

Path homologies of deep feedforward networks

Samir Chowdhury Thomas Gebhart Steve Huntsman Matvey Yutin
Stanford University University of Minnesota BAE Systems FAST Labs UC Berkeley and BAE Systems FAST Labs
samirc@stanford.edu gebhart@umn.edu steve.huntsman@baesystems.com myutin@berkeley.edu

Abstract—We provide a characterization of two types of directed homology for fully-connected, feedforward neural network architectures. These exact characterizations of the directed homology structure of a neural network architecture are the first of their kind. We show that the directed flag homology of deep networks reduces to computing the simplicial homology of the underlying undirected graph, which is explicitly given by Euler characteristic computations. We also show that the path homology of these networks is non-trivial in higher dimensions and depends on the number and size of the layers within the network. These results provide a foundation for investigating homological differences between neural network architectures and their realized structure as implied by their parameters.

I. INTRODUCTION

Deep neural networks have emerged as an effective class of machine learning algorithms across a wide range of domains [1]. The superior performance of these algorithms in comparison to other traditional machine learning methods may be attributed to their structure, consisting of a layer-wise composition of nonlinear functions parameterized by a set of real-valued weight matrices which define connectivity between layers [2], [3].

These parameters, governing connectivity and information flow within the network, can then be optimized by gradient descent for performance on a given task. For certain tasks, enforcing particular connectivity patterns by altering the network architecture can benefit the performance of the algorithms. These architectural priors are relevant to the functionality of deep neural networks on particular tasks as they define *a priori* a partitioning of the input information as it is processed throughout the network. Because the architecture is fixed during training, gradient descent searches for an optimal parameterization for the task given this connectivity structure.

While certain connectivity biases have proven successful in various domains [4], [5], the problem of defining an optimal connectivity pattern for a given task is still an open one. Neural architecture search [6] attacks this problem by attempting to optimize over network architectures using a variety of search strategies. However, the space of possible architectures is combinatorially large. This fact, combined with the high computational costs of training a single architectural instantiation on a task, means these search methods are severely restricted in their ability to properly span the possible space of network architectures. Better topological priors on optimal connectivity for a given task are needed to better constrain this search space. Recent work has shown that even in popular

network architectures that achieve near-optimal performance on a task, a substructure with a drastically reduced parameter set that achieves similar task performance is likely to exist [7]. This result, in combination with the effectiveness of pruning techniques [8], [9], implies that many neural networks are overparameterized by architectures that do not properly constrain the partitioning of the input across the network. However, gradient descent is still able to approximate this optimal topology, and this is reflected within the trained parameters of the network [7]. A method for determining the extent to which network-topological structures of a trained network differ from the topological structure implied by its architecture can provide insight on the proper connectivity structure for a given task. We provide an initial step in this direction with the characterization of two directed homologies of fully-connected neural networks. This framework provides the means to compare the actualized connectivity structure of parameterized neural networks to the homological structure defined by their architectures.

Understanding the topological structure of the network architecture naturally translates to a question of understanding the topological structure of the underlying directed graph. Recent developments in this direction with strong mathematical foundations include the theories of *path homology* and *directed flag complex (DFC) homology*. Path homology was developed by Grigor'yan, Lin, Muranov, and Yau [10], and an accompanying theory of digraph homotopy was later developed in [11]. This in turn was consistent with earlier notions of homotopy of graphs [12]. DFC homology was popularized via [13], and is built on top of a notion of ordered simplicial homology that has concrete mathematical foundations [14]. In both of these cases, persistent-homological frameworks have been developed recently [15], [16]. These recent developments have thus provided novel tools for approaching the problem of understanding neural architecture.

A. Contributions and statement of results

In this paper, we provide a characterization of the reduced path homology of fully-connected, feedforward neural networks (i.e. multilayer perceptrons (MLPs)) in terms of their architecture. This exact characterization of the path homology structure of a neural network is the first of its kind. Additionally, we provide a characterization of the DFC homology of MLPs, and show that it reduces to computing simplicial homology of the underlying undirected graph viewed as a simplicial complex. These results provide a starting point for investigating

differences between the inherent homological structure of a neural network architecture versus its realized (persistent) homological structure as implied by its learned parameters.

Specifically, let $K_{n_1, \dots, n_L}^{\rightarrow}$ denote the directed acyclic graph corresponding to the architecture of an MLP with L layers of widths $\{n_1, n_2, \dots, n_L\}$. See Figure 1 for an example. Individual layers (as sets of vertices) are denoted K_1, K_2, \dots, K_L . In graph-theoretic terms, $K_{n_1, \dots, n_L}^{\rightarrow}$ has node set $\cup_{i=1}^L K_i$ and edge set $\{(v, v') : v \in K_i, v' \in K_{i+1}, 0 \leq i \leq L-1\}$. Also let K_{n_1, \dots, n_L} denote the underlying undirected graph, viewed as a simplicial complex. We show that $K_{n_1, \dots, n_L}^{\rightarrow}$ has nontrivial reduced path homology (with field coefficients) precisely in degree $(L-1)$, and that this homology group has rank $\prod_{i=1}^L (n_i - 1)$.

