomorphism from  $\tilde{U}$  onto  $\tilde{V}$  satisfying

$$\varphi_\theta \left( [(\mathbb{R}^{F_\theta} \times \mathbb{R}^B) \times \{0\}^{N_C}] \cap \tilde{U} \right) = \Sigma_1^* \cap \tilde{V}.$$

Let us first show that  $\varphi_\theta$  is a  $C^\infty$ -diffeomorphism from a neighborhood of  $(\tau_1, 0)$  in  $(\mathbb{R}^{F_\theta} \times \mathbb{R}^B) \times \mathbb{R}^{N_C}$  onto a neighborhood of  $\eta_0$  in  $\mathbb{R}^{\mathcal{P} \times V_L}$ . As shown in Proposition 23,  $\psi^\theta$  is  $C^\infty$  and  $i$  is a linear function, so  $\varphi_\theta$  is  $C^\infty$ . Let us prove that the differential  $D\varphi_\theta(\tau_1, 0)$  is injective. For all  $(\tau, x) \in (\mathbb{R}^{F_\theta} \times \mathbb{R}^B) \times \mathbb{R}^{N_C}$ ,

$$D\varphi_\theta(\tau_1, 0) \cdot (\tau, x) = D\psi^\theta(\tau_1) \cdot \tau + i(x).$$

Since  $D\psi^\theta(\tau_1) \cdot \tau \in T$ ,  $i(x) \in G$ , and  $T$  and  $G$  are in direct sum, if  $D\varphi_\theta(\tau_1, 0) \cdot (\tau, x) = 0$ , then we have

$$\begin{cases} D\psi^\theta(\tau_1) \cdot \tau = 0 \\ i(x) = 0. \end{cases}$$

Since as shown in Proposition 23  $D\psi^\theta(\tau_1)$  is injective, and since  $i$  is invertible, we have

$$(\tau, x) = (0, 0).$$

Hence,  $D\varphi_\theta(\tau_1, 0)$  is injective. Since  $\dim((\mathbb{R}^{F_\theta} \times \mathbb{R}^B) \times \mathbb{R}^{N_C}) = |F_\theta| + |B| + N_C = |\mathcal{P}|N_L$ , the differential  $D\varphi_\theta(\tau_1, 0)$  is bijective. Using the inverse function theorem, there exists an open set  $U \subset U_\theta \times \mathbb{R}^{N_C}$  containing  $(\tau_1, 0)$ , an open set  $V \subset \mathbb{R}^{\mathcal{P} \times V_L}$  containing  $\eta_0$  such that  $\varphi_\theta$  is a  $C^\infty$ -diffeomorphism from  $U$  onto  $V$ .

We have

$$\varphi_\theta \left( [(\mathbb{R}^{F_\theta} \times \mathbb{R}^B) \times \{0\}^{N_C}] \cap U \right) \subset V_\theta \cap V.$$

In fact, if  $V$  is small enough, this inclusion is an equality. We are going to construct open subsets  $\tilde{U} \subset U$  and  $\tilde{V} \subset V$  so that it is the case. Let us define

$$O = \{\tau \in U_\theta, (\tau, 0) \in U\}.$$

Since  $U$  is an open set containing  $(\tau_1, 0)$ ,  $O$  is an open set containing  $\tau_1 = (\psi^\theta)^{-1}(\eta_0)$ . Since, according to Proposition 20,  $\psi^\theta$  is a homeomorphism,  $\psi^\theta(O)$  is an open subset of  $V_\theta$  so there exists  $\epsilon > 0$  such that

$$V_\theta \cap B_\infty(\eta_0, \epsilon) \subset \psi^\theta(O). \quad (41)$$

We can now define  $\tilde{V} = V \cap B_\infty(\eta_0, \epsilon)$ , and  $\tilde{U} = \{(\tau, x) \in U, \varphi_\theta(\tau, x) \in \tilde{V}\}$ , which are open sets such that  $(\tau_1, 0) \in \tilde{U}$ ,  $\eta_0 \in \tilde{V}$ , and  $\varphi_\theta$  is a  $C^\infty$ -diffeomorphism from  $\tilde{U}$  onto  $\tilde{V}$ . Let us show that

$$\varphi_\theta \left( [(\mathbb{R}^{F_\theta} \times \mathbb{R}^B) \times \{0\}^{N_C}] \cap \tilde{U} \right) = V_\theta \cap \tilde{V}. \quad (42)$$

The direct inclusion is immediate: if  $(\tau, 0) \in [(\mathbb{R}^{F_\theta} \times \mathbb{R}^B) \times \{0\}^{N_C}] \cap \tilde{U}$ , then

$$\varphi_\theta(\tau, 0) = \psi^\theta(\tau) \in V_\theta \cap \tilde{V}.$$

For the reciprocal inclusion, if  $\tau \in U_\theta$  is such that  $\psi^\theta(\tau) \in V_\theta \cap \tilde{V}$ , then by definition of  $\epsilon$  and  $\tilde{V}$ , (41) guarantees, since  $\psi^\theta$  is injective, that  $\tau \in O$ . By definition of  $O$ , we have  $(\tau, 0) \in U$ , and since

$$\varphi_\theta(\tau, 0) = \psi^\theta(\tau) \in \tilde{V},$$

this shows  $(\tau, 0) \in \tilde{U}$ . This shows the reciprocal inclusion, and thus (42) holds.

