

УДК 004.8

## БАРКОДЫ КАК РЕЗЮМЕ ТОПОЛОГИИ ФУНКЦИЙ ПОТЕРЬ

© 2023 г. С. А. Баранников<sup>1,3,\*</sup>, А. А. Коротин<sup>1,2</sup>, Д. А. Оганесян<sup>1</sup>,  
Д. И. Емцев<sup>1,4</sup>, Е. В. Бурнаев<sup>1,2</sup>

Представлено академиком РАН А. Л. Семеновым

Поступило 02.09.2023 г.

После доработки 08.09.2023 г.

Принято к публикации 18.10.2023 г.

Статья посвящена изучению поверхностей потерь нейронных сетей с помощью методов топологического анализа данных. Мы применяем бар-коды комплексов Морса для исследования топологии поверхностей потерь. Описан алгоритм для расчета бар-кодов локальных минимумов функции потерь. Мы провели эксперименты по расчету бар-кодов локальных минимумов для бенчмарк функций и для поверхностей потерь небольших нейронных сетей. Наши эксперименты подтверждают два наших основных наблюдения для поверхностей потерь нейронных сетей. Во-первых, бар-коды локальных минимумов расположены в небольшой нижней части диапазона значений функции потерь нейронных сетей. Во-вторых, увеличение глубины и ширины нейронной сети уменьшает бар-коды локальных минимумов. Это дает естественные следствия для обучения нейронной сети и для ее обобщающих свойств.

*Ключевые слова:* поверхность потерь, персистентные гомологии, персистентные бар-коды, теория Морса, нейронные сети

**DOI:** 10.31857/S268695432360177X, **EDN:** ANZNRU

### 1. ВВЕДЕНИЕ

Поиск минимумов функции потерь является основной стратегией для большинства алгоритмов машинного обучения. График функции потерь, который часто называется **поверхностью потерь**, обычно имеет сложную структуру [1–3]: невыпуклость, множество локальных минимумов, седловые точки, плоские области. Эти препятствия затрудняют исследование поверхности потерь.

Поиск минимумов современных нейронных сетей в основном осуществляется с помощью алгоритмов, основанных на градиентном спуске. Каким образом эти алгоритмы могут достигать почти нулевого значения функции потерь, несмотря на невыпуклость этой функции, остается недостаточно понятным.

Глобальные топологические характеристики множества траекторий градиентного потока опи-

сываются комплексом Морса через декомпозицию пространства параметров на ячейки, где поток однороден [4–6]. Бар-коды комплекса Морса составляют основное резюме топологии потока градиентного векторного поля [7–9]. Бар-коды дают декомпозицию изменения топологии подуровневых множеств функции потерь на элементарные пары метаморфоз “рождение”–“исчезновение”.

Мы утверждаем, см. раздел 5, что топологическое препятствие, образованное данным локальным минимумом для алгоритмов обучения на основе градиентных методов, можно количественно оценить длиной сегмента соответствующего этому минимуму в бар-коде. Для алгоритмов обучения на основе градиентных методов это количество измеряет *обязательный штраф за переход от данного локального минимума к точке с меньшим значением функции потерь*.

Расчет бар-кодов для различных функций составляет суть топологического анализа данных. В настоящее время доступные программные пакеты для расчета бар-кодов функций, также называемых “подуровневой персистентностью”, включают GUDHI, Dionysus, PHAT. Они основаны на общем алгоритме, требующем построения симплициального комплекса и имеющем худшую временную сложность  $O(N^3)$  по числу точек для

<sup>1</sup>Сколковский институт науки и технологий, Москва, Россия

<sup>2</sup>Научно-исследовательский институт искусственного интеллекта AIRI, Москва, Россия

<sup>3</sup>CNRS, IMJ, Paris Cité University, Париж, Франция

<sup>4</sup>ETH Цюрих, Швейцария

\*E-mail: s.barannikov@skoltech.ru

вычисления бар-кода наименьшей степени. Эти пакеты в настоящее время могут обрабатывать расчеты бар-кодов для функций, определенных на сетке, до  $10^6$  точек и в размерностях до шести. Таким образом, все существующие пакеты сталкиваются с проблемами масштабируемости.

Мы описываем алгоритм для вычисления бар-кодов наименьшей степени для функций в произвольных размерностях. В отличие от упомянутых методов на основе сетки, наш алгоритм работает с функциями, определенными на произвольно выбранных **облаках точек**. Известно, что методы на основе облаков точек работают лучше, чем методы на основе сетки, в задачах оптимизации [10]. Для вычисления бар-кодов наименьшей степени мы используем тот факт, что их определение можно переформулировать в геометрических терминах, см. определение 1 в разделе 2. Большинство в настоящее время доступных программных пакетов основаны на более алгебраическом подходе, как в определении 2 из раздела 2. Основная часть нашего алгоритма имеет худшую временную сложность  $O(N \log N)$ , наш алгоритм был протестирован в размерностях до 15 и с числом точек до  $10^9$ .

Предложенная методология описывает свойства поверхности потерь нейронных сетей через топологические признаки локальных минимумов. Мы хотели бы подчеркнуть, что значение функции потерь в минимуме составляет только половину его топологической характеристики из бар-кода. Другая половина может быть описана как значение функции потерь в 1-седловой точке, которая естественным образом связана с каждым локальным минимумом, см. раздел 2. 1-седловая точка  $q$ , связанная с минимумом  $p$ , является точкой, в которой связная компонента подуровневого множества  $\Theta_{f \leq c} = \{\theta \in \Theta | f(\theta) \leq c\}$ , содержащая  $p$ , объединяется с другой связной компонентой подуровневого множества, содержащей *более низкий* минимум. Это соответствие между локальными минимумами и 1-седловыми точками, уничтожающими связные компоненты  $\Theta_{f \leq c}$ , является взаимно однозначным.

Сегмент  $[f(p), f(q)]$ , где  $q$  — это 1-седловая точка, связанная с  $p$ , является инвариантом, который в бар-коде связывается с минимумом  $p$ . Разница  $f(q) - f(p)$  является топологическим инвариантом минимума, количественно оценивающим его как препятствие для алгоритмов обучения на основе градиентных методов. Множество всех таких сегментов для всех минимумов — это бар-код наименьшей размерности функции  $f$ .

Основной вклад данной статьи заключаются в следующем:

**1) Применение бар-кодов минимумов и соответствия между минимумами и 1-седловыми точками**

**для исследования поверхностей потерь.** Для каждого локального минимума  $p$  канонически определена 1-седловая точка  $q$  (см. Раздел 2). Множество всех сегментов  $[f(p), f(q)]$ , где  $p$  — локальный минимум, а  $q$  — соответствующая 1-седловая точка  $f$ , является устойчивым топологическим инвариантом функции потерь. Это множество инвариантно, в частности, относительно действия гомеоморфизмов на  $\Theta$ . Это множество сегментов является частью полного бар-кода. Полный бар-код дает полное описание топологии функции потерь и глобальной структуры ее градиентного потока.

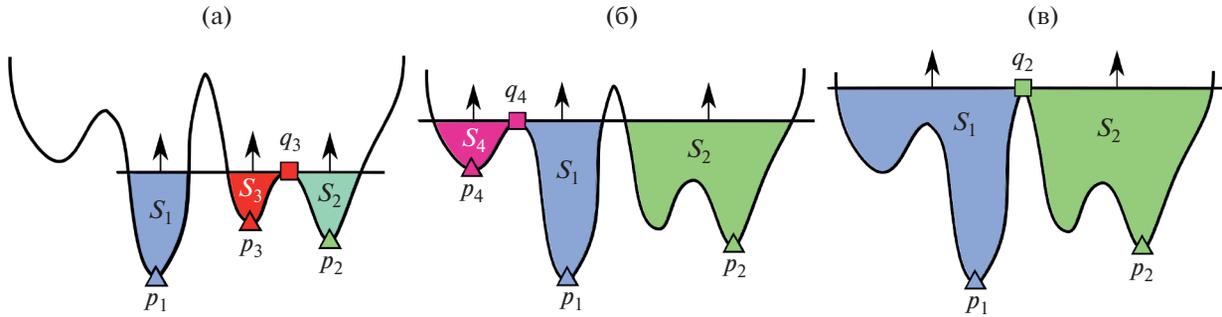
**2) Алгоритм для расчета бар-кодов минимумов.** Мы описываем и анализируем алгоритм для расчета бар-кодов минимумов. Алгоритм принимает на вход случайно или специальным образом выбранный набор точек и значения функции потерь на этом наборе. Затем алгоритм использует процедуру HNSW для расчета графа соседей. Следующим шагом является вычисление бар-кода минимумов для функции, определенной на графе. Локальные минимумы порождают кластеры точек в подуровневых множествах. Затем алгоритм работает, рассматривая соседей каждой выбранной точки с более низкими значениями функции и решая, принадлежит ли эта точка существующему кластеру, порождает новый кластер (минимум) или объединяет два или более кластера (1-седловая точка). Вторая часть алгоритма, работающего с функцией на графе, аналогична частному 0-мерному случаю общего алгоритма, описанного в [7].

**3) Эксперименты, подтверждающие наблюдения о поведении бар-кодов функций потерь нейронных сетей.** Мы вычисляем бар-коды минимумов для малых полносвязных нейронных сетей имеющих до трех скрытых слоев и подтверждаем, что все сегменты бар-кода минимумов принадлежат небольшой нижней части общего диапазона значений функции потерь и что с увеличением ширины и глубины нейронной сети бар-коды минимумов опускаются ниже.

Наш код доступен в репозитории <https://github.com/grag90/BarcodeCalc>.

Полезность нашего подхода и алгоритмов не ограничивается задачами оптимизации. Наш алгоритм позволяет быстро вычислять персисентные бар-коды многих функций, которые до сих пор были недоступны. Эти подуровневые персисентные бар-коды успешно применялись в различных дисциплинах: например в когнитивных науках [11], космологии [12], см. также ссылки в [13].