Theorem 1. *Let $K_{n_1, \dots, n_L}^{\rightarrow}$ be the MLP with L layers of widths $\{n_1, n_2, \dots, n_L\}$. Then we have*

$$\text{rank}(\text{PathHom}_p(K_{n_1, \dots, n_L}^{\rightarrow})) = \delta_p^{L-1} \prod_{i=1}^L (n_i - 1).$$

Additionally, we show that the DFC homology of $K_{n_1, \dots, n_L}^{\rightarrow}$ reduces to the simplicial homology (denoted H_p^Δ) of K_{n_1, \dots, n_L} , which in turn counts the number of loops in K_{n_1, \dots, n_L} .

Theorem 2. *We have*

$$\text{rank}(\text{DFCHom}_p(K_{n_1, \dots, n_L}^{\rightarrow})) = \text{rank}((H_p^\Delta(K_{n_1, \dots, n_L}))).$$

Specifically, this rank is 1 for $p = 0$, $(1 - \#V + \#E)$ for $p = 1$, and 0 for $p \geq 2$. Here $\#V$ and $\#E$ are the numbers of vertices and edges, respectively.

Stated differently, Theorem 2 shows that **DFCHom** picks out the structure of (undirected) loops in the MLP architecture.

II. RELATED WORK

Work on homological approaches to neural network analyses have shown promise in the ability to extract insights about network function through the investigation of topological structure of network parameters [17], [18], [19], [20], [21], [22], [23], [24]. In [25], the authors find that analyzing the 0-dimensional persistent-homological structure of neural network parameters during training provides insight into when the training process may be considered completed. Similarly, the authors of [26] analyze the 0-dimensional persistent-homological structure of neural network activations and find that this structure is closely linked to the representations used by the network to make classification decisions. Both of the above papers make use of a Vietoris-Rips filtration over the graph defined by the network architecture. Notably, this filtration does not preserve the inherent directionality of the network defined by sequential layers. The asymmetry arising from directionality within the network is important to capture in order to faithfully represent the homological properties of the flow of information through each layer. The papers also lack a homological basis from which to compare the persistent homological structures that emerge as weights are thresholded to the inherent network homology. Our result in Section I-A

provides such a homological basis for path homology of fully-connected neural networks from which one may compare the actualized homological structure of the weighted network versus the homological structure defined by its architecture.

The prior discussion relates to investigating the topological structure of the network architecture; there have also been very recent developments showing how to backpropagate a topological loss function through a deep neural network [20], [27]. These techniques build on insights developed throughout [28], [29], [30]. While our methods are grounded in constructions arising from topology, a geometric viewpoint of ReLU deep networks has been introduced in [31].

III. PRELIMINARIES

In this section, we provide background material on homology, path homology, and DFC homology. We refer the reader to [14] for additional details on homology (specifically §1.13 for ordered homology), [13] for details on DFC homology, and [11] for details on path homology.

We write \mathbb{Z}_+ to denote the nonnegative integers. Fix a field \mathbb{K} . A *chain complex* is defined to be a sequence of vector spaces $(C_p)_{p \in \mathbb{Z}}$ over \mathbb{K} and *boundary maps* $(\partial_p : C_p \rightarrow C_{p-1})_{k \in \mathbb{Z}}$ satisfying the condition $\partial_{p-1} \circ \partial_p = 0$ for each $p \in \mathbb{Z}$. We often denote a chain complex as $\mathcal{C} = (C_p, \partial_p)_{p \in \mathbb{Z}}$. Given a chain complex \mathcal{C} and any $p \in \mathbb{Z}$, one defines the following:

$$\begin{aligned} Z_p(\mathcal{C}) &:= \ker(\partial_p) = \{c \in C_p : \partial_p(c) = 0\}, \text{ the } p\text{-cycles,} \\ B_p(\mathcal{C}) &:= \text{im}(\partial_{p+1}) = \{c \in C_p : c = \partial_{p+1}(b) \text{ for some} \\ &\quad b \in C_{p+1}\}, \text{ the } p\text{-boundaries.} \end{aligned}$$

The quotient vector space $H_p(\mathcal{C}) := Z_p(\mathcal{C})/B_p(\mathcal{C})$ is called the *p-th homology vector space of the chain complex \mathcal{C}* . The dimension of $H_p(\mathcal{C})$ is called the *p-th Betti number* of \mathcal{C} , denoted $\beta_p(\mathcal{C})$. These vector spaces can be made to arise from data via the following construction.

A *simplicial complex* Σ built on a set S is (abstractly) a collection of subsets $\sigma \subseteq S$ such that whenever $\tau \subseteq \sigma \in \Sigma$, we have $\tau \in \Sigma$. The $(p+1)$ -length elements of Σ are referred to as *p-simplices*. The elements of a simplex are called *vertices*. We additionally fix an arbitrary total ordering on S . Different orderings of the vertices of a simplex are considered equivalent if they differ by an even permutation. Thus a *p-simplex* $\sigma \in S$ for $p \geq 1$ belongs to two equivalence classes, and each class is called an *orientation* of σ . For each $p \in \mathbb{Z}_+$, we write Σ_p to denote the *p-simplices* of Σ .