Let us now define

$$P_\theta : \mathbb{R}^{F_\theta} \times \mathbb{R}^B \times \mathbb{R}^{N_C} \longrightarrow \mathbb{R}^{F_\theta} \times \mathbb{R}^B \\ (\tau, x) \longmapsto \tau$$

the restriction to the first component, and let us observe that over  $V_\theta \cap \tilde{V}$ , we have

$$P_\theta \circ (\varphi_\theta)^{-1} = (\psi^\theta)^{-1}. \quad (43)$$

Indeed, if  $\eta \in V_\theta \cap \tilde{V}$ , then by (42) there exists  $\tau \in U_\theta$  such that  $(\tau, 0) \in \tilde{U}$  and  $\varphi_\theta(\tau, 0) = \eta$ . Since  $\varphi_\theta(\tau, 0) = \psi^\theta(\tau)$ , this shows that  $\tau = (\psi^\theta)^{-1}(\eta)$  and thus

$$(\psi^\theta)^{-1}(\eta) = P_\theta(\tau, 0) = P_\theta \circ (\varphi_\theta)^{-1}(\eta).$$

Now recall that  $\eta_0 = \psi^{\tilde{\theta}}(\tau_0)$ . By continuity of  $\psi^{\tilde{\theta}}$ , there exists  $\epsilon' > 0$  such that  $B_\infty(\tau_0, \epsilon') \subset (\psi^{\tilde{\theta}})^{-1}(V_\theta \cap V_{\tilde{\theta}})$  and

$$\psi^{\tilde{\theta}}(B_\infty(\tau_0, \epsilon')) \subset \tilde{V}.$$

For any  $\tau \in B_\infty(\tau_0, \epsilon')$ , we have  $\psi^{\tilde{\theta}}(\tau) \in V_\theta \cap \tilde{V}$  so, as we just proved with (43),  $(\psi^\theta)^{-1} \circ \psi^{\tilde{\theta}}(\tau) = P_\theta \circ (\varphi_\theta)^{-1} \circ \psi^{\tilde{\theta}}(\tau)$ . Since the functions  $\psi^{\tilde{\theta}}$ ,  $(\varphi_\theta)^{-1}$  and  $P_\theta$  are all  $C^\infty$ , we conclude that the transition map  $(\psi^\theta)^{-1} \circ \psi^{\tilde{\theta}}$  is  $C^\infty$  over  $B_\infty(\tau_0, \epsilon')$ , for all  $\tau_0 \in (\psi^{\tilde{\theta}})^{-1}(V_\theta \cap V_{\tilde{\theta}})$ . We conclude that  $(\psi^\theta)^{-1} \circ \psi^{\tilde{\theta}}$  is  $C^\infty$  over  $(\psi^{\tilde{\theta}})^{-1}(V_\theta \cap V_{\tilde{\theta}})$ .

We have showed that  $(V_\theta, (\psi^\theta)^{-1})_{\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus S}$  is a smooth atlas, and thus that  $\Sigma_1^*$  is a smooth submanifold of  $\mathbb{R}^{\mathcal{P} \times V_L}$ . As computed in (21), its dimension is

$$|F_\theta| + |B| = N_0 N_1 + N_1 N_2 + \cdots + N_{L-1} N_L + N_L.$$

□

## D Conditions of local identifiability

Let us restate and prove Theorem 7.

**Theorem 25.** For any  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus (S \cup \Delta_X)$ , the two following statements are equivalent:

i) There exists  $\epsilon > 0$  such that for any  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ , if  $\|\theta - \tilde{\theta}\| < \epsilon$ , then

$$f_\theta(X) = f_{\tilde{\theta}}(X) \implies \theta \sim \tilde{\theta}.$$

ii) There exists  $\epsilon' > 0$  such that

$$B_\infty(\phi(\theta), \epsilon') \cap \Sigma_1^* \cap N(X, \theta) = \{\phi(\theta)\}.$$

*Proof.*

$i) \Rightarrow ii)$  Suppose  $i)$  is satisfied for some  $\epsilon_1 > 0$ . We first construct  $\epsilon' > 0$  and then consider  $\eta \in B_\infty(\phi(\theta), \epsilon') \cap \Sigma_1^* \cap N(X, \theta)$ , and we prove that  $\eta = \phi(\theta)$ . Since  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus (S \cup \Delta_X)$  and since, according to Proposition 13,  $\Delta_X$  is closed, there exists  $\epsilon_2 > 0$  such that for any  $\tilde{\theta} \in B_\infty(\theta, \epsilon_2)$ ,

$$\alpha(X, \theta) = \alpha(X, \tilde{\theta}),$$

i.e.

$$A(X, \theta) = A(X, \tilde{\theta}).$$

Consider  $\epsilon = \min(\epsilon_1, \epsilon_2)$ . Since, according to Proposition 20,  $\rho_\theta \circ (\psi^\theta)^{-1}$  is continuous at  $\phi(\theta) \in \psi^\theta(U_\theta)$ , and since  $\rho_\theta \circ (\psi^\theta)^{-1}(\phi(\theta)) = \rho_\theta(\tau_\theta) = \theta$ , there exists  $\epsilon' > 0$  such that for all  $\tau \in U_\theta$ ,

$$\|\psi^\theta(\tau) - \phi(\theta)\|_\infty < \epsilon' \implies \|\rho_\theta(\tau) - \theta\|_\infty = \|\rho_\theta \circ (\psi^\theta)^{-1}(\psi^\theta(\tau)) - \rho_\theta \circ (\psi^\theta)^{-1}(\phi(\theta))\|_\infty < \epsilon. \quad (44)$$

Since  $\phi(\theta) = \psi^\theta(\tau_\theta)$ , Proposition 22 guarantees that, modulo a decrease of  $\epsilon'$ , we can assume that

$$B_\infty(\phi(\theta), \epsilon') \cap \Sigma_1^* \subset \psi^\theta(U_\theta). \quad (45)$$