Наша концепция также имеет применение в химии и науке о материалах, где 1-седловые точки на потенциальных энергетических ландшафтах соответствуют переходным состояниям, а мини-



**Рис. 1.** Объединение связных компонент подуровневых множеств в 1-седловых точках. (а) “Исчезновение” связной компоненты  $S_3$ , связная компонента  $S_3$  подуровневого множества объединяется со связной компонентой  $S_2$  в 1-седловой точке  $q_3$ , 1-седловая точка  $q_3$  связана с минимумом  $p_3$ . (б) “Исчезновение” связной компоненты  $S_4$ , связная компонента  $S_4$  подуровневого множества объединяется со связной компонентой  $S_1$  в 1-седловой точке  $q_4$ , 1-седловая точка  $q_4$  связана с минимумом  $p_4$ . (в) “Исчезновение” связной компоненты  $S_2$ , связная компонента  $S_2$  подуровневого множества объединяется со связной компонентой  $S_1$  в 1-седловой точке  $q_2$ , 1-седловая точка  $q_2$  связана с минимумом  $p_2$ . Обратите внимание, что 1-седловая точка  $q_2$  связана с минимумом  $p_2$ , который отделен другим минимумом от зеленого седла  $q_2$ .

мумы являются стабильными состояниями, соответствующими различным материалам или различным сворачиваниям белка, см., например, [14, 15].

Статья содержит следующие разделы. Мы начинаем с двух определений бар-кодов локальных минимумов в разделе 2. Наш алгоритм для расчета бар-кодов описан в разделе 3. В разделе 4 мы применяем наш алгоритм для расчета бар-кодов бенчмарк функций. Мы доказываем сходимость алгоритма и демонстрируем ее эмпирически в подразделе 4. В разделе 5 мы вычисляем бар-коды функций потерь малых нейронных сетей и описываем наши основные наблюдения.

## 2. ОПИСАНИЕ ТОПОЛОГИИ ПОВЕРХНОСТЕЙ ПОТЕРЬ ЧЕРЕЗ БАР-КОДЫ

Бар-коды описывают топологические признаки функций представляя изменение топологии подуровневых множеств функции как сумму пар элементарных метаморфоз “рождение”–“исчезновение”. Мы предлагаем применить эти инварианты как инструмент для изучения топологии поверхностей потерь.

В этом разделе мы описываем два определения бар-кодов минимумов для кусочно-гладких непрерывных функций. В этой работе мы сосредотачиваемся на части бар-кодов, описывающих явления “рождения”–“исчезновения” связных компонент подуровневых множеств функции потерь. Подход из этого раздела аналогично работает в контексте “почти минимумов”, т.е. для критических точек (многообразий) малых ненулевых индексов. Такие точки часто также могут являться конечными точками алгоритмов оптимизации

в высокоразмерных пространствах параметров глубоких нейронных сетей [2].

**Первое определение: объединение со связной компонентой более низкого минимума**

Пусть  $f$  является кусочно-гладкой непрерывной функцией. Значения параметра  $c$ , при которых топология подуровневых множеств

$$\Theta_{f \leq c} = \{\theta \in \Theta \mid f(\theta) \leq c\}$$

меняется, являются критическими значениями  $f$ .

Пусть  $p$  является одним из минимумов  $f$ . Когда  $c$  увеличивается от  $f(p) - \epsilon$  до  $f(p) + \epsilon$ , рождается новая связная компонента множества  $\Theta_{f \leq c}$ . Для иллюстрации этого процесса мы приводим пример на рис. 1, где связные компоненты  $S_1, S_2, S_3, S_4$  подуровневого множества рождаются в минимумах  $p_1, p_2, p_3, p_4$  соответственно.

Связные компоненты подуровневых множеств объединяются в 1-седловых критических точках.

Точка  $q$  является 1-седловой критической точкой, если пересечение множества  $\Theta_{f < f(q)}$  с любой малой окрестностью  $q$  имеет более одной связной компоненты. Без потери общности в этом разделе мы можем предположить, что 1-седловые точки  $f$  являются особенностями общего положения, так что эти пересечения имеют не более двух связных компонент. В противном случае можно всегда добавить небольшое возмущение к  $f$  или рассмотреть 1-седловые точки с кратностью, заданной числом этих компонент минус один.

Пусть  $p$  – это минимум, не являющийся глобальным. С увеличением  $c$ , связная компонент  $\Theta_{f \leq c}$ , рожденная в  $p$ , объединяется с другой связной компонентой. Затем эта объединенная связ-

ная компонента может снова объединиться с другой. После каждого объединения минимум ограничения  $f$  на объединенной связной компоненте является наименьшим из двух минимумов ограничения  $f$  на каждую из двух связных компонент перед объединением. Другими словами, связная компонента с более низким минимумом “поглощает” в точке объединения связную компоненту с более высоким минимумом, см. рис. 1. Пусть  $q$  является точкой объединения, где связная компонента с минимумом  $p$  поглощается связной компонентой, чей минимум ниже. Заметим, что пересечение множества  $\Theta_{f < f(q)}$  с любой малой окрестностью  $q$  имеет по крайней мере две связных компоненты.

**Определение 1.** Точка объединения  $q$ , где связная компонента с минимумом  $p$  поглощается связной компонентой с более низким минимумом, является 1-седлом, естественно связанным с минимумом  $p$ . Сегмент  $[f(p), f(q)]$  естественно ассоциируется в бар-коде с минимумом  $p$ .

Обратите внимание, что две связные компоненты пересечения малой окрестности такой  $q$  с  $\Theta_{f < f(q)}$  принадлежат двум разным связным компонентам всего множества  $\Theta_{f < f(q)}$ . 1-седловые этого типа называются “плюс” или типом “исчезновения”. Заметим, что

**Утверждение 1.** *Описанное соответствие между локальными минимумами и 1-седлами такого типа является взаимно однозначным.*

**Доказательство.** Соответствие в обратном направлении можно описать следующим образом. Пусть  $q$  является 1-седловой точкой такого типа, что две ветви множества  $\Theta_{f < f(q)}$  рядом с  $q$  не связаны во всем множестве  $\Theta_{f < f(q)}$ . Одна из связных компонент подуровневого множества  $\Theta_{f \leq c}$  разделяется на две, когда  $c$  уменьшается от  $f(q) + \epsilon$  до  $f(q) - \epsilon$ . Пусть  $p_1$  и  $p_2$  являются двумя минимумами ограничения  $f$  на каждую из этих двух связных компонент. Пусть  $p_1 > p_2$  является наибольшим из двух минимумов. 1-седло  $q$  связано по определению 1 с локальным минимумом  $p_1$ , так как две связных компоненты объединяются в  $q$ . Заметим что

$$p_1 \notin \left\{ \arg \min_{\substack{x \in \text{connected component} \\ \text{of } \Theta_{f \leq f(q) + \epsilon}}} f(x) \right\}$$

т.е.  $p_1$  не появляется среди минимумов  $f$  на одной из связных компонент  $\Theta_{f \leq f(q) + \epsilon}$ . Другими словами, минимум  $p_1$ , который соответствует  $q$  по определению 1, является новым минимумом, появляющимся в множестве минимумов  $f$  на связных компонентах  $\Theta_{f \leq c}$ , когда  $c$  уменьшается от  $f(q) + \epsilon$  до  $f(q) - \epsilon$ .  $\square$

Для применения в разделе 6 важно, чтобы 1-седла, связанные с минимумами, описывались следующим образом.

**Утверждение 2.** *Рассмотрим различные пути  $\gamma$ , начинающиеся от локального минимума  $p$  и идущие к более низкому минимуму. Пусть  $m_\gamma \in \Theta$  является максимумом ограничения  $f$  на таком пути  $\gamma$ . Тогда 1-седло  $q$ , соответствующее локальному минимуму  $p$  в бар-коде, является минимумом по множеству всех таких путей  $\gamma$  максимумов  $m_\gamma$ :*

$$q = \arg \left[ \min_{\substack{\gamma: [0,1] \rightarrow \Theta \\ \gamma(0)=p, f(\gamma(1)) < f(p)}} \max_t f(\gamma(t)) \right]. \quad (1)$$

**Доказательство.** Пусть связная компонента локального минимума  $p$  объединяется со связной компонентой более низкого минимума в точке  $q$ . Тогда существует путь в подуровневом множестве  $\Theta_{f \leq f(q)}$ , проходящий через  $q$  и соединяющий  $p$  с более низким минимумом. Точка  $q$  является максимумом ограничения  $f$  на этом пути. И пути в подуровневых множествах  $\Theta_{f \leq c}$ ,  $c < f(q)$ , который соединял бы  $p$  с более низким минимумом, нет. Следовательно, не существует пути с  $f(m_\gamma) < f(q)$  и  $q$  удовлетворяет уравнению (1).  $\square$

**Второе определение: инварианты фильтрованных комплексов**

Мы напоминаем здесь определение для полного бар-кода [7]. Хотя наш алгоритм из раздела 3 основан на определении 1, нам нужно определение 2 ниже, чтобы вставить вещи в общую рамку, в которой эти инварианты составляют полное описание топологии функции потерь и глобального поведения градиентного потока.

Семейства траекторий градиентного потока, исходящие из одной и той же особой точки, разлагают область  $\Theta$  на ячейки, на которых градиентный поток ведет себя равномерно [6].

Цепной комплекс является алгебраическим аналогом интуитивной идеи, представляющей сложные геометрические объекты в виде разложения на простые части. Он преобразует такое разложение в коллекцию векторных пространств и линейных отображений.

Напомним, что цепной комплекс  $(C_*, \partial_*)$  представляет собой последовательность конечномерных векторных пространств  $C_j$  (пространства “ $j$ -цепей”) и линейных операторов (“дифференциалов”)

$$\rightarrow C_{j+1} \xrightarrow{\partial_{j+1}} C_j \xrightarrow{\partial_j} C_{j-1} \rightarrow \dots \rightarrow C_0,$$

которые удовлетворяют  $\partial_j \circ \partial_{j+1} = 0$ .  $j$ -я гомология цепного комплекса  $(C_*, \partial_*)$  является частным векторных пространств  $H_j = \ker(\partial_j) / \text{im}(\partial_{j+1})$ .

Разложение сложного геометрического объекта на простые части делается часто в определенном последовательном порядке, в этом случае алгебраический аналог такого разложения представляет собой  $\mathbb{R}$ -фильтрованный цепной комплекс. Напомним что подкомплекс  $(C_*^*, \partial_*^*) \subseteq (C_*, \partial_*)$  представляет собой последовательность подпространств  $C_j^* \subseteq C_j$ , вместе с согласованными дифференциалами  $\partial_j^* = \partial_j|_{C_j^*}$ . Цепной комплекс  $C_*$  называется  $\mathbb{R}$ -фильтрованным, если в  $C_*$  имеется возрастающая последовательность подкомплексов ( $\mathbb{R}$ -фильтрация):  $F_{s_1} C_* \subset F_{s_2} C_* \subset \dots \subset F_{s_{\max}} C_* = C_*$ , индексированная конечным множеством вещественных чисел  $s_1 < s_2 < \dots < s_{\max}$ .