The standard construction of a chain complex from a simplicial complex is obtained by defining $C_p(\Sigma)$ to be the free vector space over Σ_p for each $p \geq 0$, with coefficients in \mathbb{K} , along with the relation $\sigma = -\tau$ if τ differs from σ by an odd permutation. Additionally one defines $C_{-1}(\Sigma) = \mathbb{K}$ and $C_p(\Sigma) = \{0\}$ for $p \leq -2$ (this corresponds to *reduced* homology). Finally, for any $p \in \mathbb{Z}_+$, one defines a linear map $\partial_p : C_p \rightarrow C_{p-1}$ to be the linearization of the following map on the generators of C_p :

$$\partial_p([x_0, \dots, x_p]) := \sum_{i=0}^p (-1)^i [x_0, \dots, \hat{x}_i, \dots, x_p], \quad (1)$$

for each p -simplex $[x_0, \dots, x_p] \in C_p$. Here \widehat{x}_i denotes omission of x_i from the sequence. Additionally, ∂_p is defined to be the zero map for $p \leq -1$. These constructions fully determine simplicial homology, which we denote by H_p^Δ .

An undirected graph $G = (V, E)$ has a natural representation as a simplicial complex: the nodes are 0-simplices, and the edges are 1-simplices. There are no higher-dimensional simplices. We write $\Sigma_p(G)$ to denote the p -simplices of G .

A. Path homology

Given a finite set X and any integer $p \in \mathbb{Z}_+$, an *elementary p -path over X* is a sequence (x_0, \dots, x_p) of $p+1$ elements of X . For each $p \in \mathbb{Z}_+$, the free vector space consisting of all formal linear combinations of elementary p -paths over X with coefficients in \mathbb{K} is denoted $\Lambda_p = \Lambda_p(X) = \Lambda_p(X, \mathbb{K})$. One also defines $\Lambda_{-1} := \mathbb{K}$ and $\Lambda_p := \{0\}$ for $p \leq -2$. The boundary maps are defined as in Equation (1), and we overload notation to denote them by ∂_p as before. It follows that $(\Lambda_p, \partial_p)_{p \in \mathbb{Z}}$ is a chain complex.

Next let $G = (X, E)$ be a digraph. For each $p \in \mathbb{Z}_+$, one defines an elementary p -path (x_0, \dots, x_p) on X to be *allowed* if $(x_i, x_{i+1}) \in E$ for each $0 \leq i \leq p-1$. For each $p \in \mathbb{Z}_+$, the free vector space on the collection of allowed p -paths on (X, E) is denoted $\mathcal{A}_p = \mathcal{A}_p(G) = \mathcal{A}_p(X, E, \mathbb{K})$, and is called the *space of allowed p -paths*. One further defines $\mathcal{A}_{-1} := \mathbb{K}$ and $\mathcal{A}_p := \{0\}$ for $p \leq -2$.

The allowed paths do not form a chain complex, because the image of an allowed path under ∂ need not be allowed. This is rectified as follows. Given a digraph $G = (X, E)$ and any $p \in \mathbb{Z}$, the *space of ∂ -invariant p -paths on G* is defined to be the following subspace of $\mathcal{A}_p(G)$:

$$\Omega_p = \Omega_p(G) = \Omega_p(X, E, \mathbb{K}) := \{c \in \mathcal{A}_p : \partial_p(c) \in \mathcal{A}_{p-1}\}.$$

It follows by the definitions that $\text{im}(\partial_p(\Omega_p)) \subseteq \Omega_{p-1}$ for any integer $p \geq -1$. Thus we have a chain complex:

$$\dots \xrightarrow{\partial_3} \Omega_2 \xrightarrow{\partial_2} \Omega_1 \xrightarrow{\partial_1} \Omega_0 \xrightarrow{\partial_0} \mathbb{K} \xrightarrow{\partial_{-1}} 0$$

For each $p \in \mathbb{Z}_+$, the *p -dimensional (reduced) path homology groups (denoted H_p^Ξ)* of $G = (X, E)$ are defined as:

$$H_p^\Xi(G) = H_p^\Xi(X, E, \mathbb{K}) := \ker(\partial_p) / \text{im}(\partial_{p+1}).$$

Note that this definition of path homology is slightly different from the convention in [11], where path homology refers to a version of the above (the non-reduced version) where Ω_{-1} is defined to be $\{0\}$.

B. Directed flag complex homology

The *directed flag complex* of a directed graph $G = (X, E)$ is the collection of finite sequences (x_0, x_1, \dots, x_n) , for $n \in \mathbb{Z}_+$, such that $x_i \rightarrow x_j$ whenever $i < j$. Such finite sequences are referred to as *directed n -simplices*. For each $p \in \mathbb{Z}_+$, we write $\mathcal{F} := \mathcal{F}_p(G)$ to denote the free vector space over directed p -simplices in G . Then the boundary map ∂ from Equation (1) can be overloaded to give a map $\partial_p : \mathcal{F}_p \rightarrow \mathcal{F}_{p-1}$. The

directed flag complex (DFC) homology (denoted $H_p^\mathcal{F}$) of G is then defined as:

$$H_p^\mathcal{F}(G) = H_p^\mathcal{F}(X, E, \mathbb{K}) := \ker(\partial_p) / \text{im}(\partial_{p+1}).$$

IV. DFC HOMOLOGY OF MLPs

Proof of Theorem 2. First we observe that $\mathcal{F}_p(K_{n_1, \dots, n_L}^\rightarrow) = \{0\}$ for each $p \geq 2$, as there are no ‘‘skip connections’’ from any layer i to a layer $i+j$ for $j \geq 2$. The remainder of the proof will occur at the level of chain complexes, so we introduce some notation for convenience. We write $C_*^\mathcal{F}$ to denote the chain complex arising from $\mathcal{F}_*(K_{n_1, \dots, n_L}^\rightarrow)$, and C_*^Δ to denote the chain complex arising from $\Sigma_*(K_{n_1, \dots, n_L})$. Also let $\partial_p^\mathcal{F}$ denote the boundary map applied to $C_p^\mathcal{F}$ and let ∂_p^Δ denote the boundary map applied to C_p^Δ . We will typically overload notation and just use ∂ , but the distinction will occasionally be used to clarify context. It is immediate that $C_p^\mathcal{F} = C_p^\Delta$ for $p \leq 0$ and $p \geq 2$, so we only verify this equality for $p = 1$.