Now let  $\eta \in B_\infty(\phi(\theta), \epsilon') \cap \Sigma_1^* \cap N(X, \theta)$ . Let us prove that  $\eta = \phi(\theta)$ . Using (45), there exists  $\tau \in U_\theta$  such that  $\eta = \psi^\theta(\tau)$ . Since  $\|\phi(\theta) - \eta\|_\infty < \epsilon'$ , we have using (44)

$$\|\rho_\theta(\tau) - \theta\|_\infty < \epsilon. \quad (46)$$

Since  $\epsilon < \epsilon_2$ , we have

$$A(X, \theta) = A(X, \rho_\theta(\tau)). \quad (47)$$

Since  $\psi^\theta(\tau) = \eta \in N(X, \theta)$ , we have by definition of  $N(X, \theta)$  that  $\psi^\theta(\tau) - \phi(\theta) \in \text{Ker } A(X, \theta)$ , so

$$A(X, \theta) \cdot \psi^\theta(\tau) = A(X, \theta) \cdot \phi(\theta) \quad (48)$$

Using successively (3), (47), (48) and (3) again, we have

$$\begin{aligned} f_{\rho_\theta(\tau)}(X) &= A(X, \rho_\theta(\tau)) \cdot \phi(\rho_\theta(\tau)) \\ &= A(X, \theta) \cdot \phi(\rho_\theta(\tau)) \\ &= A(X, \theta) \cdot \phi(\theta) \\ &= f_\theta(X). \end{aligned}$$

Since the hypothesis  $i)$  holds for  $\epsilon_1$ , using (46) and the fact that  $\epsilon < \epsilon_1$ , we have

$$\theta \sim \rho_\theta(\tau).$$

We conclude using Proposition 15 that

$$\eta = \phi(\rho_\theta(\tau)) = \phi(\theta),$$

which shows

$$B_\infty(\phi(\theta), \epsilon') \cap \Sigma_1^* \cap N(X, \theta) \subset \{\phi(\theta)\}.$$

The converse inclusion trivially holds and therefore  $ii)$  holds.



ii)  $\Rightarrow$  i) Suppose ii) is satisfied for some  $\epsilon' > 0$ .

We first construct  $\epsilon$  and prove i) holds. Since  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus (S \cup \Delta_X)$ , using Proposition 13, there exists  $\epsilon_1 > 0$  such that for all  $\tilde{\theta} \in B_\infty(\theta, \epsilon_1)$ ,

$$\alpha(X, \theta) = \alpha(X, \tilde{\theta}),$$

i.e.

$$A(X, \theta) = A(X, \tilde{\theta}). \quad (49)$$

Since  $\phi$  is continuous, there exists  $\epsilon_2 > 0$  such that

$$\|\theta - \tilde{\theta}\|_\infty < \epsilon_2 \implies \|\phi(\theta) - \phi(\tilde{\theta})\|_\infty < \epsilon'.$$

Using Proposition 18, there exists  $\epsilon_3 > 0$  such that

$$\theta \stackrel{R}{\sim} \tilde{\theta} \text{ and } \|\theta - \tilde{\theta}\|_\infty < \epsilon_3 \implies \theta \sim \tilde{\theta}.$$

Since  $\theta \notin S$  and  $S$  is closed, there exists  $\epsilon_4 > 0$  such that for all  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ , if  $\|\theta - \tilde{\theta}\|_\infty < \epsilon_4$ , then

$$\tilde{\theta} \notin S.$$

Let  $\epsilon = \min(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)$ . Let  $\tilde{\theta} \in B_\infty(\theta, \epsilon)$ , and suppose

$$f_\theta(X) = f_{\tilde{\theta}}(X).$$

Let us prove that  $\theta \sim \tilde{\theta}$ . Reformulating the above equality using (3) for both sides, and using the definition of  $A$  given in the beginning of Section 4, we have

$$A(X, \theta) \cdot \phi(\theta) = A(X, \tilde{\theta}) \cdot \phi(\tilde{\theta}).$$

Since  $\|\theta - \tilde{\theta}\|_\infty < \epsilon \leq \epsilon_1$ , we have the equality (49) and thus

$$A(X, \theta) \cdot \phi(\theta) = A(X, \theta) \cdot \phi(\tilde{\theta}).$$

In other words,  $\phi(\tilde{\theta}) - \phi(\theta) \in \text{Ker } A(X, \theta)$ . Since  $\epsilon < \epsilon_4$ ,  $\phi(\tilde{\theta}) \in \Sigma_1^*$ . Since  $\epsilon < \epsilon_2$ ,  $\phi(\tilde{\theta}) \in B_\infty(\phi(\theta), \epsilon')$ . Summarizing,

$$\phi(\tilde{\theta}) \in B_\infty(\phi(\theta), \epsilon') \cap \Sigma_1^* \cap N(X, \theta),$$

and using the hypothesis ii), we conclude that

$$\phi(\tilde{\theta}) = \phi(\theta).$$

By Proposition 17, we have  $\theta \stackrel{R}{\sim} \tilde{\theta}$ , and since  $\epsilon < \epsilon_3$ , we conclude that

$$\theta \sim \tilde{\theta}.$$

□

We are now going to prove Theorems 8 and 9, which we restate as Theorems 26 and 27 respectively.

**Theorem 26** (Necessary condition). *Let  $X \in \mathbb{R}^{\llbracket 1, n \rrbracket \times V_L}$  and  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus (S \cup \Delta_X)$ . If  $C_N$  is not satisfied, then for all  $\epsilon > 0$  there exists  $\tilde{\theta} \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus (S \cup \Delta_X)$  such that  $\theta \not\sim \tilde{\theta}$ ,  $\|\theta - \tilde{\theta}\|_\infty < \epsilon$  and*

$$f_\theta(X) = f_{\tilde{\theta}}(X).$$

**Theorem 27** (Sufficient condition). *Let  $X \in \mathbb{R}^{\llbracket 1, n \rrbracket \times V_L}$  and  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus (S \cup \Delta_X)$ . If  $C_S$  is satisfied, then there exists  $\epsilon > 0$  such that for all  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ , if  $\|\theta - \tilde{\theta}\|_\infty < \epsilon$ ,*

$$f_\theta(X) = f_{\tilde{\theta}}(X) \implies \theta \sim \tilde{\theta}.$$

To prove the theorems, we need to prove first the following lemmas.

