Research presentation

Edouard Balzin

Work on homotopy theory

Operads, Factorisation Algebras and Segal spaces. The problem of describing the multiplicative structure of k-fold loop spaces received a lot of attention in 1970s [1] and led to the birth of two formalisms to describe homotopy algebraic structures. One of them is the language of operads [21].

An operad $\mathcal{O}$ in a symmetric monoidal category $\mathcal{M}$ is a symmetric sequence of objects $\mathcal{O}(n)$ labelled by natural numbers, together with composition maps $\mathcal{O}(n) \otimes \mathcal{O}(n) \to \mathcal{O}(n+m-1)$ of suitable property [18]. An algebra over an operad is an object $X \in \mathcal{M}$ with compatible maps $\mathcal{O}(n) \otimes X^{\otimes n} \to X$. When $\mathcal{M}$ has a reasonable notion of weak equivalences (say in the model category sense [23]), then it is natural to consider algebras over an operad and operads themselves up to a weak equivalence, which can be done using various available homotopical techniques [23, 12, 20].

A very well known example is the little disk operad E_k, such that any k-fold loop space is naturally an E_k-algebra. The spaces $\mathrm{E}_k(n)$ can be taken to be homotopy equivalent to configuration spaces of n points in $\mathbb{R}^k$ [21], and, for $k = \infty$, $\mathrm{E}_\infty(n)$ is a contractible space with a free action of the symmetric group Σ_n.

Take the functor $C_\bullet : \mathsf{Top} \to \mathsf{DVect}_k$ from topological spaces to chain complexes of vector spaces sending each space X to its singular chain complex $C_\bullet(X)$. Applying $C_\bullet$ to E_k gives us DVect_k-operads $\mathbb{E}_k$, and studying $\mathbb{E}_k$-algebras have been of great interest in relation to formality problems [25], mathematical physics [11] and beyond.

Theorem 1 (Deligne conjecture [22, 25, 17, 11]). *The Hochschild complex $CH^\bullet(A, A)$ of a dg-algebra A is an $\mathbb{E}_2$-algebra.*

Due to the (arguably, fundamental) ambiguity in the choice of a model for $\mathbb{E}_2$, the proofs of this statement involve quite a lot of non-trivial combinatorics. There is, however, another approach to $\mathbb{E}_k$-structures, which goes back to [7] and defines $\mathbb{E}_k$-algebras in chain complexes as constructible sheaves $\mathcal{A}$ on the Ran space of a k-disk. This approach encodes the multiplication via the diagrams of the following shape

$$\mathcal{A}(1) \otimes ... \otimes \mathcal{A}(1) \xleftarrow{\sim} \mathcal{A}(n) \longrightarrow \mathcal{A}(1) \tag{1}$$

with $\mathcal{A}(1), \mathcal{A}(n)$ being certain stalks of $\mathcal{A}$ and the left arrow being a quasi-isomorphism.

These diagrams are very similar to those which appear in the delooping machinery of Segal who introduced the notion of a Γ-space in [24]. A Γ-space is a functor $X : \Gamma \to \mathsf{Top}$ from the category Γ of pointed finite sets to topological spaces, together with Segal conditions. In effect, such a functor naturally provides us diagrams

$$X(1)^n \longleftarrow X(n) \longrightarrow X(1) \tag{2}$$

where by 1 and n we also mean sets of corresponding cardinality (plus the marked point), and the Segal conditions require the left arrow to be a homotopy equivalence. Segal Γ-spaces describe the same structure as E_∞ algebras, but it is not necessarily the case that $Z(n) \cong Y^n \times \mathrm{E}_\infty(n)$ for some model of E_∞. The algebraic operations are encoded by the category Γ, and there are examples of similar "operator" categories [6] which permit to describe E_k-algebra structures. One would like to use Segal approach to study $\mathbb{E}_k$-algebras, but beyond categories of type Top with Cartesian products, the approach ceases to function. One would thus ask if there is an extension of Segal formalism to general monoidal categories, similar to factorisation algebra approach.