Note that $C_1^\mathcal{F}$ is generated by elements of the form (v, v') where $v \rightarrow v'$, and C_1^Δ is generated by elements of the form $[v, v']$ where either $v \rightarrow v'$ or $v' \rightarrow v$. By using the identity $[v, v'] = -[v', v]$, we write each chain $\sigma \in C_1^\Delta$ as $\sigma = \sum_{i=1}^k c_i [u_i, u'_i]$, where $u_i \rightarrow u'_i$. The crucial point is that in an MLP, we will only have either $v \rightarrow v'$ or $v' \rightarrow v$, but not both. Thus the chains of C_1^Δ as written above are exactly the chains of $C_1^\mathcal{F}$, and the boundary maps are exactly the same for both $C_1^\mathcal{F}$ and C_1^Δ . Thus we have $H_p^\Delta = H_p^\mathcal{F}$ for $p \geq 0$. The second statement follows from standard results on the Euler characteristic of a connected graph. \square

V. PATH HOMOLOGY OF MLPs

Prior to providing the proof of Theorem 1, we digress briefly to highlight an interesting connection. The theory of path homology admits Künneth formulas for various digraph constructions [32], and one might expect that the layered construction of feedforward neural networks would be amenable to applying such formulas. Indeed, when restricted to feedforward network architectures having two layers, the Künneth formula for join applies to give the result in Theorem 1. However, this approach seems not to work for networks with more layers. The immediate obstruction is that feedforward networks with more than two layers do not arise as the join of the individual layers.

Attempts to prove a Künneth formula for a generalized version of digraph join also seem to fail due to the structure of the boundary map ∂_p . Except for this failure (which is not obvious), a simple proof strategy along the following lines would appear convincing at first: introduce an L -ary digraph join and appeal to associativity of tensor products (or more elaborately, to an L -ary Künneth formula [33]) to argue that

$$\begin{aligned} & \text{rank}(H_p^\Xi(K_{n_1, \dots, n_L}^\rightarrow)) \\ &= \sum_{\alpha \in \mathbb{N}^L : \sum_{\ell=1}^L \alpha_\ell = p - (L-1)} \prod_{\ell=1}^L \text{rank}(H_{\alpha_\ell}^\Xi(\{1, \dots, n_\ell\})). \end{aligned}$$

From here, the desired result would follow from a straightforward calculation. It would thus be interesting to see if such a

Künneth formula could provide an alternative proof of Theorem 1. We remark that Künneth formulae in *persistent homology* have been studied in [34], [35].

We now proceed to the proof of Theorem 1.

Lemma 3 (Product rule, [32]). *Let $u \in \Lambda_p$ and $v \in \Lambda_q$. Then:*

$$\partial(uv) = (\partial u)v + (-1)^{p+1}u(\partial v).$$

Notation. We adopt some extra notation for readability. Given (the underlying digraph of) an MLP $K_{n_1, \dots, n_L}^{\rightarrow}$, we write \mathcal{A}_p^L to denote $\mathcal{A}_p(K_{n_1, \dots, n_L}^{\rightarrow})$. We define Ω_p^L and ∂_p^L analogously. Recall that K_i denotes the i th layer of $K_{n_1, \dots, n_L}^{\rightarrow}$.

To prove the theorem, we need to understand the kernel of ∂_{L-1}^L . The next proposition gives a representation of each such kernel element in terms of an $(L-1)$ -layer MLP.

Proposition 4. *Consider the map ∂_{L-1}^L defined on Ω_{L-1}^L . Any element $\gamma \in \ker(\partial_{L-1}^L)$ can be written as a finite sum*

$$\gamma = \sum_{i=1}^d w_i v_i, \quad w_i \in \ker(\partial_{L-2}^{L-1}), v_i \in K_L, d \geq 1.$$

Proof. By the structure of $K_{n_1, \dots, n_L}^{\rightarrow}$, all $(L-1)$ -paths have the form $(v^{(1)}v^{(2)} \dots v^{(L)})$, where each $v^{(i)} \in K_i$. Thus any element in $\ker(\partial_{L-1}^L)$ has the form

$$\gamma = \sum_{i=1}^d w_i v_i \in \Omega_{L-1}(K_{n_1, \dots, n_L}^{\rightarrow}), \quad w_i \in \mathcal{A}_{L-2}^{L-1}, v_i \in K_L.$$

Here we assume WLOG that the v_i are distinct. By Lemma 3:

$$0 = \partial(\gamma) = \sum_{i=1}^d (\partial w_i)v_i + (-1)^{L-1}w_i.$$

By linear independence, we must have $(-1)^{L-1} \sum_{i=1}^d w_i = 0$ and $\sum_{i=1}^d (\partial w_i)v_i = 0$. Since the v_i are all distinct, we have $\partial w_i = 0$ and hence $w_i \in \ker(\partial_{L-2}^{L-1})$ for each $1 \leq i \leq d$. This concludes the proof. \square

The next proposition further clarifies the preceding representation as a difference of basis terms.