**Теорема 3** ([7], Секция 2). *Любой  $\mathbb{R}$ -фильтрованный цепной комплекс  $C_*$  может быть приведен к канонической форме, т.е. к канонически определенной прямой сумме  $\mathbb{R}$ -фильтрованных комплексов двух типов: одномерные комплексы с тривиальным дифференциальном  $\partial_j(e_i) = 0$  и двумерные комплексы с тривиальными гомологиями  $\partial_j(e_{i_2}) = e_{i_1}$ , с помощью линейного преобразования, сохраняющего  $\mathbb{R}$ -фильтрацию. Получившаяся прямая сумма таких простых  $\mathbb{R}$ -фильтрованных комплексов единственна.*

**Определение 2.** Полный бар-код представляет собой визуализацию разложения  $\mathbb{R}$ -фильтрованного комплекса в соответствии с теоремой 3. Каждый фильтрованный 2-мерный комплекс с тривиальными гомологиями  $\partial_j(e_{i_2}) = e_{i_1}$ ,  $\langle e_{i_1} \rangle = F_{\leq s_1}$ ,  $\langle e_{i_1}, e_{i_2} \rangle = F_{\leq s_2}$  описывает один топологический признак в размерности  $j$ , который “рождается” в  $s_1$  и “исчезает” в  $s_2$ . Он представлен сегментом  $[s_1, s_2]$  в бар-коде степени  $j$ . А каждый фильтрованный одномерный комплекс с тривиальным дифференциалом,  $\partial_j e_i = 0$ ,  $\langle e_i \rangle = F_{\leq s}$  описывает топологический признак размерности  $j$ , который “рождается” в  $s$  и никогда не “исчезает”. Он представлен полупрямой  $[r, +\infty[$  в бар-коде степени  $j$ .

Доказательство Теоремы 3 приведено в [7], Секция 2, см. также [8, 9]. Мы описываем его в Приложении 9 для удобства читателя. По сути, можно привести  $\mathbb{R}$ -фильтрованный комплекс к требуемой прямой сумме простых  $\mathbb{R}$ -фильтрованных комплексов по индукции, начиная с наименьших базисных элементов степени один, таким образом, что манипуляции с базисными элементами степени  $j$  не разрушают полученное разложение на простые  $\mathbb{R}$ -фильтрованные комплексы в степени  $j-1$  и в более низких частях фильтрации в степени  $j$ .

Пусть  $f : \Theta \rightarrow \mathbb{R}$  — это гладкая, или, в более общем случае, кусочно-гладкая непрерывная

функция, такая что подуровневые множества  $\Theta_{f \leq c} = \{\theta \in \Theta \mid f(\theta) \leq c\}$  компактны.

Существуют разные фильтрованные цепные комплексы, вычисляющие гомологии топологических пространств  $\Theta_{f \leq c}$ , например, Čech комплексы, симплициальные комплексы или CW-комплексы. Без потери общности можно предположить, что кусочно-гладкая функция  $f$  является гладкой, в противном случае всегда можно заменить  $f$  на гладкое приближение. Добавив небольшое возмущение, можно также предположить, что критические точки  $f$  являются невырожденными, а критические значения различными.

Один фильтрованный цепной комплекс, естественно связанный с такой функцией  $f$ , с подкомплексами  $F_s C_*$ , вычисляющими гомологии подуровневых множеств  $\Theta_{f \leq s}$ , — это комплекс Морса, см. приложение А и [4, 5] для дополнительной информации. Комплекс Морса определяется следующим образом. Базисные элементы в  $k$ -векторных пространствах  $C_j$  соответствуют критическим точкам  $f$  индекса  $j$  с выбором ориентации. Ориентация здесь — это выбор ориентации  $j$ -мерного подпространства касательного пространства в критической точке, на котором гессиан отрицательно определен. Матрица  $\partial_j$  состоит из числа траекторий градиента, учитываемых со знаками, между критическими точками индекса  $j$  и индекса  $(j-1)$ . Естественная фильтрация на комплексе Морса задается значениями критических точек. Базисные элементы в  $F_s C_*$  соответствуют критическим точкам с критическим значением  $s$  и ниже.

**Утверждение 4.** Пусть  $r$  — минимум, который не является глобальным. Тогда  $r$  представляет тривиальный класс гомологий в комплексе Морса, вычисляющем гомологии  $\Theta_{f \leq c}$  для достаточно большого  $c$ , т.е.  $r$  является нижним базисным элементом одного из двумерных комплексов с тривиальными гомологиями в канонической форме. Таким образом,  $r$  ассоциировано с 1-седлом  $q$ . Это 1-седло из определения 1, т.е.  $q$  — это 1-седло, возле которой связная компонента подуровневого множества, соответствующая  $r$ , поглощается связной компонентой с более низким минимумом. Сегмент  $[f(p), f(q)]$ , следовательно, соответствует минимуму  $r$  по определению 1.

**Доказательство.** Пусть  $q$  — это 1-седловая точка, возле которой связная компонента, соответствующая  $r$ , поглощается связной компонентой с более низким минимумом  $r$ . Гомологии  $H_0(\Theta_{f \leq f(q)-\epsilon})$  линейно порождаются классами связных компонент  $\Theta_{f \leq f(q)-\epsilon}$ . Они соответствуют генераторам комплекса Морса, заданным минимумами ограничения  $f$  на каждую связную компоненту. В ком-

плексе Морса, вычисляющем гомологии  $\Theta_{f \leq f(q)+\epsilon}$ , генератор, соответствующий локальному минимуму  $p$ , после приведения к канонической форме, равен границе генератора, соответствующего 1-седлу  $q$ , так как класс гомологий соответствующий генератору  $p$  исчезает в  $f(q)$ . Следовательно,  $q$  ассоциирован с  $p$  в канонической форме.  $\square$

Аналогичные результаты справедливы для других типов фильтрованных комплексов, представляющих гомологии подуровневых множеств, таких как кубические или CW-комплексы, для непрерывных кусочно-гладких функций (см., например, [16]).

Общее количество различных топологических признаков в подуровневых множествах  $\Theta_{f \leq c}$  функции потерь можно непосредственно получить из бар-кода. А именно, количество пересечений горизонтальной линии на уровне  $c$  с сегментами в бар-коде индекса  $j$  дает количество независимых топологических признаков размерности  $j$  в  $\Theta_{f \leq c}$ , см. примеры в разделе 4.

### 3. АЛГОРИТМ ДЛЯ ВЫЧИСЛЕНИЯ БАР-КОДОВ МИНИМУМОВ

В этом разделе мы описываем алгоритм для вычисления бар-кодов локальных минимумов. Алгоритм использует определение 1 бар-кодов из раздела 2, основанное на эволюции связных компонент подуровневых множеств.

Для анализа поверхности заданной функции  $f : \Theta \rightarrow \mathbb{R}$  мы сначала строим ее приближение на основе конечного графа. Для этого мы рассматриваем случайно выбранное подмножество точек  $\{\theta_1, \dots, \theta_N\} \in \Theta$  и строим граф с этими точками в качестве вершин. Мы соединяем вершины ребром, если точки близки. Таким образом, для каждой вершины  $\theta_n$ , сравнивая  $f(\theta_n)$  с  $f(\theta_{n'})$  для соседей  $\theta_{n'}$ , вершины  $\theta_n$ , мы можем описать локальную топологию возле точки  $\theta_n$ . В то же время связные компоненты подуровневых множеств  $\Theta_{f \leq c}$  соответствуют, см. Утверждение 5 ниже, связным компонентам подграфа  $\Xi_{f \leq c}$  точек  $\theta_n$ , таких что  $f(\theta_n) \leq c$ .

Два технических момента здесь — это выбор точек  $\theta_n$  и определение близости, т.е. когда точки соединяются ребром. В наших экспериментах мы выбираем точки равномерно из некоторого куба. Чтобы добавить ребра, мы вычисляем ориентированный граф  $k$ -ближайших соседей на заданных точках, затем отбрасываем ориентацию ребер и проверяем, что расстояние между соседями не превышает  $c(D)N^{-\frac{1}{D}}$ , где  $D$  — размерность вход-

ных данных функции  $f$ . В наших экспериментах мы используем  $k = 2D$ .

Теперь мы описываем часть алгоритма, который вычисляет бар-коды функции по ее приближению на основе графа, описанного выше. Основная идея заключается в том, чтобы отслеживать эволюцию связных компонент подуровневых подмножеств графа  $\Xi_{f \leq c} = \{\theta_n \mid f(\theta_n) \leq c\}$  при увеличении  $c$ .

Для простоты предположим, что точки  $\theta$  упорядочены по значению функции  $f$ , т.е. для  $n < n'$  у нас есть  $f(\theta_n) < f(\theta_{n'})$ . В этом случае нас интересует эволюция связных компонент в процессе последовательного добавления вершин  $\theta_1, \theta_2, \dots, \theta_N$  в граф, начиная с пустого графа. Обозначим подграф на вершинах  $\theta_1, \dots, \theta_n$  как  $\Xi_n$ . Когда мы добавляем новую вершину  $\theta_{n+1}$  к  $\Xi_n$ , есть три возможности эволюции связных компонент:

1. Вершина  $\theta_{n+1}$  имеет нулевую степень в  $\Xi_{n+1}$ . Это означает, что  $\theta_{n+1}$  является локальным минимумом  $f$  и образует новую связную компоненту в подуровневом множестве.
2. Все соседи  $\theta_{n+1}$  в  $\Xi_{n+1}$  принадлежат одной связной компоненте в  $\Xi_n$ .
3. Все соседи  $\theta_{n+1}$  в  $\Xi_{n+1}$  принадлежат  $K \geq 2$  связным компонентам  $s_1, s_2, \dots, s_K$  в  $\Xi_n$ . Таким образом, все эти компоненты объединяются в одну связную компоненту в  $\Xi_{n+1}$ .

В третьем случае, согласно определению 1 раздела 2, точка  $\theta_{n+1}$  является дискретной 1-седловой точкой. Таким образом, одна из компонент  $s_k$  поглощает все остальные. Эта компонента, которая имеет наименьшее минимальное значение. Для оставшихся компонент это задает их бар-коды: для  $s_i, i \neq k$  пара рождение-исчезновение равна  $[\min_{\theta \in s_i} f(\theta); f(\theta_{n+1})]$ . Мы резюмируем эту процедуру в алгоритме 1.