Derived sections. As the category Cat of small categories is Cartesian, Segal formalism can be applied here. A symmetric monoidal structure on a category $\mathcal{M}$ is, up to an equivalence, the same data as [24, 20] a Segal Γ-space in the category of categories $M : \Gamma \to \mathsf{Cat}$ with $M(S) \cong \mathcal{M}^{S \setminus s}$, where $S \setminus s$ denotes the complement to the marked point $s \in S$. One can equivalently consider the associated Grothendieck opfibration [13] $p : \mathcal{M}^\otimes \to \Gamma$ with fibres $\mathcal{M}^\otimes(S) = p^{-1}(S) \cong \mathcal{M}^{S \setminus s}$. Any algebra A in a symmetric monoidal category $\mathcal{M}$ defines a section $A : \Gamma \to \mathcal{M}^\otimes$ of the functor p by the rule $S \mapsto (A, ..., A) \in \mathcal{M}^\otimes(S) = \mathcal{M}^{S \setminus s}$. This defines a functor $i : Alg(\mathcal{M}) \to \mathsf{Sect}(\Gamma, \mathcal{M}^\otimes)$ from the category of $\mathcal{M}$-algebras to p-sections, which is full and faithful and its image is easy to characterise.

However, if $\mathcal{M} = \mathsf{DVect}_k$, commutative algebras may not necessarily coincide, even up to a quasi-isomorphism, with $\mathbb{E}_\infty$-algebras. Indeed, the section formalism still does not produce diagrams of shape (1) inside $\mathcal{M}$. To tackle this, we introduce the notion of a *derived section* [2, 3].

Generalising the situation, from $\mathsf{DVect}_k^\otimes \to \Gamma$ we pass to a *model opfibration* $p : \mathcal{E} \to \mathcal{C}$, which is a Grothendieck opfibration such that each fibre $\mathcal{E}(c)$ is a model category and the transition functors $f_! : \mathcal{E}(c) \to \mathcal{E}(c')$ preserve fibrations and trivial fibrations of the model structure on $\mathcal{E}(c)$. To it, we first associate the category of presections $\mathsf{PSect}(\mathcal{C}, \mathcal{E})$ as follows.

First, the family $\mathcal{E} \to \mathcal{C}$ is encoded by a covariant functor from $\mathcal{C}$, or a contravariant functor from $\mathcal{C}^{\mathsf{op}}$. Denote by $\mathcal{E}^\top \to \mathcal{C}^{\mathsf{op}}$ the associated Grothendieck fibration. By $\mathbb{C}$, we denote the simplicial replacement [8] of $\mathcal{C}$, which is the opposite of the simplex category of the nerve $N\mathcal{C}$. An object of $\mathbb{C}$ is thus a sequence $c_0 \to ... \to c_n$ of composable maps in $\mathcal{C}$. The fibration $\mathcal{E}^\top \to \mathcal{C}^{\mathsf{op}}$ induces a Grothendieck fibration $\mathbf{E} \to \mathbb{C}$ with $\mathbf{E}(c_0 \to ... \to c_n) \cong Sect(\mathbf{c}_{[n]}, \mathcal{E}^\top)$. The category $\mathsf{PSect}(\mathcal{C}, \mathcal{E})$ is then defined as the category of sections $\mathsf{Sect}(\mathbb{C}, \mathbf{E})$ of the fibration $\mathbf{E} \to \mathbb{C}$. To each object $\mathbf{c}_{[n]} = c_0 \to ... \to c_n$ of $\mathbb{C}$, a presection X associates the diagram

$$\alpha^* X(c_0) \longleftarrow X(c_0 \to ... \to c_n) \longrightarrow X(c_n), \tag{3}$$

where $\alpha^* : \mathbf{E}(c_0) \to \mathbf{E}(\mathbf{c}_{[n]})$ is the transition functor, and we define derived sections in the following way.