Proposition 5. *Consider the map ∂_{L-2}^{L-1} defined on Ω_{L-2}^{L-1} , and let B_{L-2} be a basis for $\ker(\partial_{L-2}^{L-1})$. Let $\gamma \in \ker(\partial_{L-1}^L)$. Then we can write*

$$\gamma = \sum_{i=1}^{|B_{L-2}|} \sum_{j,k=1}^{|K_L|} c_{ijk} u_i(v_j - v_k),$$

where $c_{ijk} \in \mathbb{K}$, $u_i \in B_{L-2}$, and $v_j, v_k \in K_L$.

Proof. Using Proposition 4, we write $\gamma = \sum_{i=1}^d c_i u_i v_i$, where $d \geq 1$, $c_i \in \mathbb{K}$, $u_i \in B_{L-2}$, and $v_i \in K_L$. This allows for degeneracy, in the sense that we may have $u_i = u_j$ for $i \neq j$, and likewise for c_i and v_i . By Lemma 3, we obtain $0 = \partial(\gamma) = \sum_{i=1}^d c_i u_i$. Here we have used the relation $\partial(u_i) = 0$.

Next fix $i = 1$, and let $I_1 \subseteq \{1, 2, \dots, d\}$ denote the indices j for which $u_j = u_1$. Since $\sum_{i=1}^d c_i u_i = 0$, we have by linear independence that $\sum_{j \in I_1} c_j = 0$. Next let $I_1^+ \sqcup I_1^-$ be a

partition of I_1 into nonempty sets. Then we have $\sum_{j \in I_1^+} c_j = \sum_{k \in I_1^-} -c_k$ by the preceding observation. It follows that

$$\sum_{j \in I_1} c_j u_1 v_j = \sum_{j \in I_1} c_j u_1 v_j = \sum_{j \in I_1^+} c_j u_1 v_j - \sum_{k \in I_1^-} (-c_k) u_1 v_k.$$

We would like to write the latter as a sum of elements of the form $c_{jk} u_1(v_j - v_k)$. The problem of determining the coefficients c_{jk} can be phrased as a supply-demand problem, and we present this next. For now we assume $\sum_{j \in I_1^+} c_j = \sum_{k \in I_1^-} -c_k \neq 0$.

Define the *supply* vector r to be an $|I_1^+| \times 1$ column vector with entries c_j , $j \in I_1^+$. Also define the *demand* vector s to be the $1 \times |I_1^-|$ row vector with entries $-c_k$, $k \in I_1^-$. We need to construct a nonnegative $|I_1^+| \times |I_1^-|$ matrix T^1 with row and column sums equal to r and s , respectively. Define T^1 by writing $T_{jk}^1 := r_j s_k / \sum_{l \in I_1^+} c_l$ for each j, k .

To verify that T^1 has the desired row and column sums, recall that $\sum_{k \in I_1^-} s_k = \sum_{k \in I_1^-} -c_k = \sum_{j \in I_1^+} c_j$, and so

$$\sum_{k \in I_1^-} T_{jk}^1 = r_j \sum_{k \in I_1^-} s_k / \sum_{l \in I_1^+} c_l = r_j \sum_{j \in I_1^+} c_j / \sum_{l \in I_1^+} c_l = r_j.$$

Similarly, recall $\sum_{j \in I_1^+} r_j = \sum_{j \in I_1^+} c_j$, and so

$$\sum_{j \in I_1^+} T_{jk}^1 = s_k \sum_{j \in I_1^+} r_j / \sum_{l \in I_1^+} c_l = s_k \sum_{j \in I_1^+} c_j / \sum_{l \in I_1^+} c_l = s_k.$$

Thus we obtain:

$$\sum_{j \in I_1^+} c_j u_1 v_j - \sum_{k \in I_1^-} (-c_k) u_1 v_k = \sum_{j \in I_1^+} \sum_{k \in I_1^-} T_{jk}^1 u_1 (v_j - v_k).$$

It may be the case that $v_j = v_{j'}$ for $j \neq j' \in I_1^+$, and likewise for I_1^- . By summing such terms and by padding T^1 with zeros if necessary, we create a $|K_L| \times |K_L|$ matrix C^1 satisfying:

$$\sum_{j \in I_1^+} \sum_{k \in I_1^-} T_{jk}^1 u_1 (v_j - v_k) = \sum_{j,k=1}^{|K_L|} C_{jk}^1 u_1 (v_j - v_k). \quad (2)$$

Now we return to the case $\sum_{j \in I_1^+} c_j = \sum_{k \in I_1^-} -c_k = 0$. In this case, we further subdivide I_1^+ into nonempty sets $I_1^{++} \sqcup I_1^{+-}$. If $\sum_{j \in I_1^{++}} c_j = \sum_{k \in I_1^{+-}} -c_k \neq 0$, then we can proceed as before to obtain a decomposition as in Equation (2), and otherwise we can continue subdividing the index set. Since there are only finitely many terms, the subdivision operation must terminate in a finite number of steps. Similarly one subdivides I_1^- as necessary to collect terms in the form of Equation (2).