### 4. БАР-КОДЫ БЕНЧМАРК ФУНКЦИЙ И СХОДИМОСТЬ АЛГОРИТМА

В этом разделе мы применяем наш алгоритм для описания топологии тестовых функций. В подразделе 4 мы применяем алгоритм к визуализируемым 2-мерным примерам, а в подразделе 4 мы проверяем сходимость на бенчмарк функциях более высоких размерностей. В разделе 5 мы применяем наш алгоритм к анализу поверхностей потерь **малых нейронных сетей**.

Мы применяем наш алгоритм к нескольким функциям  $f : \mathbb{R}^D \rightarrow \mathbb{R}$  из **Global Optimization Benchmark** (см., например, [17]). Эти функции предназначены для обмана алгоритмов глобаль-

ной оптимизации, они очень сложные, имеют много локальных минимумов и седловых точек даже для малых размерностей. Таким образом,

вычисление бар-кодов и соответствия минимум-седло для этих функций также представляет собой достаточно трудную задачу.

---

**Алгоритм 1:** Вычисление бар-кодов минимумов для функции на графе
 

---

**Input:** Неориентированный граф  $G = (V, E)$ ; функция  $f$  на вершинах графа

**Output:** Bar-codes: список пар “рождение”–“исчезновение”

$S \leftarrow \{\}$

**for**  $\theta \in V$  *в возрастающем порядке*  $f(\theta)$  **do**

$S' \leftarrow \{s \in S \mid \exists \theta' \in s \text{ такие что } (\theta, \theta') \in E \text{ и } f(\theta) > f(\theta')\}$

**if**  $S' = \emptyset$  **then**

$S \leftarrow S \sqcup \{\{\theta\}\}$

**else**

$f^* \leftarrow \min f(\theta')$  для  $\theta' \in \prod_{s \in S'} s$

**for**  $s \in S'$  **do**

$f^s \leftarrow \min f(\theta')$  для  $\theta' \in s$

**if**  $f^s \neq f^*$  **then**

                Bar-codes  $\leftarrow$  Bar-codes  $\sqcup \{(f^s, f(\theta))\}$

**end**

$s_{\text{new}} \leftarrow \left( \prod_{s \in S'} s \right) \sqcup \{\theta\}$

$S \leftarrow (S \setminus S') \sqcup \{s_{\text{new}}\}$

**end**

**end**

**for**  $s \in S$  **do**

$f^s \leftarrow \min f(\theta')$  для  $\theta' \in s$

    Bar-codes  $\leftarrow$  Bar-codes  $\sqcup \{(f^s, \infty)\}$

**end**

**Return** Bar-codes

---

### Бар-коды 2D бенчмарк функций

Для начала мы вычисляем бар-коды нескольких двумерных бенчмарк функций. Мы визуализируем полученные бар-коды и соответствие минимум-седло. Затем мы проводим эксперименты, чтобы оценить, как количество точек, используемых для вычисления бар-кодов, влияет на качество ответа. Мы рассматриваем следующие тестовые целевые функции:

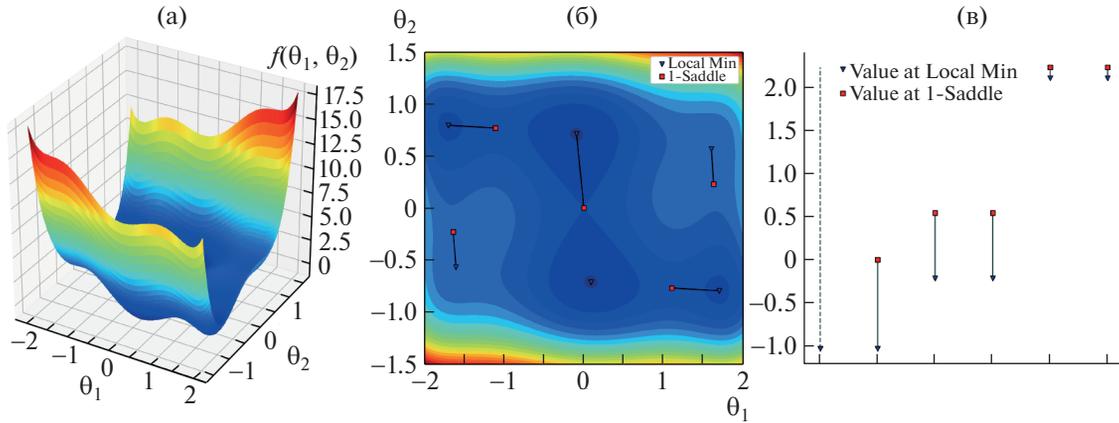
1. **HumpCamel6** функция  $f : [-2, 2] \times [-1.5, 1.5] \rightarrow \mathbb{R}$  с 6 локальными минимумами (рис. 2):

$$f(\theta_1, \theta_2) = (4 - 2.1\theta_1^2 + \theta_1^4/3)\theta_1^2 + \theta_1\theta_2 + (-4 + 4\theta_2^2)\theta_2^2.$$

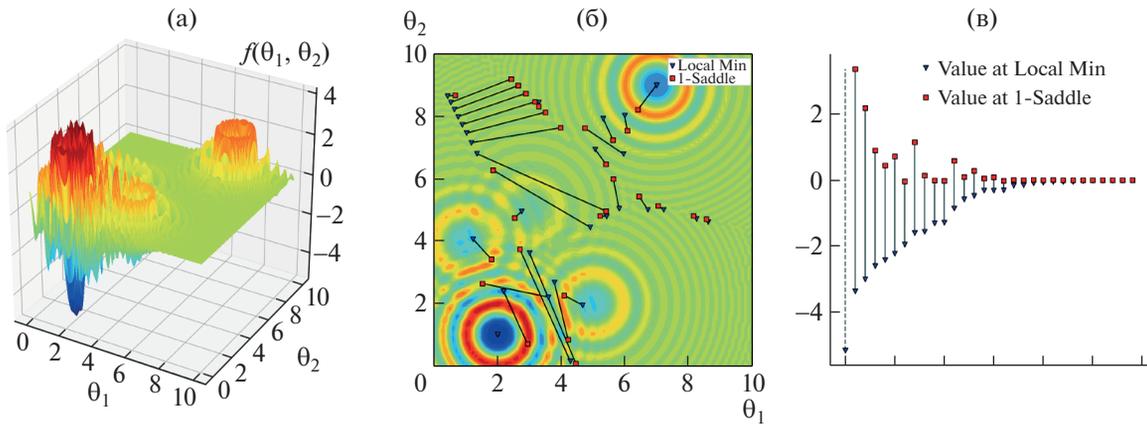
2. **Langermann** целевая функция  $f : [0, 10]^2 \rightarrow \mathbb{R}$  (рис. 3):

$$f(\theta_1, \theta_2) = -\sum_{i=1}^5 \frac{c_i \cos(\pi[(\theta_1 - a_i)^2 + (\theta_2 - b_i)^2])}{\exp\left(\frac{1}{\pi}[(\theta_1 - a_i)^2 + (\theta_2 - b_i)^2]\right)}.$$

Чтобы сохранить графики простыми, мы отображали соответствие минимум-седло только для верхних 30 сегментов в бар-коде, отсортированных по размеру кластера (количество точек в момент поглощения кластера). Как мы видим на рис. 2, соответствие между минимумом и седлом не тривиально. Для многих минимумов соответствующие седла находятся довольно далеко. В частности, это наблюдение иллюстрирует замеча-



**Рис. 2.** Функция HumpCamel6, ее соответствие минимум-седло и бар-код, вычисленный Алгоритмом 1. (а) График поверхности 3D, (б) соответствие между минимумами и 1-седлами, бар-код локальных минимумов.



**Рис. 3.** Функция Langermann, ее соответствие минимум-седло и бар-коды, вычисленные Алгоритмом 1. (а) График поверхности 3D, (б) соответствие между минимумами и 1-седлами, бар-код локальных минимумов.

ние, обсуждаемое на рис. 1: соответствующее минимуму 1-седло часто не является ближайшим седлом.

3. **Wavy** тестовая целевая функция  $f : [-\pi, \pi]^2 \rightarrow \mathbb{R}$  (рис. 4):

$$f(\theta_1, \theta_2) = 1 - \frac{1}{2} \sum_{d=1}^2 \cos(10\theta_d) \cdot e^{-\frac{\theta_d^2}{2}}.$$

Функция Wavy симметрична относительно диэдральной группы  $D_4$  порядка 8. Таким образом, ее пары критических значений минимум-седло образуют мультиплеты, формируя простые представления группы  $D_4$ .

4. **HolderTable** тестовая целевая функция  $f : [-10, 10]^2 \rightarrow \mathbb{R}$  (рис. 5):

$$f(\theta_1, \theta_2) = - \left| e^{\left| \frac{\sqrt{\theta_1^2 + \theta_2^2}}{\pi} \right|} \sin(\theta_1) \cos(\theta_2) \right|.$$

**Сходимость Алгоритма**

В этом подразделе мы доказываем и эмпирически проверяем сходимость нашего алгоритма, когда количество точек  $N$ , используемых для построения бар-кодов, стремится к бесконечности. Для сравнения двух бар-кодов мы принимаем **расстояние Bottleneck** [9, 18], также известное как расстояние Вассерштейн- $\infty$ ,  $\mathbb{W}_\infty$ , на диаграммах персистентности (двумерные облака точек пар рождение-исчезновение):

$$\mathbb{W}_\infty(\mathcal{D}, \mathcal{D}') := \inf_{\pi \in \Gamma(\mathcal{D}, \mathcal{D}')} \sup_{a \in \mathcal{D} \cup \Delta} |a - \pi(a)|. \tag{2}$$

Здесь  $\Delta$  обозначает “диагональ” (пары с “рождением” равным “исчезновению”), а  $\Gamma(\mathcal{D}, \mathcal{D}')$  обозначает набор частичных соответствий между  $\mathcal{D}$  и  $\mathcal{D}'$ , определенных как биекции между  $\mathcal{D} \cup \Delta$  и  $\mathcal{D}' \cup \Delta$ .