Definition 2. Let $\mathcal{E} \to \mathcal{C}$ be a model opfibration. The category of derived sections $\mathsf{DSect}(\mathcal{C}, \mathcal{E})$ is a full subcategory of $\mathsf{PSect}(\mathcal{C}, \mathcal{E})$ consisting of those X such that the left arrow in diagrams like (3) is a weak equivalence in $\mathcal{E}(c_n)$.

Thus in the case of $\mathcal{M}^\otimes \to \Gamma$, a derived section gives us diagrams of the shape $Z^{\otimes n} \overset{\sim}{\longleftarrow} Y(n) \longrightarrow Z$ with left arrow a weak equivalence.

Working with DSect: Reedy Model Structure. In order to work with the category $\mathsf{DSect}(\mathcal{C}, \mathcal{E})$ or even $\mathsf{PSect}(\mathcal{C}, \mathcal{E})$ homotopically it is necessary to have some structure. I prove the following result, sketched in [2] and fleshed out in [4]:

Theorem 3. *For a model opfibration $\mathcal{E} \to \mathcal{C}$, the category $\mathsf{PSect}(\mathcal{C}, \mathcal{E})$ carries a model structure with pointwise weak equivalences.*

The model structure provided by this theorem is very concrete and resembles a lot the ordinary Reedy model structure. Returning to model opfibrations, the localisation [8, 12] $\operatorname{Ho}\mathsf{PSect}(\mathcal{C}, \mathcal{E})$ of the presection category is thus well under control, and we denote by $\operatorname{Ho}\mathsf{DSect}(\mathcal{C}, \mathcal{E})$ the corresponding subcategory of derived sections.

Higher-categorical sections and derived sections. One can go further and ask the following question: what is the link between the derived sections $\mathsf{DSect}(\mathcal{C}, \mathcal{E})$ of a model opfibration $\mathcal{E} \to \mathcal{C}$, and the quasicategory [19] of sections $\mathsf{Sect}(\mathcal{C}, L\mathcal{E})$ of the associated higher-categorical opfibration $L\mathcal{E} \to \mathcal{C}$, obtained by localising $\mathcal{E}$? Our more recent work [4] provides the answer to this question.

It has been remarked in the preprint [16] that given a model bifibration $\mathcal{M} \to \mathcal{R}$ over a Reedy category $\mathcal{R}$, the associated model category of sections $\mathsf{Sect}(\mathcal{R}, \mathcal{E})$ should in fact be a model for the higher-categorical sections. We make this result concrete in [4]:

Theorem 4. *Let $\mathcal{M} \to \mathcal{R}$ be a model bifibration over a Reedy category $\mathcal{R}$, and let $L\mathcal{M} \to \mathcal{R}$ be the biCartesian fibration [20] obtained through the procedure of the infinity-localisation [15]. Then the induced ∞-functor $L\mathsf{Sect}(\mathcal{R}, \mathcal{M}) \to \mathsf{Sect}(\mathcal{R}, L\mathcal{M})$ is an equivalence of ∞-categories.*

We prove this theorem in a slightly greater generality. The original idea for Theorem 4 is due to [16]. The authors however lacked the tools necessary to prove the needed higher-categorical statements; we fill this gap using the modern machinery of [19]. In [5], we then prove

Theorem 5. *Let $\mathcal{E} \to \mathcal{C}$ be a model opfibration. Then the infinity-category $L\mathsf{DSect}(\mathcal{C}, \mathcal{E})$ is equivalent to $\mathsf{Sect}(\mathcal{C}, L\mathcal{E})$.*

This result allows for great flexibility when working with derived sections. To prove a statement about objects of $\mathsf{DSect}(\mathcal{C}, \mathcal{E})$, we can either treat them as higher-categorical sections, or as ordinary sections over the simplicial replacement $\mathbb{C}$. As a consequence, one can rather immediately deduce the existence of (homotopy co)limits in $\mathsf{Sect}(\mathcal{C}, L\mathcal{E})$ by using the fact that in the language of derived sections, the questions of (co)completeness are much easier to treat, see [5] for detail.

We conclude by presenting a collaboration that we would like to pursue once the COVID situation relaxes.