**Lemma 28.** Let us denote by  $T = \text{Im } D\psi^\theta(\tau_\theta)$  the direction of the tangent plane to  $\Sigma_1^*$  at  $\phi(\theta)$ . Let us denote by  $H$  the intersection  $\text{Ker } A(X, \theta) \cap T$ . We have

$$\dim(H) = |F_\theta| + |B| - R_\Gamma. \quad (50)$$

*Proof.* Let  $\eta \in T$ . There exists  $h \in \mathbb{R}^{F_\theta} \times \mathbb{R}^B$  such that  $\eta = D\psi^\theta(\tau_\theta) \cdot h$ . We have the following equivalence:

$$\begin{aligned} \eta \in \text{Ker } A(X, \theta) &\iff A(X, \theta) \cdot \eta = 0 \\ &\iff A(X, \theta) \circ D\psi^\theta(\tau_\theta) \cdot h = 0 \\ &\iff \Gamma(X, \theta) \cdot h = 0 \\ &\iff h \in \text{Ker } \Gamma(X, \theta). \end{aligned}$$

This shows that  $D\psi^\theta(\tau_\theta)^{-1}(\text{Ker } A(X, \theta) \cap T) = \text{Ker } \Gamma(X, \theta) \subset \mathbb{R}^{F_\theta} \times \mathbb{R}^B$ .

Since  $D\psi^\theta(\tau_\theta)$  is injective, we thus have

$$\dim(H) = \dim(\text{Ker } \Gamma(X, \theta)) = |F_\theta| + |B| - R_\Gamma. \quad \square$$

**Lemma 29.** Let  $G$  be a supplementary subspace of  $\text{Ker } A(X, \theta)$  such that

$$H \oplus G = \text{Ker } A(X, \theta). \quad (51)$$

If  $R_\Gamma = R_A$ , there exist an open set  $\mathcal{O} \subset U_\theta \times G$  containing  $(\tau_\theta, 0)$  and an open set  $\mathcal{V} \subset \mathbb{R}^{\mathcal{P} \times V_L}$  containing  $\phi(\theta)$  such that

$$\xi : \begin{array}{ccc} \mathcal{O} & \longrightarrow & \mathcal{V} \\ (\tau, g) & \longmapsto & \psi^\theta(\tau) + g \end{array}$$

is a diffeomorphism from  $\mathcal{O}$  onto  $\mathcal{V}$ .

*Proof.* Let us first show that

$$T \oplus G = \mathbb{R}^{\mathcal{P} \times V_L}. \quad (52)$$

Indeed, since  $\text{Ker } A(X, \theta) = H \oplus G$  and  $T \cap \text{Ker } A(X, \theta) = H$ , we have  $T \cap G = \{0\}$ . We of course have

$$T \oplus G \subset \mathbb{R}^{\mathcal{P} \times V_L}. \quad (53)$$

Let us show that  $\dim(G) = \dim(\mathbb{R}^{\mathcal{P} \times V_L}) - \dim(T)$ . First note that we have

$$\dim(\text{Ker } A(X, \theta)) = \dim(\mathbb{R}^{\mathcal{P} \times V_L}) - \text{rank}(A(X, \theta)) = |\mathcal{P}|N_L - R_A. \quad (54)$$

Using (51) and (54), we have

$$\begin{aligned} \dim(G) &= \dim(\text{Ker } A(X, \theta)) - \dim(H) \\ &= |\mathcal{P}|N_L - R_A - \dim(H). \end{aligned}$$

Using (50) and the hypothesis  $R_\Gamma = R_A$  we thus have

$$\begin{aligned} \dim(G) &= |\mathcal{P}|N_L - R_A + R_\Gamma - |F_\theta| - |B| \\ &= |\mathcal{P}|N_L - |F_\theta| - |B| \\ &= |\mathcal{P}|N_L - \dim(T), \end{aligned}$$

where the last equality comes from the injectivity of  $D\psi^\theta(\tau_\theta)$ , shown in Proposition 23. Together with (53), this proves (52).

Let us now consider the function

$$\xi : \begin{array}{ccc} U_\theta \times G & \longrightarrow & \mathbb{R}^{\mathcal{P} \times V_L} \\ (\tau, g) & \longmapsto & \psi^\theta(\tau) + g. \end{array}$$

For all  $(h, g) \in (\mathbb{R}^{F_\theta} \times \mathbb{R}^B) \times G$ , we have

$$D\xi(\tau_\theta, 0) \cdot (h, g) = D\psi^\theta(\tau_\theta)h + g.$$

The differential  $D\xi(\tau_\theta, 0)$  is injective. Indeed, if

$$D\xi(\tau_\theta, 0) \cdot (h, g) = 0,$$

then since  $D\psi^\theta(\tau_\theta)h \in T$  and  $g \in G$ , we have

$$\begin{cases} D\psi^\theta(\tau_\theta)h = 0 \\ g = 0, \end{cases}$$

and since  $D\psi^\theta(\tau_\theta)$  is injective,  $h = 0$  and  $D\xi(\tau_\theta, 0)$  is injective. Since, using (52),

$$\dim(\mathbb{R}^{F_\theta} \times \mathbb{R}^B) + \dim(G) = |\mathcal{P}|N_L,$$

$D\xi(\tau_\theta, 0)$  is bijective.

We can thus apply the inverse function theorem: there exists an open set  $\mathcal{O} \subset U_\theta \times G$  containing  $(\tau_\theta, 0)$ , an open set  $\mathcal{V} \subset \mathbb{R}^{\mathcal{P} \times V_L}$  containing  $\phi(\theta)$  such that  $\xi$  is a diffeomorphism from  $\mathcal{O}$  into  $\mathcal{V}$ .  $\square$

We can now prove the theorems.