Для функции  $f : \Theta \rightarrow \mathbb{R}$  пусть  $\mathcal{D}_f^*$  обозначает ее бар-код. Наш алгоритм использует конечное число  $N$  случайно выбранных точек  $\Theta_N = \{\theta_1, \dots, \theta_N\}$  для

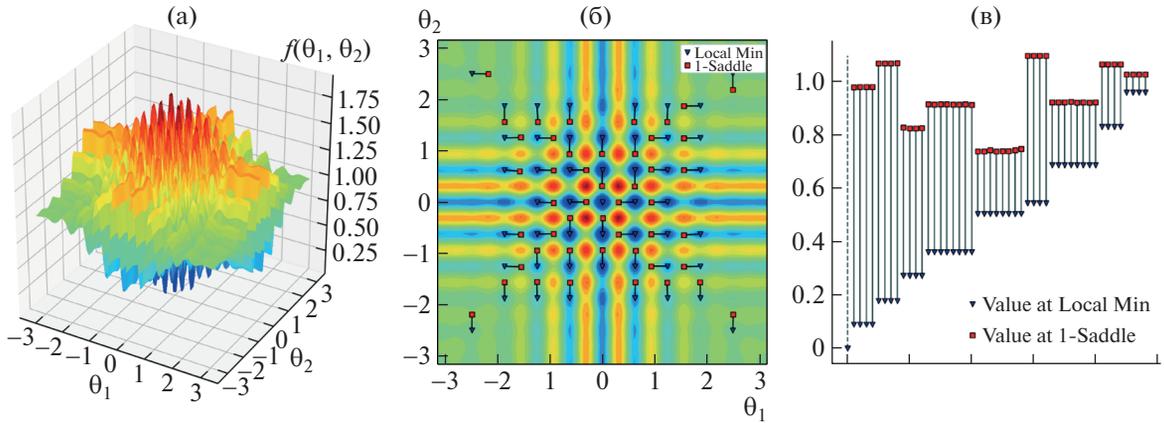


Рис. 4. Функция Wavy, ее соответствие минимум-седло и бар-код, вычисленный с помощью Алгоритма 1. (а) График поверхности 3D, (б) соответствие между минимумами и 1-седлами, бар-код локальных минимумов.

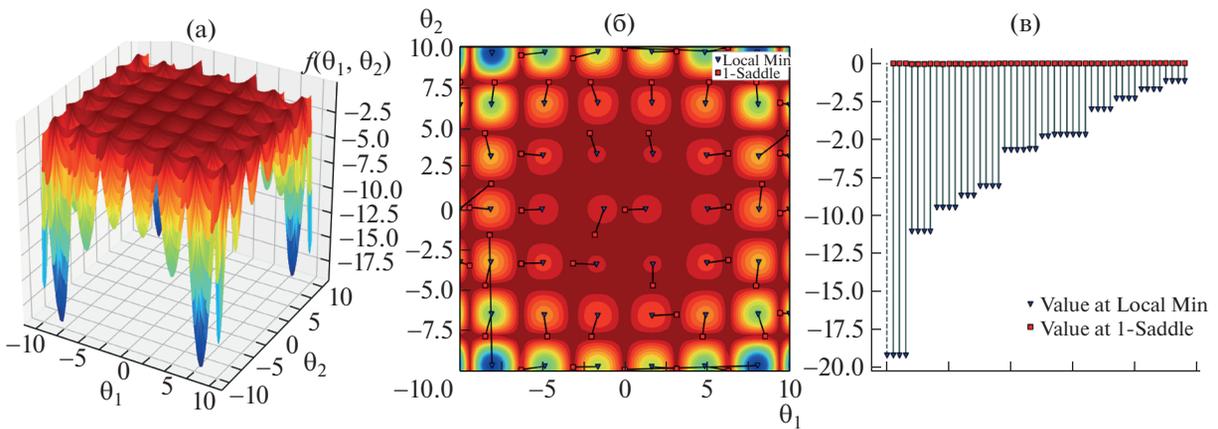


Рис. 5. Функция HolderTable, ее соответствие минимум-седло и бар-код, вычисленный с помощью Алгоритма 1. (а) График поверхности 3D, (б) соответствие между минимумами и 1-седловыми точками, (в) бар-код локальных минимумов.

вычисления приближения  $\hat{\mathcal{D}}_f(\Theta_N)$  истинного бар-кода  $\mathcal{D}_f^*$ . Пусть  $C$  обозначает константу Липшица для  $f$ .

**Утверждение 5.** Пусть для любого  $\theta \in \Theta$  существует  $\theta_i \in \Theta_N$  такое, что  $|\theta - \theta_i| < \varepsilon$ . Тогда

$$\mathbb{W}_\infty(\hat{\mathcal{D}}_f(\Theta_N), \mathcal{D}_f^*) < C\varepsilon.$$

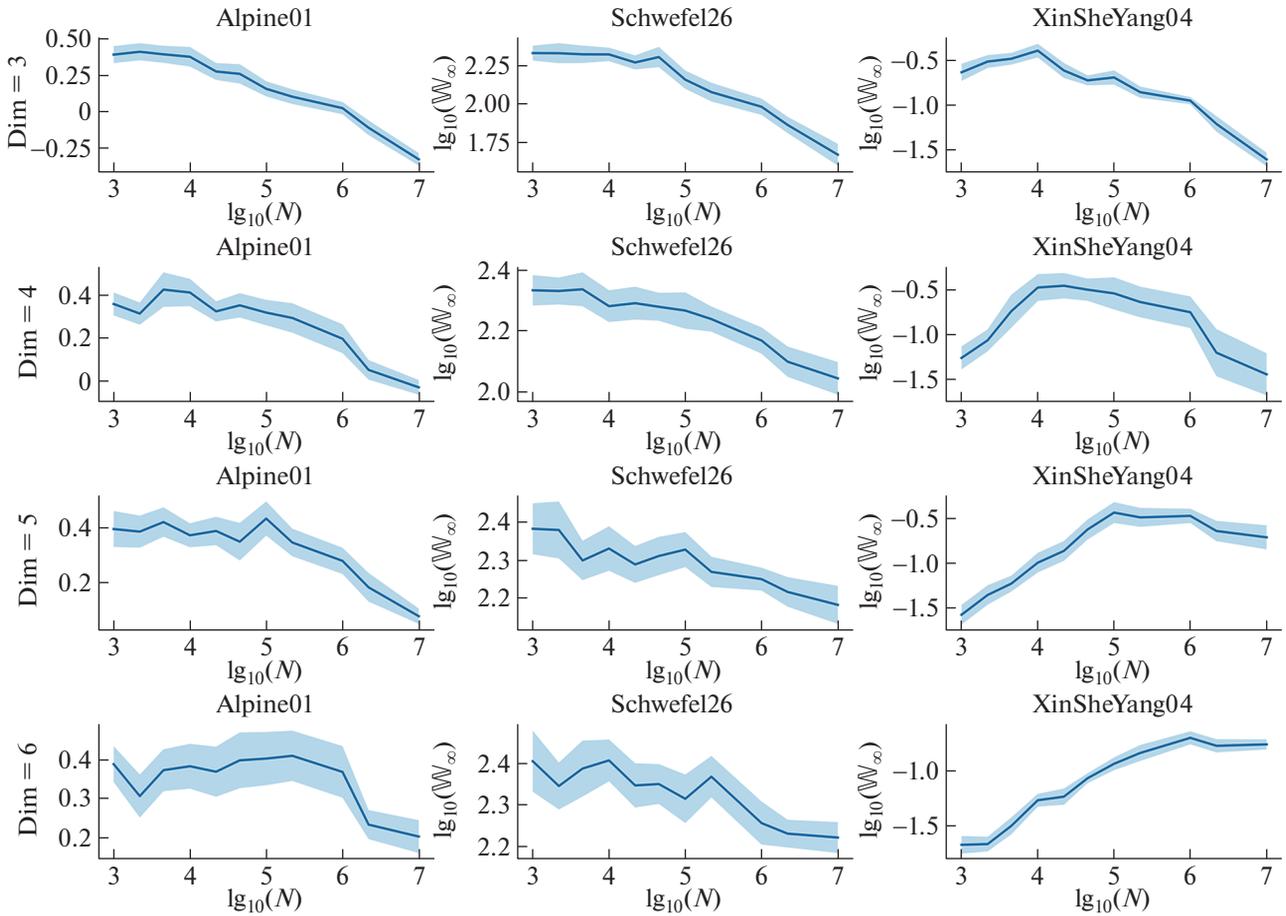
*Доказательство.* Это следует, например, из ([16], Лемма 1).  $\square$

Отметим, что для типичной выборки  $\Theta_N$  максимальное расстояние от  $\theta \in \Theta$  до  $\Theta_N$  составляет  $\sim N^{-\frac{1}{D}}$ . Из этого следует, что  $\hat{\mathcal{D}}_f(\Theta_N)$  сходится к  $\mathcal{D}_f^*$  при  $N \rightarrow \infty$  для любой типичной последовательности выборок  $\{\Theta_N\}$ . И в частности,

$$\mathbb{W}_\infty(\hat{\mathcal{D}}_f(\Theta_N), \hat{\mathcal{D}}_f(\Theta_{N'})) \rightarrow 0 \quad (3)$$

при  $N \rightarrow \infty$  для любой типичной пары последовательностей  $\{\Theta_N\}, \{\Theta_{N'}\}$ .