The generalised Deligne conjecture. This is a collaboration with Y. Harpaz. Let us rethink what approach may be adapted in order to better understand various aspects of higher algebra. First, one would really prefer to fix what it means to be an "operation-indexing" category. The formalism of [6] is very limiting (one discovers immediately [2] that many examples do not exactly fit the definitions of [6]). Recently however, a more satisfactory definition was proposed by Harpaz [14], bearing the name of a weak operad or algebraic pattern [9]. Without going into technicalities, the main idea is that a weak operad $\mathcal{O}$ is an infinity-category with a factorisation system, that allows to write what it means for an (infinity)-functor $\mathcal{O} \to \mathcal{M}$ to be Segal. An advantage of this notion is that it allows to consider at the same time both the infinity-operads of Lurie and the categories like $\Gamma, \Delta^{\mathsf{op}}$ or some associated constructions. We use this notion to work on the following.

Given an algebra A, its endomorphism operad $\operatorname{End}(A)^\otimes$ is naturally an operad over and under the associative operad, which is the fact that retains the information about the algebraic structure on A. From the observations of [6], one expects that given any (infinity)-operad $\mathcal{O}$, there is an associated weak operad that we call $\Delta_\mathcal{O}$. Its meaning is similar to the category of simplices $\mathbb{C}$ introduced above, but now it puts elements of $\mathcal{O}$ over the simplices; one could also say that it is a certain version of an S-construction applied to the operad $\mathcal{O}$. Using this language the full description of the structure of an enriched operad over $\mathcal{O}$ can be done in a way similar to [10], describing enriched operads over $\mathcal{O}$ as $\Delta_{\mathcal{O}}$-algebras. The claim is then as follows: if $\mathcal{O}$ is the associative operad, then $\mathrm{End}(A)^{\otimes}$ can be naturally realised as a $\Delta_{\mathcal{O}}$-algebra in DVect_k, and the localisation of the weak operad $\Delta_{\mathcal{O}}$ along unary operations gives the operad $\mathbb{E}_2$. This would be, in our opinion, a most natural proof of Deligne's conjecture, and we hope to also use it to verify that all other solutions to the conjecture produce the same answer. Along the way we hope to revisit the category of trees T as appearing in [2] and understand its links with the marked planar dendroidal category Ω_*. The greater goal however is to siginficantly generalise Deligne conjecture, replacing the associative operad by any reduced unital coherent infinity-operad.

The are many other things that one can do while working with the weak operads, such as studying the generalisation of twisted arrow categories, Segal categories over weak operads, and many more.

Work on artificial intelligence

Definition 6. A *(feed-forward) neural network* is a function $N : \mathbb{R}^n \to \mathbb{R}^k$ that can be written as a composition $N = \sigma \circ f$ where $f : \mathbb{R}^n \to \mathbb{R}^k$ is an affine map and $\sigma : \mathbb{R}^k \to \mathbb{R}^k$, called an *activation function*, is obtained by coordinate-wise application of a fixed function $\tilde{\sigma} : \mathbb{R} \to \mathbb{R}$. A composition of a number of neural networks is called a *deep neural network.*

In practice, $\tilde{\sigma}$ is often chosen to be $x \mapsto max(0, x)$ or $x \mapsto tanh(x)$. The affine functions are often of specific form (a fairly well-known class is convolutional neural networks).

Thanks to the recent boom, deep neural networks have become extremely popular in industrial practice, computer science and applied mathematics. The practical reason for this is that deep neural network are able to approximate various functions, or even demonstrate functional dependencies, via a machine-implementable procedure of gradient descent, that allows to pick the parameters for the affine functions in the definition above. The algorithm minimises a certain functional, called the *loss function* [30], defined on the space of all deep neural networks of specified type. Searching for functional dependencies and performing numerical experiments may also be useful in pure mathematics, where machine learning is gradually gaining its territory [33].