Repeating this process for $i = 2, \dots, d$, we obtain:

$$\gamma = \sum_{i=1}^d c_i u_i v_i = \sum_{i=1}^{|B_{L-2}|} \sum_{j,k=1}^{|K_L|} C_{jk}^i u_i (v_j - v_k). \quad \square$$

Proposition 6. *Given the setup of Proposition 5 and $\gamma \in \ker(\partial_{L-1}^L)$, we can further write:*

$$\gamma = \sum_{i=1}^{|B_{L-2}|} \sum_{j=2}^{|K_L|} c_{ij} u_i (v_1 - v_j),$$

where $c_{ij} \in \mathbb{K}$, $u_i \in B_{L-2}$, and $v_j \in K_L$. Consequently, $B_{L-1} := \{u_i(v_1 - v_j) : u_i \in B_{L-2}, 2 \leq j \leq |K_L|\}$ forms a basis for $\ker(\partial_{L-1}^L)$.

Proof. We repeat the first few steps of Proposition 5; namely we use Proposition 4 to write $\gamma = \sum_{i=1}^d c_i u_i v_i$, observe $\sum_{i=1}^d c_i u_i = 0$, and collect u_i terms to write

$$\gamma = \sum_{i=1}^{|B_{L-2}|} u_i \sum_{j=1}^{|K_L|} c_{ij} v_j.$$

We further write

$$\gamma = \sum_{i=1}^{|B_{L-2}|} u_i \left(\sum_{j=1}^{|K_L|} c_{ij} (v_j - v_1) + \sum_{j=1}^{|K_L|} c_{ij} v_1 \right)$$

and observe that, by the preceding observation:

$$\sum_{i=1}^{|B_{L-2}|} u_i \sum_{j=1}^{|K_L|} c_{ij} v_1 = \sum_{i=1}^d c_i u_i v_1 = 0. \quad \square$$

Proposition 7. *Let K be a two-layer MLP, and consider the boundary map ∂_1^2 on $\Omega_1(K)$. Let u_1, u_2, \dots, u_{n_1} and v_1, v_2, \dots, v_{n_2} denote the vertices of the first and second layers of K , respectively. Then $\ker(\partial_1^2)$ is generated by the elements $\{(u_1 - u_j)(v_1 - v_k) : 2 \leq j \leq n_1, 2 \leq k \leq n_2\}$. In particular, $\dim(\ker(\partial_1^2)) = (n_1 - 1)(n_2 - 1)$.*

Proof. Let K_1 denote the first layer of K , and recall that ∂_0^1 denotes the boundary map defined on $\Omega_0(K_1)$. Since we are computing reduced homology, $\ker(\partial_0^1)$ is generated by elements of the form $u_j - u_k$. We rewrite this as $u_j - u_k = u_j - u_1 + u_1 - u_k = -(u_1 - u_j) + (u_1 - u_k)$. Thus a basis for $\ker(\partial_0^1)$ is given by $\{u_1 - u_j : 2 \leq j \leq |K_1|\}$. An application of Proposition 6 completes the proof. \square

Proposition 8. *Given an MLP $K_{n_1, \dots, n_L}^{\rightarrow}$ with L layers, we have $\ker(\partial_j^L) = \text{im}(\partial_{j+1}^L)$ for each $0 \leq j \leq L - 2$.*

Proof. Let $0 \leq j \leq L - 2$, and let $\gamma \in \ker(\partial_j^L)$. Then γ is a linear combination of paths of length $(j+1)$. The endpoints of these paths may belong to layers K_{j+1}, \dots, K_L . By collecting paths ending at the same layer, write $\gamma = \sum_{l=j+1}^L \gamma_l$, where each γ_l consists of the summands of γ ending at K_l (and is 0 if there are no such summands). By linear independence, we must individually have $\partial(\gamma_l) = 0$ for each $k+1 \leq l \leq L$.

Let $k+1 \leq l < L$, and let $z_{l+1} \in K_{l+1}$. By Lemma 3,

$$\partial(\gamma_l z_{l+1}) = (\partial \gamma_l) z_{l+1} + (-1)^{j+1} \gamma_l (\partial z_{l+1}) = (-1)^{j+1} \gamma_l.$$

Next fix $l = L$, and note that γ_L is a linear combination of paths that start at K_{L-j} and end at K_L . Fix $z' \in K_{L-(j+1)}$. Then we have:

$$\partial(z' \gamma_L) = (\partial z') \gamma_L + (-1)^j z' (\partial \gamma_L) = \gamma_L.$$

Finally define $\zeta = \sum_{l=j+1}^{L-1} (-1)^{j+1} \gamma_l z_{l+1} + z' \gamma_L$. By the previous work, we have $\partial(\zeta) = \sum_{l=j+1}^L \gamma_l = \gamma$. By construction, $\zeta \in \Omega_{j+1}^L$. This shows that $\ker(\partial_j^L) = \text{im}(\partial_{j+1}^L)$ and concludes the proof. \square



Fig. 1. (L) A MLP with $L = 3$ layers and $(n_1, n_2, n_3) = (4, 10, 3)$, with weight magnitudes indicated by arc thickness and signs indicated by color (red = negative; blue = positive). (R) A subgraph obtained by removing weights with magnitude below the median.