*Proof of Theorem 26.* If  $C_N$  is not satisfied, then we have  $R_\Gamma = R_A < |F_\theta| + |B|$ . We can thus apply Lemma 29: there exist an open set  $\mathcal{O} \subset U_\theta \times G$  containing  $(\tau_\theta, 0)$  and an open set  $\mathcal{V} \subset \mathbb{R}^{\mathcal{P} \times V_L}$  containing  $\phi(\theta)$  such that

$$\begin{aligned} \xi : \quad \mathcal{O} &\longrightarrow \mathcal{V} \\ (\tau, g) &\longmapsto \psi^\theta(\tau) + g \end{aligned}$$

is a diffeomorphism from  $\mathcal{O}$  onto  $\mathcal{V}$ .

Consider  $\epsilon > 0$ . We define the open set  $\tilde{\mathcal{O}} = \mathcal{O} \cap (\psi^\theta)^{-1}(B(\phi(\theta), \epsilon) \times G)$  and its image  $\tilde{\mathcal{V}} = \xi(\tilde{\mathcal{O}})$ .

Using the computation of  $\dim(H)$  shown in Lemma 28, we have

$$\dim(H) = |F_\theta| + |B| - R_\Gamma > 0,$$

so there exists a nonzero  $w \in H$  such that  $\phi(\theta) + w \in \tilde{\mathcal{V}}$ . Since  $\xi$  induces a diffeomorphism from  $\tilde{\mathcal{O}}$  onto  $\tilde{\mathcal{V}}$ , there exists  $(\tau, g) \in \tilde{\mathcal{O}}$  such that

$$\phi(\theta) + w = \psi^\theta(\tau) + g$$

i.e.

$$\psi^\theta(\tau) - \phi(\theta) = w - g. \tag{55}$$

Let us denote  $\tilde{\theta} = \rho_\theta(\tau)$  and let us show that Theorem 25.ii) does not hold. By definition,  $\phi(\tilde{\theta}) = \psi^\theta(\tau)$  and since  $(\tau, g) \in \tilde{\mathcal{O}}$ ,  $\|\phi(\theta) - \phi(\tilde{\theta})\|_\infty < \epsilon$ . Since  $H \cap G = \{0\}$ ,  $w \in H$ ,  $g \in G$  and  $w \neq 0$ , (55) shows that

$$\phi(\tilde{\theta}) - \phi(\theta) \neq 0.$$

Furthermore, since  $w \in H \subset \text{Ker } A(X, \theta)$  and  $g \in G \subset \text{Ker } A(X, \theta)$ , (55) shows that

$$\phi(\tilde{\theta}) - \phi(\theta) \in \text{Ker } A(X, \theta),$$

so

$$\phi(\tilde{\theta}) \in N(X, \theta).$$

Summarizing, for any  $\epsilon > 0$  there exists  $\tilde{\theta} \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus \mathcal{S}$  such that  $\phi(\tilde{\theta}) \in B_\infty(\phi(\theta), \epsilon) \cap \Sigma_1^* \cap N(X, \theta) \setminus \{\phi(\theta)\}$ . The second item of Theorem 25 does not hold. Since it is equivalent, the first item of Theorem 25 does not hold either. In other words, the conclusion of Theorem 26 holds.  $\square$

*Proof of Theorem 27.* Suppose that  $C_S$  is satisfied. Using Lemma 28 and using  $C_S$ , we obtain

$$\dim(T \cap \text{Ker } A(X, \theta)) = |F_\theta| + |B| - R_\Gamma = 0.$$

We thus have

$$T \cap \text{Ker } A(X, \theta) = \{0\}. \quad (56)$$

In order to apply Theorem 25, let us show by contradiction that there exists  $\epsilon > 0$  such that

$$B_\infty(\phi(\theta), \epsilon) \cap \Sigma_1^* \cap N(X, \theta) = \{\phi(\theta)\}. \quad (57)$$

More precisely, we suppose that for all  $n \in \mathbb{N}^*$ , there exists  $\phi_n \in N(X, \theta) \cap \Sigma_1^*$  such that  $\phi_n \neq \phi(\theta)$  and  $\|\phi(\theta) - \phi_n\|_\infty < \frac{1}{n}$  and prove that it leads to  $T \cap \text{Ker } A(X, \theta) \neq \{0\}$ , which contradicts (56).

Using Proposition 22, there exists  $n_0 \in \mathbb{N}^*$  such that for all  $n \geq n_0$ , there exists  $\tau_n \in U_\theta$  such that  $\phi_n = \psi^\theta(\tau_n)$ . Since  $\psi^\theta$  is a homeomorphism and  $\psi^\theta(\tau_\theta) = \phi(\theta)$ ,

$$\phi_n \rightarrow \phi(\theta)$$

implies that

$$\tau_n \rightarrow \tau_\theta.$$

Moreover, for all  $n \geq n_0$ ,  $\tau_n \neq \tau_\theta$ .

When  $n$  tends to infinity, we can thus write

$$\phi_n - \phi(\theta) = \psi^\theta(\tau_n) - \psi^\theta(\tau_\theta) = D\psi^\theta(\tau_\theta) \cdot (\tau_n - \tau_\theta) + o(\tau_n - \tau_\theta).$$

Let us apply  $A(X, \theta)$  and divide by  $\|\tau_n - \tau_\theta\|$ .

$$\frac{1}{\|\tau_n - \tau_\theta\|} A(X, \theta) \cdot (\phi_n - \phi(\theta)) = A(X, \theta) \circ D\psi^\theta(\tau_\theta) \cdot \left( \frac{\tau_n - \tau_\theta}{\|\tau_n - \tau_\theta\|} \right) + \frac{1}{\|\tau_n - \tau_\theta\|} A(X, \theta) \circ o(\tau_n - \tau_\theta). \quad (58)$$

Since  $\phi_n \in N(X, \theta)$  for all  $n \in \mathbb{N}^*$ ,

$$\frac{1}{\|\tau_n - \tau_\theta\|} A(X, \theta) \cdot (\phi_n - \phi(\theta)) = 0.$$