Neural semigroups. One wide class of mathematical structures that could be treated as data for neural networks is produced by studying various algebraic structures on finite sets. A particular type of examples is given by *semigroups*: sets with associative binary operation. Ordinary groups and monoids have underlying semigroups, and one can also produce examples from categories. Given a category $\mathcal{C}$, take $S := \mathrm{Mor}\,\mathcal{C} \cup \{0\}$ and define $0 \cdot 0 := 0$ and $f \cdot g := f \circ g$ if composition is possible and 0 otherwise. Unlike groups, no classification for finite semigroups is known: at best, there are lists of some small cardinality semigroups [28], with 53 trillion classes of isomorphic and anti-isomorphic semigroups of cardinality 9. This number suggests that the total classification is out of reach, with machine tools like [32, 31] often taking hours to verify existence of a finite semigroup satisfying some set of equations.

A deep neural network operating on (a suitable embedding into $\mathbb{R}^N$ of) the set Semi_n of all semigroups of n elements could potentially learn to solve minimisation problems operating on a very small set of training examples. What kind of functions could one potentially construct using deep neural networks defined on Semi_n? What information about semigroups could be inferred from those functions? Can one use various clusterisation results known in deep learning to attempts at describing certain classes of "large" finite semigroups?

Completing partial tables. Our work [26] started as an attempt to teach a neural network to (re)construct semigroup structures. To pose the problem, let us first explain how we present semigroups as data. Given a set S, one can consider the free vector space $\mathbb{R}[S]$; a semigroup structure on S equips $\mathbb{R}[S]$ with a structure of a non-unital algebra over $\mathbb{R}$. If we write $S = \{e_i\}_{i=1}^n$, then in $\mathbb{R}[S]$ one has $e_i \cdot e_j = \sum_k M_{ijk} e_k$. The structure constants M_{ijk} are coefficients of the multiplication tensor in the canonical basis on $\mathbb{R}[S]^{*\otimes 2} \otimes \mathbb{R}[S]$, and indeed any algebra structure on a finite vector space amounts to specifying a basis and a set of such coefficients. So, as a representation of our semigroup structure we simply take $(M_{ijk}) \in \mathbb{R}^{n^3}$. This admits an immediate generalisation to any finite number of k-ary operations on S.

The coefficients (M_{ijk}) have a couple of properties. They are normalised, in the sense that for given i, j $M_{ijk} = 0$ except for exactly one k, and they also satisfy the associativity condition: $\sum_m (M_{ijm} M_{mkl} - M_{iml} M_{jkm}) = 0$. These coefficients are commonly depicted using multiplication tables. The problem of classifying semigroups is then tied to the question of finding (M_{ijk}) with such properties.

One way to generate a semigroup consists of specifying on $S = \{e_i\}$ a certain set of equations of the form $e_i \cdot e_j = e_k$. If we in addition put $M_{ijk} = 1/n$ for those (i, j) that do not have an equation specified, we get a tensor in $\mathbb{R}^{n^3}$ that can be viewed as a partially filled multiplication table, with uniform probability distributions added in place of unknown multiplications. A neural network generation procedure then consists of writing a deep neural network function $G : \mathbb{R}^{n^3} \to \mathbb{R}^{n^3}$ that, given such a partial set of coefficients, produces $G(M_{ijk})$ that satisfy the semigroup coefficients properties and such that $G(M_{ijk}) = M_{ijk}$ for the i, j with specified multiplications. It should be noted that not any set of equations in S can be extended to a semigroup structure. However, there are indications (found by the second author of [26]) that if the number of equations is less than $n = |S|$, then a semigroup structure with such equations exists with probability almost one.

The preprint of [26] describes such a neural network for the test case of $n = 5$. The classification is known here, so we simply generate partial multiplication tables by taking full tables and masking certain cells. A deep neural network $G : \mathbb{R}^{125} \to \mathbb{R}^{125}$ accepting such partial tables as inputs is then forced to not change known cells and is penalised if the output is not associative. This is done by the means of interpreting the associativity equation in probability terms: our loss function is written as the Kullback-Leibler divergence between $\sum_m G(M_{ijm})G(M_{mkl})$ and $\sum_m G(M_{iml})G(M_{jkm})$. Training a network on only 10 percent of all (anti)-isomorphism classes of semigroups while masking around one-half of all cells produces a generator G that succeeds to construct a semigroup from a half-filled table with rate of success 80 percent.