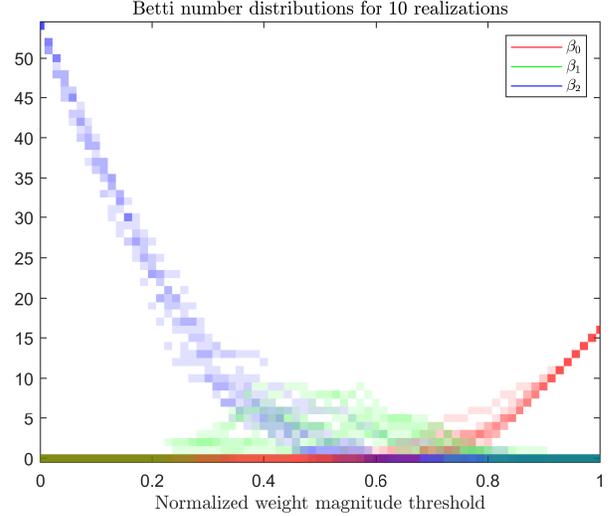


Fig. 2. Path homology Betti numbers β_p of a trained 3-layer fully-connected network $((n_1, n_2, n_3) = (4, 10, 3))$ across normalized weight magnitude thresholds and for 10 realizations of random initial weights. The distribution of Betti numbers is indicated by opacity. The threshold normalization sends the j th smallest threshold value to j/T , where T is the number of nontrivial threshold values. Note that the network starts with path homology concentrated in degree 2, as predicted by Theorem 1. Surprising features include the gradual downward cascade of β_2 as well as the ‘‘bump’’ in β_1 .

Now we proceed to the proof of Theorem 1.

Proof. $K_{n_1, \dots, n_L}^{\rightarrow}$ has no j -paths for $j \geq L$. Thus by Proposition 8, there can only be nontrivial reduced path homology in degree $L - 1$. Applying Proposition 6 inductively while using Proposition 7 as a base case, we see that $\ker(\partial_{L-1}^L)$ has dimension $\prod_{i=1}^L (n_i - 1)$. Since there are no L -paths, $\text{im}(\partial_L^L)$ is trivial. The result follows. \square

VI. DISCUSSION

In this work, we provide characterization results for two types of digraph homologies as applied to feedforward neural network architectures. Our results show that these two homology theories, while similar in structure, yield quite different outputs (with different interpretations) when applied to deep networks. From this perspective, it is important to utilize both types of digraph homology when studying neural architectures.

As an example of the utility of path homology to characterize neural networks, we used Fisher’s classical iris data set [36] to train a MLP with $L = 3$ layers and $(n_1, n_2, n_3) = (4, 10, 3)$ using MATLAB’s `patternnet` function. We performed multiple training runs with different realizations of random initial conditions. For each realization, we extracted the trained

weight matrix A . We then computed the path homology of the DAG obtained by removing arcs corresponding to weights less than a given nontrivial value in A (details of this and other experimental results obtained using the path homology algorithm and its implementation will be provided elsewhere). The results are shown in Figure 2. The figure shows that the path homology of a filtered subgraph of a fully-connected network does not suddenly vanish as edges are removed. Rather, the Betti numbers gradually “cascade down” from top to zero dimension. This phenomenology suggests that path homology can give particularly detailed topological characterizations of neural networks. For example, given a weight matrix, the corresponding filtered (or persistent) path homology of subnetworks obtained by restricting attention to a moving window of a few adjacent layers can help identify notional sub-networks that exhibit functional specificity.

With further homological characterization of directed graph motifs like those presented in this paper, persistent homological structure derived empirically from analyses of network parameters can be related back to network architectures. This link between parameterized neural network topology and neural network architecture provides actionable insight in the architectural design process and can constrain the architecture search space, as one can define architectures that *a priori* better suit the observed homological structure learned by a network for a given task.

Acknowledgements. We would like to thank Guilherme Vituri, Michael Robinson, and Bernard McShea for useful discussions. We are very grateful to the anonymous reviewers for their useful suggestions.