Мы проверили условие (3) на нескольких тестовых функциях из набора для глобальной оптимизации. Каждая из тестовых функций представляет собой серию функций произвольной размерности. Даже для небольших размерностей они чрезвычайно сложны (содержат экспоненциальное количество локальных минимумов и седловых точек). Для каждой функции  $a$  и размерности  $D = 3, 4, 5, 6$  мы рассматриваем различные размеры выборок  $\log_{10} N \in [3, 3.5, 4, \dots, 7]$ . Для каждой тройки  $(f, D, N)$  мы случайно выбираем  $R$  пар  $(\{\Theta_N\}_r, \{\Theta_{N'}\}_r)$  облаков точек размера  $N$ . Затем для каждого из  $2R$  облаков точек мы вычисляем бар-код нашим алгоритмом и в каждой паре измеряем расстояние  $\mathbb{W}_\infty$  между бар-кодами. Наконец, мы берем среднее значение



**Рис. 6.** Уменьшение логарифма расстояния Bottleneck ( $\mathbb{W}_\infty$ ) между парами диаграмм персистентности на выборках размера  $N$  для функций Alpine01, Schwefel26, XinSheYang04 из набора для глобальной оптимизации в размерностях  $D \in \{3, 4, 5, 6\}$ .

$$\begin{aligned} \mathbb{E}_{\Theta_N, \Theta'_N} \mathbb{W}_\infty(\hat{\mathcal{D}}_f(\Theta_N), \hat{\mathcal{D}}_f(\Theta'_N)) &\approx \\ \approx \frac{1}{R} \sum_{r=1}^R \mathbb{W}_\infty(\hat{\mathcal{D}}_f((\Theta_N)_r), \hat{\mathcal{D}}_f((\Theta'_N)_r)). \end{aligned} \quad (4) \quad f(\theta_1, \dots, \theta_D) = \left[ \sum_{d=1}^D \sin^2(\theta_d) - e^{-\sum_{d=1}^D \theta_d^2} \right] e^{-\sum_{d=1}^D \sin^2 \sqrt{|\theta_d|}}.$$

Рассматриваемые функции:

1. **Alpine01** функция  $f : [-10, 10]^D \rightarrow \mathbb{R}$  определена как

$$f(\theta_1, \dots, \theta_D) = \sum_{d=1}^D |\theta_d \sin \theta_{x_d} + 0.1 \theta_d|.$$

2. **Schwefel26** функция  $f : [-500, 500]^D \rightarrow \mathbb{R}$  определена как:

$$f(\theta_1, \dots, \theta_D) = 418.9829 \cdot D - \sum_{d=1}^D \theta_d \sin(\sqrt{|\theta_d|}).$$

3. **XinSheYang04** функция  $f : [-10, 10]^D \rightarrow \mathbb{R}$  определена как:

Для каждой пары  $(f, D)$  мы резюмируем зависимость от размера облака  $N$  в виде графика в десятичной логарифмической шкале. Мы эмпирически наблюдаем, что для достаточно больших  $N$

$$\log(\mathbb{W}_\infty(\hat{\mathcal{D}}_f(\Theta_N), \hat{\mathcal{D}}_f(\Theta'_N))) \sim -\frac{1}{D} \log(N),$$

как и ожидалось на основе Предложения 5.

Результаты представлены на рис. 6.

В табл. 1 время выполнения Алгоритма 1, включая поиск HNSW NN, на случайных выборках из 106 точек, сравнивается со временем выполнения для платформы GUDHI на сетке из 106 точек для функций Alpine01 и Schwefel26 в размерностях  $D \in \{3, 4, 5, 6, 7\}$ .

**Таблица 1.** Время выполнения на  $10^6$  точек, на Intel(R) Xeon(R) CPU E5-2698 v4 @2.20GHz

	Alpine01	Schwefel26								
	3D	3D	4D	4D	5D	5D	6D	6D	7D	7D
Gudhi	9	10	28	29	72	79	160	181	767	863
Alg 1	12	12	14	14	16	16	18	18	22	21

## 5. ТОПОЛОГИЯ ФУНКЦИЙ ПОТЕРЬ НЕЙРОННЫХ СЕТЕЙ

Очевидно, что некоторые минимумы могут больше мешать алгоритмам, основанным на градиентных методах. Например, два минимума на рис. 7 локально выглядят одинаково, но представляют разные препятствия для оптимизации на основе градиентного спуска. Траектория алгоритма, основанного на градиентном спуске, может испытывать трудности при попытке выйти из окрестности локального минимума. Чтобы достичь точки с меньшим значением целевой функции, путь, исходящий из окрестности данного локального минимума, должен подниматься к точкам с большим значением функции. Минимальное значение такого более высокого значения целевой функции точно соответствует значению функции в 1-седле, связанном с минимумом, из определения бар-кода. Чем больше сегмент минимума в бар-коде, тем сложнее траектории обучения, основанной на градиентном спуске, выйти из окрестности минимума, чтобы достичь точек с меньшим значением целевой функции.

В этом разделе мы вычисляем и анализируем бар-коды небольших полносвязных нейронных сетей, имеющих не более трех скрытых слоев.

Для нескольких архитектур нейронных сетей известно множество результатов о поверхности потерь (см., например, [19, 20] и ссылки в них). Различные геометрические и топологические свойства поверхностей потерь были изучены в [21–24].

Мы анализировали нейронные сети, которые являются небольшими. Однако наш метод позволяет полностью исследовать поверхность потерь, в отличие от стохастического исследования поверхностей потерь более высокой размерности. Подчеркнем, что даже с практической точки зрения важно сначала понять поведение бар-кодов в простейших примерах, где все схемы оптимизации гиперпараметров несложно отключить.

Для каждой анализируемой нейронной сети функция потерь представляет собой среднеквадратичную ошибку при прогнозировании (случайно выбранной) функции  $g : [-\pi, \pi] \rightarrow \mathbb{R}$ , заданной как

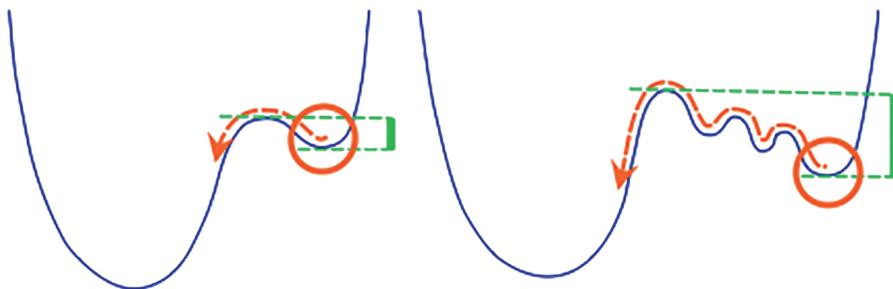
$$g(x) = 0.31 \cdot \sin(-x) - 0.72 \cdot \sin(-2x) - 0.21 \cdot \cos(x) + 0.89 \cdot \cos(2x)$$

плюс  $l_2$ -регуляризация. Ошибка вычисляется для прогноза на равномерно распределенных входах

$$x \in \left\{ -\pi + \frac{2\pi}{100} k \mid k = 0, 1, \dots, 100 \right\}.$$

## 6. ТОПОЛОГИЯ ФУНКЦИЙ ПОТЕРЬ НЕЙРОННЫХ СЕТЕЙ

Мы рассмотрели полносвязные нейронные сети с одним скрытым слоем из 2 и 3 нейронов, двумя скрытыми слоями с  $2 \times 2$ ,  $3 \times 2$  и  $3 \times 3$  нейронами, и тремя скрытыми слоями с  $2 \times 2 \times 2$  и  $3 \times 2 \times 2$  нейронами с активацией *ReLU*. Мы вычислили бар-коды функций потерь на гиперкубических множествах  $\Theta$ , которые были выбраны на



**Рис. 7.** Бар-коды количественно оценивают топологические препятствия, создаваемые локальными минимумами для оптимизации на основе градиентных методов. Здесь два локальных минимума локально выглядят одинаково, но явно представляют разные препятствия для обучения градиентными методами. Эти препятствия количественно оцениваются длинами зеленых сегментов, которые связаны с локальными минимумами в бар-коде.

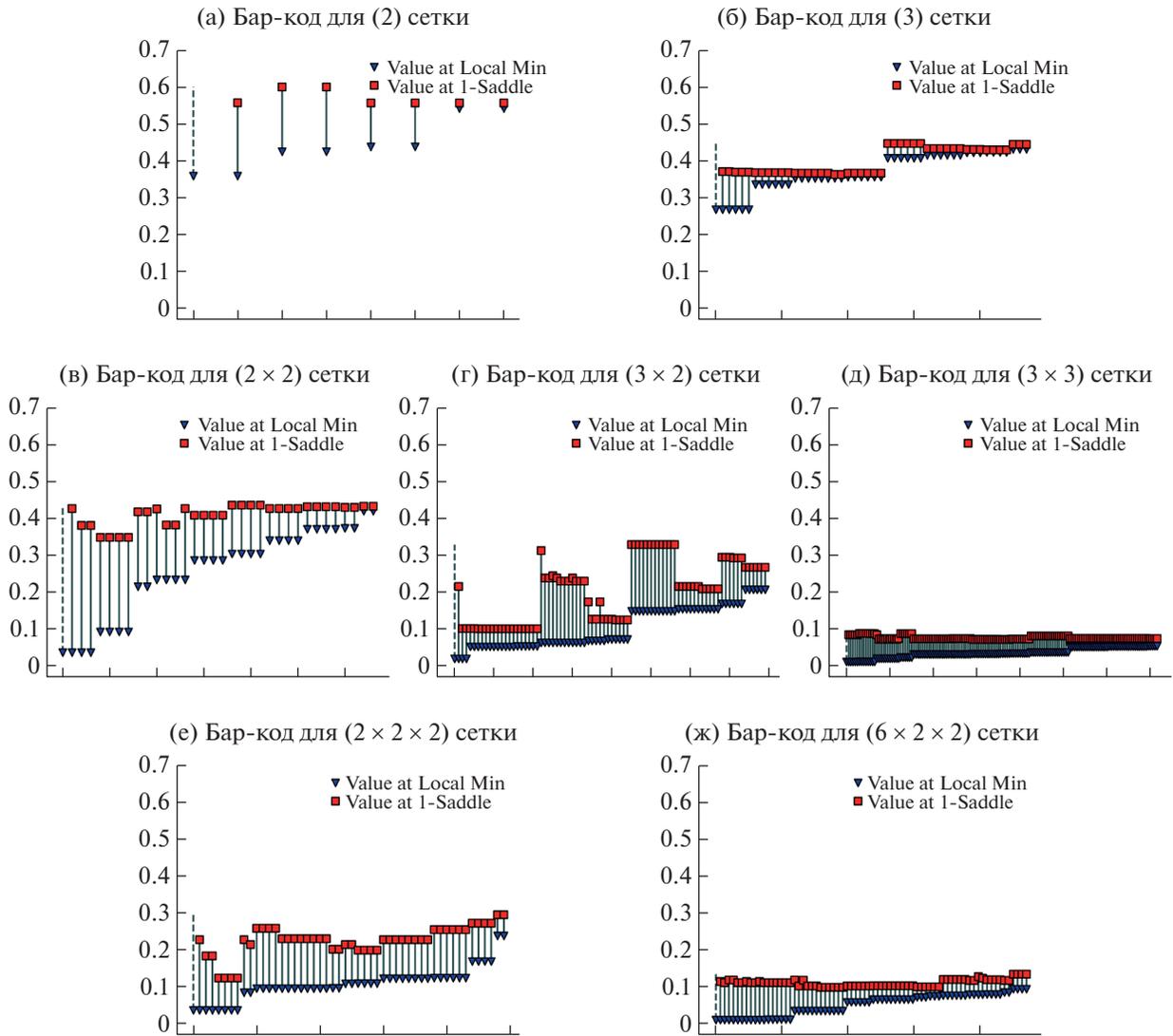


Рис. 8. Бар-коды функций потерь нейронных сетей.