There are many questions arising from our experiment. It can be generalised to other cardinalities present in the classification list [28], and even in case of higher cardinalities, one can use [32, 31] to create some examples of semigroups that one could use as data. The generator G for $n = 5$ reconstructs the original table in around 14 percent of cases, does this number provide any information about the number of semigroups? Can one use G, applied repeatedly to random inputs, in order to construct any semigroup (some small-order verification of ours has been so far positive)? Can one greatly restrict the training set, using for example only nilpotent semigroups, to get to all semigroups via G? How can one mathematically formalise these experiments? What other structures could be studied using the similar method, and to what end? We can only continue our work to find out.

Homological invariants of neural networks. We conclude by mentioning another recently started project in collaboration with B. Shminke (3IA, Nice) and M. Prasma (private sector research, Berlin). Another interaction between deep learning and mathematics consists of studying neural networks as mathematical objects. The existing approaches usually involve a lot of analysis [33] or statistics, yet more recently there have been attempts at applying the methods of homology and category theory in order to understand, in particular, the procedure of supervised learning.

Definition 7. Given a neural network $N : \mathbb{R}^n \to \mathbb{R}^k$ and $x \in \mathbb{R}^n$, its *m-th activation value at x*, $a_m(x)$, is defined as the m-th of $N(x)$. For a deep neural network $D = N_l \circ ... \circ N_1 : \mathbb{R}^n \to \mathbb{R}^k$, one defines its *$(i, m)$-th activation value at x*, $a_{(m,i)}(x)$, to be the m-th coordinate of $N_i \circ ... \circ N_1(x)$. The assignments $x \mapsto a_i(x)$ and $x \mapsto a_{(m,i)}(x)$ will be called *activation functions*.

One usually thinks of activation functions as of "neurons", their values corresponding to the reactions of the neuron to the data. In effect, one can see even in simplest examples [34] that neurons appear to become responsible for certain aspects of the data set as the network trains. One can attempt to understand leairning by studying the interaction of activation functions.

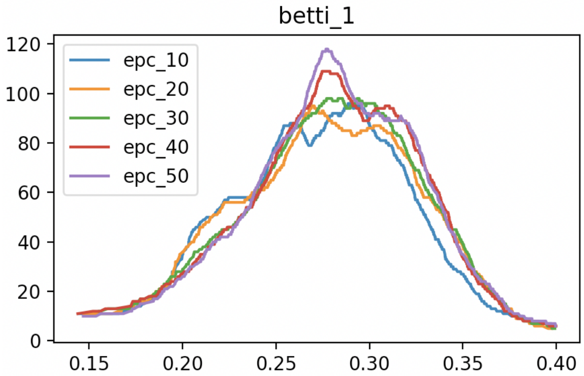
In [27] one proposes then the following construction. Fix a deep neural network D. Given a finite ordered subset $\mathbf{x} = \{x_j\}_{j=1}^J \subset \mathbb{R}^n$ that usually comes from a training set, we can form vectors $a_{m,i}(\mathbf{x}) = (a_{m,i}(x_1), ..., a_{m,i}(x_J)) \in \mathbb{R}^J$ for all possible activation functions coming from D. Given two such *activation vectors* $\mathbf{p} = a_{(m,i)}(\mathbf{x})$ and $\mathbf{p}' = a_{(m',i')}(\mathbf{x})$, define their correlation distance to be

$$d(\mathbf{p}, \mathbf{p}') := 1 - corr(\mathbf{p}, \mathbf{p}')$$

where $corr(\mathbf{p}, \mathbf{p}')$ is the standard Pearson correlation coefficient (that is set to zero if the standard deviation for either $\mathbf{p}$ or $\mathbf{p}'$ is zero). This turns the finite set $A(D, \mathbf{x}) = \{\mathbf{p} = a_{(m,i)}(\mathbf{x})\}_{m,i}$ of all possible activation vectors into a metric space.