REFERENCES

- [1] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *nature*, vol. 521, no. 7553, p. 436, 2015.
- [2] K. Hornik, “Approximation capabilities of multilayer feedforward networks,” *Neural networks*, vol. 4, no. 2, pp. 251–257, 1991.
- [3] H. Mhaskar, Q. Liao, and T. Poggio, “When and why are deep networks better than shallow ones?,” in *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.
- [4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems*, pp. 1097–1105, 2012.
- [5] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
- [6] T. Elsken, J. H. Metzen, and F. Hutter, “Neural architecture search: A survey,” *Journal of Machine Learning Research*, vol. 20, no. 55, pp. 1–21, 2019.
- [7] J. Frankle and M. Carbin, “The lottery ticket hypothesis: Finding sparse, trainable neural networks,” *arXiv preprint arXiv:1803.03635*, 2018.
- [8] S. Han, J. Pool, J. Tran, and W. J. Dally, “Learning both weights and connections for efficient neural networks. 2015,” URL <https://arxiv.org/abs/1506.02626>.
- [9] H. Li, A. Kadav, I. Durdanovic, H. Samet, and H. P. Graf, “Pruning filters for efficient convnets,” *arXiv preprint arXiv:1608.08710*, 2016.
- [10] A. Grigor’yan, Y. Lin, Y. Muranov, and S.-T. Yau, “Homologies of path complexes and digraphs,” *arXiv preprint arXiv:1207.2834*, 2012.
- [11] A. Grigor’yan, Y. Lin, Y. Muranov, and S.-T. Yau, “Homotopy theory for digraphs,” *Pure and Applied Mathematics Quarterly*, vol. 10, no. 4, pp. 619–674, 2014.
- [12] E. Babson, H. Barcelo, M. de Longueville, and R. Laubenbacher, “Homotopy theory of graphs,” *Journal of Algebraic Combinatorics*, vol. 24, no. 1, pp. 31–44, 2006.
- [13] M. W. Reimann, M. Nolte, M. Scolamiero, K. Turner, R. Perin, G. Chindemi, P. Dłotko, R. Levi, K. Hess, and H. Markram, “Cliques of neurons bound into cavities provide a missing link between structure and function,” *Frontiers in computational neuroscience*, vol. 11, p. 48, 2017.
- [14] J. R. Munkres, *Elements of algebraic topology*, vol. 7. Addison-Wesley Reading, 1984.
- [15] S. Chowdhury and F. Mémoli, “Persistent path homology of directed networks,” in *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 1152–1169, SIAM, 2018.
- [16] K. Turner, “Rips filtrations for quasimetric spaces and asymmetric functions with stability results,” *Algebraic & Geometric Topology*, vol. 19, no. 3, pp. 1135–1170, 2019.
- [17] G. Carlsson and R. B. Gabrielsson, “Topological approaches to deep learning,” *arXiv preprint arXiv:1811.01122*, 2018.
- [18] Q. Nguyen, “On connected sublevel sets in deep learning,” in *Proceedings of ICMML*, 2019.
- [19] M. Gabella and N. Afambo, “Topology of learning in artificial neural networks,” <https://arxiv.org/abs/1902.08160>, 2019.
- [20] R. Brüel-Gabrielsson, B. J. Nelson, A. Dwaraknath, P. Skraba, L. J. Guibas, and G. Carlsson, “A topology layer for machine learning,” *arXiv preprint arXiv:1905.12200*, 2019.
- [21] W. H. Guss and R. Salakhutdinov, “On characterizing the capacity of neural networks using algebraic topology,” *arXiv preprint arXiv:1802.04443*, 2018.
- [22] C. Hofer, R. Kwitt, M. Niethammer, and A. Uhl, “Deep learning with topological signatures,” in *Advances in Neural Information Processing Systems*, pp. 1634–1644, 2017.
- [23] M. Moor, M. Horn, B. Rieck, and K. Borgwardt, “Topological autoencoders,” *arXiv preprint arXiv:1906.00722*, 2019.
- [24] C. Hofer, R. Kwitt, M. Dixit, and M. Niethammer, “Connectivity-optimized representation learning via persistent homology,” *arXiv preprint arXiv:1906.09003*, 2019.
- [25] B. Rieck, M. Togninalli, C. Bock, M. Moor, M. Horn, T. Gumbsch, and K. Borgwardt, “Neural persistence: A complexity measure for deep neural networks using algebraic topology,” in *International Conference on Learning Representations (ICLR)*, 2019.
- [26] T. Gebhart, P. Schrater, and A. Hylton, “Characterizing the shape of activation space in deep neural networks,” <https://arxiv.org/abs/1901.09496>, 2019.
- [27] X. Hu, L. Fuxin, D. Samaras, and C. Chen, “Topology-preserving deep image segmentation,” *arXiv preprint arXiv:1906.05404*, 2019.
- [28] C. Chen, X. Ni, Q. Bai, and Y. Wang, “A topological regularizer for classifiers via persistent homology,” *arXiv preprint arXiv:1806.10714*, 2018.
- [29] A. Poulernard, P. Skraba, and M. Ovsjanikov, “Topological function optimization for continuous shape matching,” in *Computer Graphics Forum*, vol. 37, pp. 13–25, Wiley Online Library, 2018.
- [30] R. B. Gabrielsson, V. Ganapathi-Subramanian, P. Skraba, and L. J. Guibas, “Topology-aware surface reconstruction for point clouds,” *arXiv preprint arXiv:1811.12543*, 2018.
- [31] N. Lei, Z. Luo, S.-T. Yau, and D. X. Gu, “Geometric understanding of deep learning,” *arXiv preprint arXiv:1805.10451*, 2018.
- [32] A. Grigor’yan, Y. V. Muranov, and S.-T. Yau, “Homologies of graphs and künneth formulas,” *Comm. Anal. Geom.*, vol. 25, p. 969, 2017.
- [33] T. W. Hungerford, “Multiple Künneth formulas for abelian groups,” *Trans. Am. Math. Soc.*, vol. 118, p. 257, 1965.
- [34] G. Carlsson and B. Filippenko, “Persistent homology of the sum metric,” *arXiv preprint arXiv:1905.04383*, 2019.
- [35] H. Gakhar and J. Perea, “Künneth formulae in persistent homology,” *arXiv preprint arXiv:1910.05656*, 2019.
- [36] R. A. Fisher, “The use of multiple measurements in taxonomic problems,” *Ann. Eugenics*, vol. 7, p. 179, 1936.