Since  $\frac{\tau_n - \tau_\theta}{\|\tau_n - \tau_\theta\|}$  belongs to the unit sphere, we can extract a subsequence that converges to a limit  $h$  with norm equal to 1. Taking the limit in (58) according to this subsequence, we obtain

$$0 = A(X, \theta) \circ D\psi^\theta(\tau_\theta) \cdot h,$$

which shows that  $D\psi^\theta(\tau_\theta) \cdot h \in \text{Ker } A(X, \theta)$ . Since  $h \neq 0$  and  $D\psi^\theta(\tau_\theta)$  is injective,  $D\psi^\theta(\tau_\theta)h \neq 0$  and

$$T \cap \text{Ker } A(X, \theta) \neq \{0\}.$$

This is in contradiction with (56).

We have proven (57). We can now conclude thanks to Lemma 25: there exists  $\epsilon' > 0$  such that for any  $\tilde{\theta} \in \mathbb{R}^E \times \mathbb{R}^B$ , if  $\|\theta - \tilde{\theta}\| < \epsilon'$ , then

$$f_\theta(X) = f_{\tilde{\theta}}(X) \implies \theta \sim \tilde{\theta}.$$

□

## E Checking the conditions numerically

We restate and prove Proposition 11.

**Proposition 30.** *Let  $X \in \mathbb{R}^{n \times V_0}$  and  $\theta \in \mathbb{R}^E \times \mathbb{R}^B$ . We have*

$$R_A = N_L \text{rank}(\alpha(X, \theta)).$$

*Proof.* Let  $\eta \in \mathbb{R}^{\mathcal{P} \times V_L}$ . We have

$$A(X, \theta) \cdot \eta = \alpha(X, \theta)\eta.$$

If we denote by  $\eta^1, \dots, \eta^{N_L} \in \mathbb{R}^{\mathcal{P}}$  the  $N_L$  columns of  $\eta$ , the columns of  $A(X, \theta) \cdot \eta$  are  $\alpha(X, \theta)\eta^1, \dots, \alpha(X, \theta)\eta^{N_L}$ . If we consider the matrix  $\eta$  as a family of  $N_L$  vectors of  $\mathbb{R}^{\mathcal{P}}$  and the matrix  $A(X, \theta) \cdot \eta$  as a family of  $N_L$  vectors of  $\mathbb{R}^n$ , the operator  $A(X, \theta)$  can then be equivalently described as

$$A(X, \theta) : \begin{array}{ccc} (\mathbb{R}^{\mathcal{P}})^{N_L} & \longrightarrow & (\mathbb{R}^n)^{N_L} \\ (\eta^1, \dots, \eta^{N_L}) & \longmapsto & (\alpha(X, \theta)\eta^1, \dots, \alpha(X, \theta)\eta^{N_L}). \end{array}$$

The rank of such an operator is  $N_L \text{rank}(\alpha(X, \theta))$ .  $\square$

We restate and prove Proposition 10.

**Proposition 31.** *Let  $X \in \mathbb{R}^{n \times V_0}$  and  $\theta \in (\mathbb{R}^E \times \mathbb{R}^B) \setminus (S \cup \Delta_X)$ . The function*

$$\begin{array}{ccc} U_\theta & \longrightarrow & \mathbb{R}^{n \times V_L} \\ \tau & \longmapsto & f_{\rho_\theta(\tau)}(X) \end{array}$$

*is differentiable in a neighborhood of  $\tau_\theta$ , and we denote by  $D_\tau f_\theta(X)$  its differential at  $\tau_\theta$ . We have*

$$D_\tau f_\theta(X) = \Gamma(X, \theta). \quad (59)$$

*Proof.* Using (3) at  $\rho_\theta(\tau)$  and the definition of  $\psi^\theta$  in (11), we have

$$f_{\rho_\theta(\tau)}(X) = A(X, \theta) \cdot \psi^\theta(\tau).$$

Taking the differential of

$$\begin{array}{ccc} U_\theta & \longrightarrow & \mathbb{R}^{n \times V_L} \\ \tau & \longmapsto & f_{\rho_\theta(\tau)}(X) \end{array}$$

at  $\tau_\theta$ , and using (13), we obtain

$$D_\tau f_\theta(X) = A(X, \theta) \circ D\psi^\theta(\tau_\theta) = \Gamma(X, \theta). \quad \square$$

To finish with, the following proposition gives explicit expressions of the coefficients of  $\Gamma(X, \theta)$ . These expressions are given for the sake of theoretical completeness. Note that when it comes to computing  $\Gamma(X, \theta)$  in practice (in order to compute  $R_\Gamma$ ), the correct approach is using backpropagation as described in Section 5 rather than evaluating the expressions in Proposition 32 which involve sums with very large numbers of summands.

**Proposition 32.** *If we decompose it in the canonical bases of  $\mathbb{R}^{F_\theta} \times \mathbb{R}^B$  and  $\mathbb{R}^{\llbracket 1, n \rrbracket \times V_L}$ ,  $\Gamma(X, \theta)$  is a  $(nN_L) \times (|F_\theta| + |B|)$  matrix. For lighter notations, let us drop the dependency in  $(X, \theta)$  and denote by  $\gamma^{i, v_L}$  the lines of  $\Gamma(X, \theta)$ , for  $i \in \llbracket 1, n \rrbracket$  and  $v_L \in V_L$ , which satisfy  $(\gamma^{i, v_L})^T \in \mathbb{R}^{F_\theta} \times \mathbb{R}^B$ . For any  $(i, v_L) \in \llbracket 1, n \rrbracket \times V_L$ , let us express the coefficients of  $\gamma^{i, v_L}$ , i.e. express  $\gamma_{v_l \rightarrow v_{l+1}}^{i, v_L}$  for any  $v_l \rightarrow v_{l+1} \in F_\theta$  and express  $\gamma_{v_l}^{i, v_L}$  for any  $v_l \in B$ .*