Definition 8. Let $D : \mathbb{R}^n \to \mathbb{R}^k$ be a deep neural network and $\mathbf{x} = \{x_j\}_{j=1}^J$ be an ordered set of vectors in $\mathbb{R}^n$. Then the *q-th persistent homology group* $H_q(D, \mathbf{x}, \varepsilon)$ of persistence parameter ε is defined as the q-th homology group of the Vietoris-Rips complex of the finite metric space $(A(D, \mathbf{x}), d)$ with distance at most ε.

The project [27] provides code that implements the computation of these homology groups during the process of learning for a few examples of convolutional [30] neural networks usually used in computer vision. For instance, here is its output of $\dim H_1(D, \mathbf{x}, \varepsilon)$ for D being the LeNet neural network trained over MNIST dataset with different epochs meaning different iterations of passes of gradient descent algorithm:



The authors of [27] suggest that the formation of the peak as above in various homology dimensions corresponds to finding a suitable training minimum, and that further phenomena (such as overfitting to the training data set) can also be witnessed in the behaviour of homology dimension curves.

The idea of aggregating neurons is not new and was already studied using the techniques of mean field theory. The homological perspective is however fresh, and we would argue that it could become more transparent conceptually, if developed. Of course, currently many questions remain unanswered, beginning with the basic mathematical underpinnings of the construction. The relation of the defined homology groups to other homological invariants in mathematics is, for now, completely out of scope. The homology dimension curves shown above seem to change continiously with training epochs, but to formalise it, one would have to likely develop a mixture of analytical and category-theoretical techniques similar to [29]. It is not obvious how to formulate the relation between minimisation and emergence of peaks. It is equally unclear why one should choose $1 - corr(\mathbf{p}, \mathbf{p}')$ as metric and not say its square root (that also defines a metric) or a Wasserstein distance. In attacking all these (and many more other) questions, wea are hoping to see if it is possible to partially reverse these considerations and develop a version of homological learning, both in formal and in practical terms.