основе типичного диапазона параметров минимумов. Результаты представлены на рис. 8.

Наши результаты дают следующие два наблюдения:

1. Бар-коды расположены в очень малой нижней части диапазона значений; обычно максимальное значение функции было около 200 и выше, и седла, соединенные с минимумами, лежат гораздо ниже 1.

2. С увеличением глубины и ширины нейронной сети бар-коды опускаются ниже.

Например, верхние границы бар-кодов однослойной сети (2) находятся в диапазоне [0.55, 0.65], двухслойной сети (2 × 2) — в диапазоне [0.35, 0.45], и трехслойной сети (2 × 2 × 2) — в диапазоне [0.1, 0.3].

**Следствия для обучения**

Все минимумы в рассмотренных случаях расположены в нижней части диапазона функции потерь, и они опускаются ниже по мере увеличения глубины и ширины нейронной сети. Траектории градиентного потока всегда имеют минимумы в качестве конечных точек для всех, кроме подмножества начальных точек меры ноль. Следовательно, по существу, любая конечная точка градиентного потока дает достаточно хорошее решение задачи регрессии. Точность решения увеличивается с увеличением глубины и ширины нейронной сети.

Во время обучения траектория градиентного спуска не может застрять на высоких локальных минимумах, так как по существу все минимумы расположены в очень малой нижней части диапазона функции. Возможно, существуют некоторые

шумовые минимумы, которые находятся немного выше, но из них легко выйти во время обучения, так как их бар-код низок, что означает, что всегда существует путь ухода с низким препятствием.

Если длина сегмента бар-кода мала для любого локального минимума, то всегда существует путь с малым значением функции потерь к более низкому минимуму, что означает, что методы оптимизации на основе градиентного спуска могут в принципе достичь самого низкого минимума во время обучения.

Если бар-коды всех локальных минимумов низки, это также означает, что любые два таких минимума могут быть соединены путем с малым значением функции потерь. Это означает, что для любых двух локальных минимумов существует непрерывное преобразование с небольшими значениями потерь между прогнозами одного минимума и прогнозами другого минимума, что подразумевает, что их прогнозы в определенном смысле эквивалентны.

### Стратегия вычисления бар-кодов для глубоких нейронных сетей

Исследование поверхности потерь с использованием методов топологического анализа данных может быть расширено на глубокие нейронные сети. Стратегия для глубоких нейронных сетей основана на предложении 2 выше. Идея заключается в том, чтобы действовать методом градиентного спуска по пути, начиная с данного минимума и идущего к точке с более низким минимумом. Деформация пути под действием градиентного потока задается действием компонента градиента, который ортогонален касательному направлению пути. Компонент градиента, который параллелен направлению касания пути, поглощается перепараметризацией пути. Затем формула (1) дает оценку значения 1-седла, соответствующего этому минимуму в бар-коде.

#### Следствия для обобщения

Мы провели некоторые эксперименты, которые показывают, что существует корреляция между высотой бар-кодов групп низких минимумов с приблизительно одним и тем же значением функции потерь, и их ошибками обобщения. Таким образом, по крайней мере при определенных условиях, чем ниже бар-код для минимумов с одинаковым значением функции потерь, тем лучше их свойства обобщения.

## 7. ЗАКЛЮЧЕНИЕ

В этой работе мы представили методологию для анализа графиков функций, в частности, поверхностей потерь нейронных сетей. Методология основана на вычислении топологических инвариантов, называемых бар-кодами.

Для вычисления бар-кодов мы использовали конструкцию на графах, для аппроксимации функции потерь. Затем мы применили разработанный нами алгоритм для вычисления бар-кодов локальных минимумов функций. Наши экспериментальные результаты вычисления бар-кодов для небольших нейронных сетей привели к двум основным наблюдениям.

Во-первых, все бар-коды находятся в очень малой нижней части общего диапазона функции. Во-вторых, бар-коды опускаются ниже по мере увеличения глубины и ширины нейронной сети. С практической точки зрения это означает, что оптимизация градиентного спуска не может застрять в высоких локальных минимумах, и также, что нетрудно перейти от одного локального минимума к другому (с меньшим значением) во время обучения.

Разработанный нами метод имеет несколько направлений для дальнейших исследований. Хотя мы тестировали метод на небольших нейронных сетях, его можно применить к современным нейронным сетям, таким как сверточные сети (например, ResNet, VGG, AlexNet, U-Net, см. [25]) для задач обработки изображений. Однако в этом случае графовая аппроксимация, которую мы используем, требует разумного выбора представителей вершин графа, так как плотное заполнение области точками вычислительно неприемлемо. Очевидно, существуют также связи, заслуживающие дальнейшего исследования, между бар-кодами локальных минимумов и оптимальными значениями коэффициентов скорости обучения, или скоростями сходимости во время обучения.

### Приложение

#### А. ГРАДИЕНТНЫЙ КОМПЛЕКС МОРСА

Градиентный комплекс Морса  $(C_*, \partial_*)$  определяется следующим образом.

Для общего  $f$  критические точки  $p_\alpha$ , где  $df|_{T_{p_\alpha}} = 0$ , являются изолированными. Вблизи каждой критической точки  $p_\alpha$  функция  $f$  может быть записана как  $f = \sum_{l=1}^j -(x^l)^2 + l = j \sum_{l=1}^n (x^l)^2$  в

некоторых локальных координатах. Индекс критической точки определяется как размерность множества направлений, указывающих вниз в этой точке, или отрицательного подпространства Гесса:  $\text{index}(p_\alpha) = j$ . Обозначим

$$C_j = \bigoplus_{\text{index}(p_\alpha)=j} [p_\alpha, \text{or}(T_{p_\alpha}^-)],$$

где  $\text{or}$  — ориентация на отрицательном подпространстве  $T_{p_\alpha}^- = T_{p_\alpha}^- \oplus T_{p_\alpha}^+$  Гесса  $\partial^2 f$ .

Пусть

$$\mathcal{M}(p_\alpha, p_\beta) = \left\{ \gamma : \mathbb{R} \rightarrow M^n \mid \dot{\gamma} = -(\text{grad}_g f)(\gamma(t)), \right. \\ \left. \lim_{t \rightarrow -\infty} \gamma(t) = p_\alpha, \lim_{t \rightarrow +\infty} \gamma(t) = p_\beta \right\} / \mathbb{R}$$

это множеством градиентных траекторий, соединяющих критические точки  $p_\alpha$  и  $p_\beta$ , где естественное действие  $\mathbb{R}$  происходит путем сдвига  $\gamma(t) \mapsto \gamma(t + \tau)$ .

Если  $\text{index}(p_\beta) = \text{index}(p_\alpha) - 1$ , то в общем случае множество  $\mathcal{M}(p_\alpha, p_\beta)$  конечно. Пусть

$$\#\mathcal{M}([p_\alpha, \text{or}], [p_\beta, \text{or}])$$

в этом случае обозначает количество траекторий, подсчитанных со знаками, учитывая выбор ориентации, между критическими точками  $p_\alpha$  и  $p_\beta$ .

Оператор  $\partial_j$  определяется следующим образом:

$$\partial_j [p_\alpha, \text{or}] = \sum_{\text{index}(p_\beta) = j-1} [p_\beta, \text{or}] \#\mathcal{M}(p_\alpha, p_\beta).$$

Естественная фильтрация на построенном цепном комплексе  $(C_*, \partial_*)$  задается значениями функции в критических точках.

Описание критических точек на многообразии  $\Theta$  с непустой границей  $\partial\Theta$  немного модифицируется следующим образом. Связные компоненты подуровней функции рождаются также в локальном минимуме ограничения  $f$  на границе  $f|_{\partial\Theta}$ , если  $\text{grad}f$  указывает внутрь многообразия  $\Theta$ . Объединение двух связных компонент также может произойти в 1-седле ограничения  $f|_{\partial\Theta}$ , если  $\text{grad}f$  указывает внутрь  $\Theta$ . Когда мы говорим о минимумах и 1-седлах, это также означает такие критические точки  $f|_{\partial\Theta}$ . Аналогично множество генераторов индекса  $j$  в комплексе Морса включает критические точки индекса  $j$   $f|_{\partial\Theta}$  с  $\text{grad}f$ , указывающим внутрь  $\Theta$ . Дифференциал также модифицируется аналогичным образом, чтобы учесть траектории, включающие такие критические точки.

### В. ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 3

**Теорема.** ([7], Раздел 2) *Любой  $\mathbb{R}$ -фильтрованный цепной комплекс  $C_*$  над полем  $k$  может быть приведен линейным преобразованием, сохраняющим  $\mathbb{R}$ -фильтрацию, к “канонической форме”, канонически определенной прямой сумме неразложимых  $\mathbb{R}$ -фильтрованных комплексов двух типов:*

- 1-мерный  $\mathbb{R}$ -фильтрованный комплекс с тривиальным дифференциалом,  $\partial \tilde{e}_i^{(j)} = 0$ ,  $\langle \tilde{e}_i^{(j)} \rangle = F_{\leq r}$ ,  $r \in \mathbb{R}$ ,

- 2-мерный  $\mathbb{R}$  – фильтрованный комплекс с тривиальной гомологией  $\partial \tilde{e}_2^{(j+1)} = \tilde{e}_1^{(j)}$ ,  $\langle \tilde{e}_1^{(j)} \rangle = F_{\leq s_1}$ ,  $\langle \tilde{e}_2^{(j)}, \tilde{e}_2^{(j+1)} \rangle = F_{\leq s_2}$ ,  $s_1, s_2 \in \mathbb{R}$ .

*Полученная каноническая форма для данного фильтрованного комплекса единственна*

*Доказательство.* ([7], Раздел 2)

Пусть  $\{e_i^{(n)}\}$  – базис в векторных пространствах  $C_n$ , совместимый с фильтрацией, так что каждое подпространство  $F_r C_n$  является линейной оболочкой  $\langle e_1^{(n)}, \dots, e_r^{(n)} \rangle$ . Заметим, что фильтрация определяет естественный порядок на множестве базовых элементов.

Пусть  $\partial e_l^{(n)}$  имеет требуемую форму для  $n = j$  и  $l \leq i$ , или  $n < j$  и всех  $l$ . Т.е. либо  $\partial e_l^{(n)} = 0$ , либо  $\partial e_l^{(n)} = e_{m(l)}^{(n-1)}$ , где  $m(l) \neq m(l')$  для  $l \neq l'$ .