- For any  $l \in \llbracket 0, L-1 \rrbracket$  and any  $(v_l, v_{l+1}) \in V_l \times V_{l+1}$  such that  $v_l \rightarrow v_{l+1} \in F_\theta$ ,

$$\begin{aligned} \gamma_{v_l \rightarrow v_{l+1}}^{i, v_L} &= \sum_{v_0 \in V_0} x_{v_0}^i \bar{w}_{v_0 \rightarrow v_1} \bar{a}_{v_l}(x^i, \theta) \prod_{\substack{1 \leq k \leq L-1 \\ k \neq l}} a_{v_k}(x^i, \theta) w_{v_k \rightarrow v_{k+1}} \\ &\quad \vdots \\ &\quad v_{l-1} \in V_{l-1} \\ &\quad v_{l+2} \in V_{l+2} \\ &\quad \vdots \\ &\quad v_{L-1} \in V_{L-1} \\ &+ \sum_{l'=1}^L \sum_{v_{l'} \in V_{l'}} b_{v_{l'}} \bar{a}_{v_l}(x^i, \theta) \prod_{\substack{l' \leq k \leq L-1 \\ k \neq l}} a_{v_k}(x^i, \theta) w_{v_k \rightarrow v_{k+1}}, \quad (60) \\ &\quad \vdots \\ &\quad v_{l-1} \in V_{l-1} \\ &\quad v_{l+2} \in V_{l+2} \\ &\quad \vdots \\ &\quad v_{L-1} \in V_{L-1} \end{aligned}$$

where  $\bar{w}_{v_0 \rightarrow v_1} = w_{v_0 \rightarrow v_1}$  and  $\bar{a}_{v_l}(x^i, \theta) = a_{v_l}(x^i, \theta)$  except when  $l = 0$  in which case  $\bar{w}_{v_0 \rightarrow v_1} = 1$  and  $\bar{a}_{v_l}(x^i, \theta) = 1$ .

- For any  $l \in \llbracket 1, L \rrbracket$  and any  $v_l \in V_l$ ,

$$\gamma_{v_l \rightarrow v_{l+1}}^{i, v_L} = \sum_{v_{l+1} \in V_{l+1}} \prod_{l \leq k \leq L-1} a_{v_k}(x^i, \theta) w_{v_k \rightarrow v_{k+1}}. \quad (61)$$

$$\vdots$$

$$v_{L-1} \in V_{L-1}$$

*Proof.* Let  $(i, v_L) \in \llbracket 1, n \rrbracket \times V_L$ .

Let us compute  $\gamma_{v_l \rightarrow v_{l+1}}^{i, v_L}$ , for  $l \in \llbracket 0, L-1 \rrbracket$  and  $(v_l, v_{l+1}) \in V_l \times V_{l+1}$  such that  $v_l \rightarrow v_{l+1} \in F_\theta$ .  $\gamma_{v_l \rightarrow v_{l+1}}^{i, v_L}$  is the coefficient corresponding to the line  $(i, v_L)$  and the column  $(v_l \rightarrow v_{l+1})$  of  $\Gamma(X, \theta)$ . Let us denote by  $h^{v_l \rightarrow v_{l+1}} \in \mathbb{R}^{F_\theta} \times \mathbb{R}^B$  the vector whose component indexed by  $v_l \rightarrow v_{l+1}$  is equal to 1 and whose other components are zero. Let us denote by  $d^{i, v_L} \in \mathbb{R}^{n \times V_L}$  the element whose component indexed by  $(i, v_L)$  is equal to 1 and whose other components are zero. Let us denote by  $\langle \cdot, \cdot \rangle_{\mathbb{R}^{n \times V_L}}$  the scalar product of the euclidean space  $\mathbb{R}^{n \times V_L}$ . We have

$$\begin{aligned} \gamma_{v_l \rightarrow v_{l+1}}^{i, v_L} &= \left\langle d^{i, v_L}, \Gamma(X, \theta) \cdot h^{v_l \rightarrow v_{l+1}} \right\rangle_{\mathbb{R}^{n \times V_L}} \\ &= \left\langle d^{i, v_L}, A(X, \theta) \circ D\psi^\theta(\tau_\theta) \cdot h^{v_l \rightarrow v_{l+1}} \right\rangle_{\mathbb{R}^{n \times V_L}} \\ &= \left\langle d^{i, v_L}, A(X, \theta) \cdot \frac{\partial \psi^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta) \right\rangle_{\mathbb{R}^{n \times V_L}} \\ &= \left\langle d^{i, v_L}, \alpha(X, \theta) \frac{\partial \psi^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta) \right\rangle_{\mathbb{R}^{n \times V_L}} \\ &= \left[ \alpha(X, \theta) \frac{\partial \psi^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta) \right]_{i, v_L}, \end{aligned}$$

where  $\left[ \alpha(X, \theta) \frac{\partial \psi^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta) \right]_{i, v_L}$  denotes the coefficient  $(i, v_L)$  of the product  $\alpha(X, \theta) \frac{\partial \psi^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta)$ . Let us remind the dimensions in this product. For the left factor, recalling the definition given in the beginning of Section D, we have  $\alpha(X, \theta) \in \mathbb{R}^{n \times \mathcal{P}}$ . Concerning the right factor, since for any  $\tau \in U_\theta$ , we have  $\psi^\theta(\tau) \in \mathbb{R}^{\mathcal{P} \times V_L}$ , the partial derivative satisfies  $\frac{\partial \psi^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta) \in \mathbb{R}^{\mathcal{P} \times V_L}$ . Hence, the dimension of the product is

$$\alpha(X, \theta) \frac{\partial \psi^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta) \in \mathbb{R}^{n \times V_L}.$$

