Neural Networks and Multipersistence

Maria Walch

August 2022

1 Introduction

In fig 1 an overview over the process starting from data to multiparamter persistence modules is shown. In order to fit data processed in neuronal networks into this framework, the first step is to find a multifiltration.

$$\boxed{\text{Data}} \Rightarrow \boxed{\text{Multifiltration}} \xRightarrow{\text{Homology}} \boxed{\text{Multiparameter Persistence Module}}$$

Figure 1: Overview over the Multiparamter persistence process. Figure taken from [BL22].

Definition 1.1 (Filtration)**.** Let $\mathcal{I}$ be one of the poset categories $\mathbb{N}$, $\mathbb{Z}$ or $\mathbb{R}$ or the respective opposite categories thereof. Then an $\mathcal{I}$-indexed filtration is a functor $F : \mathcal{I} \to \textbf{Top}$ such that $F_r \subset F_s$ whenever $r \leq s$.

1.1 Filtration Dimension 1: Filtration over the Chain Complex

Definition 1.2 (Sublevel/ Superlevel Filtration)**.** For T any topological space and $\gamma : T \to \mathbb{R}$, define the *sublevel-filtration* $S^{\uparrow}(\gamma)_r = \{p \in T \mid \gamma(p) \leq r\}$. Likewise, define the *superlevel filtration* to be the $\mathbb{R}^{op}$-indexed filtration given by $S^{\downarrow}(\gamma)_r = \{p \in T \mid \gamma(p) \geq r\}$.

For TDA, the most common sublevel filtration is the following: Let $P \subset \mathbb{R}^n$ be a finite set (we might also call it a *point cloud*), and let $d_P : \mathbb{R}^n \to [0, \infty)$ be the distance to P, i.e.

$$d_P(x) = \min_{y \in P} ||x - y||, \tag{1}$$

then $S^{\uparrow}(d_P)_r$ is the union of balls of radius r centered at P. This is also known as *offset filtration* of P and is topologically equivalent to the *Čech filtration* and a sub-filtration thereof called the *Delaunay filtration*. Another common and also for higher dimensions computable filtration is the *Rips filtration*. Let X be a finite metric space, $r \in [0, \infty)$ and $N(X)_r$ be the neighborhood graph of X, i.e. the graph with vertex set X containing the edge $[x, y]$ if and only if $d(x, y) \leq r$. Let $\mathrm{Rips}(X)_r$ denote the clique complex on $N(X)_r$, i.e. the largest simplicial complex with 1-skeleton $N(X)_r$. These simplicial complexes assemple into a filtration $\mathrm{Rips}(X) : [0, \infty) \to \mathbf{Simp}$ called the (Vietoris-)Rips filtration.

1.2 Filtration Dimension 2: Filtration over the Network Function

Point cloud data sets $P \subset \mathbb{R}^n$ endowed with a function $\gamma : P \to \mathbb{R}$ give a natural example for multi-parameter filtrations. For time-varying point cloud data $\gamma(x)$ may be the time of appearance of a data point. In the network context, there are several scenarios possible that lead to a multifiltration of the input data point cloud P_{int}. In the following, we will focus on the case of *bifiltrations*, meaning that our indexing poset $\mathcal{I}$ is two-dimensional. Possible functors are then : $F : \mathcal{I} \to \mathbf{Top}$, where $\mathcal{I} = \mathbb{R}^2$, $\mathcal{I} = \mathbb{R}^{op} \times \mathbb{R}$, or $\mathcal{I} = \mathbb{N}^{op} \times \mathbb{R}$ as bifiltrations, assuming $F_s \subset F_t$ whenever $s \leq t$ (componentwise).

1.2.1 Network Function

We assume our network to be a function $\mathcal{F} : \mathbb{R}^n \to \mathbb{R}^k$, which takes an input $i \in \mathbb{R}^n$ to an output $o \in \mathbb{R}^k$, where k might or might not be different from n, depending on the network architecture respectively.

1.2.2 Layers and Weights

An activation function is a function $\phi : \mathbb{R} \to \mathbb{R}$ with certain additional properties depending on its concrete choice. Let $h^t \in \mathbb{R}^n$ be the input vector to one concrete layer. Then the p-th column of the weight matrix, together with the subsequent application of the activation function define a function $l : \mathbb{R}^n \to \mathbb{R}$ determining the p-th entry of the layer l's output vector y^t as follows:

$$y_p^t = \phi(\sum_l h_l^t W_{lp}). \tag{2}$$

By this definition, we have not one, but for N columns of the weight matrix N functions mapping one layer input to $\mathbb{R}$.

1.2.3 Loss Function

If we consider the target $t \in mathbbR^k$ to be fixed, thus we can define our loss function $\mathcal{L}$ as $\mathcal{L} : \mathbb{R}^k \to \mathbb{R}$. Thus for one forward pass, we have a function $\mathcal{L} \circ \mathcal{F} : \mathbb{R}^n \to \mathbb{R}$. By such, we have defined a function γ, i.e. $\mathcal{L} \circ \mathcal{F} : P \to \mathbb{R}$. We can use this to define a sublevel-Rips bifiltration $S^{\uparrow}(\gamma) : \mathbb{R}^2 \to \mathbf{Simp}$:

$$S^{\uparrow}(\gamma)_{a,r} = \gamma^{-1}(-\infty, a]. \tag{3}$$

In the same way, a superlevel-Rips bifiltration can be defined via $S^{\downarrow}(\gamma) : \mathbb{R}^{op} \times \mathbb{R} \to \mathbf{Simp}$:

$$S^{\downarrow}(\gamma)_{a,r} = \gamma^{-1}[a, \infty). \tag{4}$$

References

[BL22] Magnus Bakke Botnan and Michael Lesnick. *An Introduction to Multiparameter Persistence*. 2022. DOI: 10.48550/ARXIV.2203.14289. URL: https://arxiv.org/abs/2203.14289.