References

Homotopy theory: **[1]** John F. Adams, *Infinite Loop Spaces*, Princeton University Press, 1978 **[2]** Eduard Balzin, *Grothendieck Fibrations and Homotopical Algebra*, PhD Thesis available at `https://tel.archives-ouvertes.fr/tel-01369024v1` **[3]** Edouard Balzin, *Derived sections of Grothendieck fibrations and the problems of homotopical algebra*, `http://arxiv.org/abs/1410.3387`, Applied Categorical Structures (2017), DOI: 10.1007/s10485-017-9483-1 **[4]** Edouard Balzin, *Reedy Model Structures in Families*, preprint `https://arxiv.org/abs/1803.00681` **[5]** Edouard Balzin, *The formalism of Segal sections*, preprint `https://arxiv.org/abs/1811.09601` **[6]** Clark Barwick, *From operator categories to topological operads*, preprint `http://arxiv.org/abs/1302.5756` **[7]** Alexander Beilinson, Vladimir Drinfeld, *Chiral Algebras*, AMS 2004 **[8]** Aldridge K. Bousfield, Daniel M. Kan, *Homotopy limits, completions and localizations,* Lecture Notes in Mathematics, Vol. 304. Springer-Verlag, Berlin, 1972 **[9]** Hongyi Chu, Rune Haugseng, *Homotopy-coherent algebra via Segal conditions*, preprint `https://arxiv.org/abs/1907.03977v1` **[10]** Hongyi Chu, Rune Haugseng, *Enriched ∞-operads*, preprint `https://arxiv.org/abs/1707.08049v2` **[11]** Kevin Costello, *Topological conformal field theories and Calabi-Yau categories,* Advances in Mathematics Volume 210, Issue 1, 20 March 2007, pp 165-214 **[12]** William G. Dwyer, Philip S. Hirschhorn, Daniel M. Kan, and Jeffrey H. Smith, *Homotopy Limit Functors on Model Categories and Homotopical Categories*, AMS 2004 **[13]** Alexander Grothendieck, Michèle Raynaud et al., *Revêtements étales et groupe fondamental (SGA I)*, Lecture Notes in Mathematics 224, Springer 1971 **[14]** Yonatan Harpaz, *Little cube algebras and factorization homology*, lecture notes `https://www.math.univ-paris13.fr/~harpaz/lecture_notes.pdf` **[15]** Vladimir Hinich, *Dwyer-Kan localization revisited*, Homology, Homotopy and Applications, Volume 18 (2016), Number 1 **[16]** André Hirschowitz, Carlos Simpson, *Descente pour les n-champs (Descent for n-stacks)*, preprint `http://arxiv.org/abs/math/9807049` **[17]** Maxim Kontsevich, Yan Soibelman, *Notes on A-infinity algebras, A-infinity categories and non-commutative geometry. I,* Homological Mirror Symmetry, Volume 757 of the series Lecture Notes in Physics pp 1-67, 20 August 2008, `http://arxiv.org/abs/math/0606241` **[18]** Tom Leinster, *Higher Operads, Higher Categories*, London Mathematical Society Lecture Note Series 298, Cambridge University Press (2004), ISBN 0-521-53215-9 **[19]** Jacob Lurie, *Higher Topos Theory*, book, Annals of mathematics studies ; no. 170 **[20]** Jacob Lurie, *Higher Algebra*, online book, available at `http://www.math.harvard.edu/~lurie/` **[21]** J. Peter May, *The geometry of iterated loop spaces,* Springer-Verlag, Berlin, 1972, Lectures Notes in Mathematics, Vol. 271 **[22]** James E. McClure and Jeffrey H. Smith, *A solution of Deligne's Hochschild cohomology conjecture,* Recent Progress in Homotopy Theory (Baltimore, MD, 2000), volume 293 of Contemp. Math., pp 153-193. Amer. Math. Soc., Providence, RI, 2002. **[23]** Daniel Quillen, *Higher K-theory: I,* Lecture Notes in Mathematics Volume 341, 1973, pp 85-147 **[24]** Graeme Segal, *Categories and Cohomology Theories*, Topology Vol. 13 pp. 293-312, Pergamon Press 1974 **[25]** Dmitry E. Tamarkin, *Another proof of M. Kontsevich formality theorem,* `http://arxiv.org/abs/math/9803025`

Artificial intelligence: **[26]** Edouard Balzin, Boris Shminke, *A neural network for semigroups*, preprint and code `https://github.com/inpefess/neural-semigroups/` **[27]** Ciprian Cornelanu *DNN-topology*, `https://github.com/cipriancorneanu/dnn-topology`, C. A. Corneanu, M. Madadi, S. Escalera, and A. M. Martinez, *What does it mean to learn in deep networks? and, how does one detect adversarial attacks?* Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 4757–4766, 2019 **[28]** Andreas Distler, James Mitchell, *GAP Package Smallsemi*, `https://www.gap-system.org/Packages/smallsemi.html` **[29]** Brendan Fong, David I. Spivak and Rémy Tuyéras, *Backprop as Functor: A compositional perspective on supervised learning*, LICS 2019, `https://arxiv.org/abs/1711.10455v3` **[30]** Ian Goodfellow, Yoshua Bengio and Aaron Courville, *Deep Learning Book*, Mit Press 2016, `https://www.deeplearningbook.org` **[31]** *Isabelle generic proof assistant*, `https://isabelle.in.tum.de` **[32]** *Prover9 and Mace4,* `https://www.cs.unm.edu/~mccune/prover9/` **[33]** Sanya conference on Algebraic Geometry and Machine Learning, `https://www.youtube.com/channel/UCNakPYZGk8PPsORGB9GbbNA`, `https://sites.google.com/view/agmlsanya/talks` **[34]** *TensorFlow playground*, `http://playground.tensorflow.org/`