To obtain the coefficient  $(i, v_L)$  of this product, we keep the  $i^{th}$  line of the left factor, which is equal to  $\alpha(x^i, \theta)$ , and the column  $v_L$  of the right factor, which is equal to  $\frac{\partial \psi_{v_L}^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta)$ . We thus have

$$\left[ \alpha(X, \theta) \frac{\partial \psi^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta) \right]_{i, v_L} = \alpha(x^i, \theta) \frac{\partial \psi_{v_L}^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta) = \sum_{p \in \mathcal{P}} \alpha_p(x^i, \theta) \frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta).$$

Let  $p \in \mathcal{P}$ . If  $p = (v_0, \dots, v_L) \in \mathcal{P}_0$ , looking at the case 1 in the proof of Proposition 23, we have

$$\frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta) = \mathbb{1}_{\{v_l \rightarrow v_{l+1} \in p\}} \prod_{\substack{k \in \llbracket 0, L-1 \rrbracket \\ k \neq l}} w_{v_k \rightarrow v_{k+1}}.$$

Recalling the definition of  $\alpha_p(x^i, \theta)$  in the case  $p \in \mathcal{P}_0$ , given in (2), we also have

$$\alpha_p(x^i, \theta) = x_{v_0}^i \prod_{k=1}^{L-1} a_{v_k}(x^i, \theta),$$

and thus

$$\alpha_p(x^i, \theta) \frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta) = \mathbb{1}_{\{v_l \rightarrow v_{l+1} \in p\}} x_{v_0}^i \prod_{k=1}^{L-1} a_{v_k}(x^i, \theta) \prod_{\substack{k \in [0, L-1] \\ k \neq l}} w_{v_k \rightarrow v_{k+1}}. \quad (62)$$

Now if  $p = (v_{l'}, \dots, v_L) \in \mathcal{P}_{l'}$ , for  $l' \in \llbracket 1, \dots, L-1 \rrbracket$ , looking at the case 2 in the proof of Proposition 23, we have

$$\frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta) = \mathbb{1}_{\{v_l \rightarrow v_{l+1} \in p\}} b_{v_{l'}} \prod_{\substack{k \in [l', L-1] \\ k \neq l}} w_{v_k \rightarrow v_{k+1}}.$$

Recalling the definition of  $\alpha_p(x^i, \theta)$  in the case  $p \in \mathcal{P}_{l'}$ , given in (2), we also have

$$\alpha_p(x^i, \theta) = \prod_{k=l'}^{L-1} a_{v_k}(x^i, \theta),$$

and thus

$$\alpha_p(x^i, \theta) \frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta) = \mathbb{1}_{\{v_l \rightarrow v_{l+1} \in p\}} b_{v_{l'}} \prod_{k=l'}^{L-1} a_{v_k}(x^i, \theta) \prod_{\substack{k \in [l', L-1] \\ k \neq l}} w_{v_k \rightarrow v_{k+1}}. \quad (63)$$

Finally, if  $p = \beta$ , looking at the case 3 in the proof of Proposition 23, we have

$$\frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta) = 0,$$

and thus

$$\alpha_p(x^i, \theta) \frac{\partial \psi_{p, v_L}^\theta}{\partial \tau_{v_l \rightarrow v_{l+1}}}(\tau_\theta) = 0. \quad (64)$$

Assembling (62), (63) and (64), we can sum over all  $p \in \mathcal{P}$ , and obtain

$$\begin{aligned} \gamma_{v_{l+1} \rightarrow v_l}^{i, v_L} &= \sum_{\substack{p \in \mathcal{P}_0 \\ p = (v_0, \dots, v_{L-1})}} \mathbb{1}_{\{v_l \rightarrow v_{l+1} \in p\}} x_{v_0}^i \prod_{k=1}^{L-1} a_{v_k}(x^i, \theta) \prod_{\substack{k \in [0, L-1] \\ k \neq l}} w_{v_k \rightarrow v_{k+1}} \\ &\quad + \sum_{l'=1}^L \sum_{\substack{p \in \mathcal{P}_{l'} \\ p = (v_{l'}, \dots, v_{L-1})}} \mathbb{1}_{\{v_l \rightarrow v_{l+1} \in p\}} b_{v_{l'}} \prod_{k=l'}^{L-1} a_{v_k}(x^i, \theta) \prod_{\substack{k \in [l', L-1] \\ k \neq l}} w_{v_k \rightarrow v_{k+1}} \end{aligned}$$

which can be reformulated, getting rid of the zero sums when  $v_l \rightarrow v_{l+1} \notin p$ , as

$$\begin{aligned} \gamma_{v_{l+1} \rightarrow v_l}^{i, v_L} &= \sum_{v_0 \in V_0} x_{v_0}^i \bar{w}_{v_0 \rightarrow v_1} a_{v_1}(x^i, \theta) \prod_{\substack{k \in [1, L-1] \\ k \neq l}} a_{v_k}(x^i, \theta) w_{v_k \rightarrow v_{k+1}} \\ &\quad \vdots \\ &\quad v_{l-1} \in V_{l-1} \\ &\quad v_{l+2} \in V_{l+2} \\ &\quad \vdots \\ &\quad v_{L-1} \in V_{L-1} \\ &\quad + \sum_{l'=1}^L \sum_{\substack{v_{l'} \in V_{l'} \\ \vdots \\ v_{l-1} \in V_{l-1} \\ v_{l+2} \in V_{l+2} \\ \vdots \\ v_{L-1} \in V_{L-1}}} a_{v_{l'}}(x^i, \theta) b_{v_{l'}} \prod_{\substack{k \in [l', L-1] \\ k \neq l}} a_{v_k}(x^i, \theta) w_{v_k \rightarrow v_{k+1}}, \end{aligned}$$

which shows (60).

The proof of (61) is similar to the one of (60).  $\square$