Пусть

$$\partial e_{i+1}^{(j)} = \sum_k e_k^{(j-1)} \alpha_k.$$

Переместим все члены с  $e_k^{(j-1)} = \partial e_q^j$ ,  $q \leq i$ , с правой стороны на левую. Мы получим

$$\partial \left( e_{i+1}^{(j)} - \sum_{q \leq i} e_q^{(j)} \alpha_{k(q)} \right) = \sum_k e_k^{(j-1)} \beta_k.$$

Если  $\beta_k = 0$  для всех  $k$ , то определим

$$\tilde{e}_{i+1}^{(j)} = e_{i+1}^{(j)} - \sum_{q \leq i} e_q^{(j)} \alpha_{k(q)},$$

так что

$$\partial \tilde{e}_{i+1}^{(j)} = 0,$$

и  $\partial e_l^{(n)}$  имеет требуемую форму для  $l \leq i+1$  и  $n = j$ , и для  $n < j$  и всех  $l$ .

В противном случае пусть  $k_0$  будет максимальным  $k$  с  $\beta_k \neq 0$ . Тогда

$$\partial(e_{i+1}^{(j)} - \sum_{q \leq i} e_q^{(j)} \alpha_{k(q)}) = e_{k_0}^{(j-1)} \beta_{k_0} + \sum_{k < k_0} e_k^{(j-1)} \beta_k,$$

$\beta_{k_0} \neq 0$ . Определим

$$\tilde{e}_{i+1}^{(j)} = \left( e_{i+1}^{(j)} - \sum_{q \leq i} e_q^{(j)} \alpha_{k(q)} \right) / \beta_{k_0},$$

$$\tilde{e}_{k_0}^{(j-1)} = e_{k_0}^{(j-1)} + \sum_{k < k_0} e_k^{(j-1)} \beta_k / \beta_{k_0}.$$

Тогда

$$\partial \tilde{e}_{i+1}^{(j)} = \tilde{e}_{k_0}^{(j-1)}$$

и для  $n = j$  и  $l \leq i + 1$ , или  $n < j$  и всех  $l$ ,  $\partial e_l^{(n)}$  имеет требуемую форму. Если комплекс был приведен к “канонической форме” на подкомплексе  $\bigoplus_{n \leq j} C_n$ , то аналогично приводим  $\partial e_l^{(j+1)}$  и так далее.

Единственность канонической формы следует, по существу, из единственности на каждом предыдущем этапе. Пусть  $\{a_i^{(j)}\}$ ,  $\left\{b_i^{(j)} = \sum_{k \leq i} a_k^{(j)} \alpha_k\right\}$  будут двумя базисами  $C_*$  для двух разных канонических форм. Предположим, что для всех индексов  $p < j$  и всех  $n$ , и  $p = j$  и  $n \leq i$  канонические формы согласуются. Пусть  $\partial a_{i+1}^{(j)} = a_m^{(j-1)}$  и  $\partial b_{i+1}^{(j)} = b_l^{(j-1)}$  с  $m > l$ ,  $a_m^{(j-1)}$  не находится в фильтрационном подпространстве, соответствующем  $b_l^{(j-1)}$ .

Следовательно,

$$\partial \left( \sum_{k \leq i+1} a_k^{(j)} \alpha_k \right) = \sum_{n \leq l} a_n^{(j-1)} \beta_n,$$

где  $\alpha_{i+1} \neq 0$ ,  $\beta_l \neq 0$ . Таким образом,

$$\partial a_{i+1}^{(j)} = \sum_{n \leq l} a_n^{(j-1)} \beta_n / \alpha_{i+1} - \sum_{k \leq i} \partial a_k^{(j)} \alpha_k / \alpha_{i+1}.$$

С другой стороны,  $\partial a_{i+1}^{(j)} = a_m^{(j-1)}$ , с  $m > l$ , и  $\partial a_k^{(j)}$  для  $k \leq i$  либо равны нулю, либо равны некоторым базовым элементам  $a_n^{(j-1)}$ , отличным от  $a_m^{(j-1)}$ . Это дает противоречие.

Аналогично, если  $\partial b_{i+1}^{(j)} = 0$ , то

$$\partial a_{i+1}^{(j)} = - \sum_{k \leq i} \partial a_k^{(j)} \alpha_k / \alpha_{i+1},$$

что снова дает противоречие по тем же аргументам. Это доказывает единственность канонической формы.  $\square$

**Замечание 1.** Бар-код  $\mathbb{R}$ -фильтрованного цепного комплекса состоит из сегментов, представляющих неразложимые  $\mathbb{R}$ -фильтрованные цепные комплексы, как мы напомним в Определении 2 в разделе 2. Существует стандартный лексикографический порядок на множестве таких сегментов. Прямая сумма из утверждения теоремы — это стандартное векторное пространство “прямой суммы по множеству”.

#### ИСТОЧНИК ФИНАНСИРОВАНИЯ

Данная работа была частично поддержана грантом РФФИ 21-51-12005 ННИО\_а.

#### СПИСОК ЛИТЕРАТУРЫ

1. Li H., Xu Z., Taylor G., Studer C., Goldstein T. Visualizing the loss landscape of neural nets, in: Advances in

Neural Information Processing Systems. 2018. P. 6389–6399.

2. Dauphin Y.N., Pascanu R., Gulcehre C., Cho K., Ganguli S., Bengio Y. Identifying and attacking the saddle point problem in high-dimensional non-convex optimization, Advances in Neural Information Processing Systems. 2014. V. 27. P. 2933–2941.

3. Choromanska A., Henaff M., Mathieu M., Arous G.B., LeCun Y. The loss surfaces of multilayer networks, JMLR Workshop and Conference Proceedings. 2015. V. 38.

4. Bott R. Lectures on Morse theory, old and new, Bulletin of the American mathematical society. 1982. V. 7 (2). P. 331–358.

5. Smale S. Differentiable dynamical systems, Bulletin of the American mathematical Society. 1967. V. 73 (6). P. 747–817.

6. Thom R. Sur une partition en cellules associée à une fonction sur une variété, Comptes Rendus de l'Académie des Sciences. 1949. V. 228 (12). P. 973–975.

7. Barannikov S. Framed Morse complexes and its invariants., Advances in Soviet Mathematics. 1994. V. 21. P. 93–116.  
<https://doi.org/10.1090/advsov/021/03>

8. Le Peutrec D., Nier F., Viterbo C. Precise Arrhenius law for p-forms: The Witten Laplacian and Morse–Barannikov complex, Annales Henri Poincaré. 2013. V. 14 (3). P. 567–610.  
<https://doi.org/10.1007/s00023-012-0193-9>

9. Le Roux F., Seyfaddini S., Viterbo C. Barcodes and area-preserving homeomorphisms, arXiv:1810.03139. 2018. P. 1–109.

10. Bergstra J., Bengio Y. Random search for hyper-parameter optimization, Journal of Machine Learning Research. 2012. V. 13. P. 281–305.

11. Chung P.B.M.K., Kim P.T. Persistence diagrams of cortical surface data, Information Processing in Medical Imaging. 2009. V. 5636. P. 386–397.

12. Sousbie T., Pichon C., Kawahara H. The persistent cosmic web and its filamentary structure – II. Illustrations, Monthly Notices of the Royal Astronomical Society. 2011. V. 414 (1). P. 384–403.  
<https://doi.org/10.1111/j.1365-2966.2011.18395.x>

13. Pun C.S., Xia K., Lee S.X. Persistent-homology-based machine learning and its applications – a survey, preprint arxiv: 1811.00252. 2018. arXiv:1811.00252.

14. Dellago C., Bolhuis P.G., Geissler P.L. Transition Path Sampling, John Wiley & Sons, Ltd, 2003. P. 1–78.  
<https://doi.org/10.1002/0471231509.ch1>

15. Oganov A.R., Valle M., How to quantify energy landscapes of solids, The Journal of Chemical Physics. 2009. V. 130 (10). P. 104504.  
<https://doi.org/10.1063/1.3079326>

16. Chazal F., Guibas L., Oudot S., Skraba P. Scalar field analysis over point cloud data, Discrete & Computational Geometry. 2011. V. 46 (4). P. 743.

17. Jamil M., Yang X.-S. A literature survey of benchmark functions for global optimization problems, International Journal of Mathematical Modelling and Numerical Optimisation. 2013. V. 4 (2). P. 150–194.

18. *Efrat A., Itai A., Katz M.J.* Geometry helps in bottleneck matching and related problems, *Algorithmica*. 2001. V. 31 (1). P. 1–28.
19. *Kawaguchi K.* Deep learning without poor local minima, in: *Advances in neural information processing systems*. 2016. P. 586–594.
20. *Gori M., Tesi A.* On the problem of local minima in backpropagation, *IEEE Transactions on Pattern Analysis & Machine Intelligence*. 1992. V. 14 (1). P. 76–86.
21. *Cao J., Wu Q., Yan Y., Wang L., Tan M.* On the flatness of loss surface for two-layered relu networks, in: *Asian Conference on Machine Learning*. 2017. P. 545–560.
22. *Yi M., Meng Q., Chen W., Ma Z.-m., Liu T.-Y.* Positively scale-invariant flatness of relu neural networks, *arXiv:1903.02237*. 2019.
23. *Chaudhari P., Choromanska A., Soatto S., LeCun Y., Baldassi C., Borgs C., Chayes J., Sagun L., Zecchina R.* Entropy-sgd: Biasing gradient descent into wide valleys, in: *International Conference on Learning Representations (ICLR)*, 2017.
24. *Dinh L., Pascanu R., Bengio S., Bengio Y.* Sharp minima can generalize for deep nets, in: *Proceedings of the 34th International Conference on Machine Learning, Proceedings of Machine Learning Research, PMLR*. 2017. P. 1019–1028.
25. *Alom M.Z., Taha T.M., Yakopcic C., Westberg S., Sidike P., Nasrin M.S., Hasan M., Van Essen B.C., Awwal A.A., Asari V.K.* A state-of-the-art survey on deep learning theory and architectures, *Electronics*. 2019. V. 8 (3). P. 292.

## BARCODES AS SUMMARY OF LOSS FUNCTION TOPOLOGY

S. Barannikov<sup>a,c</sup>, A. Korotin<sup>a,b</sup>, D. Oganessian<sup>a</sup>, D. Emtsev<sup>a,d</sup>, and E. Burnaev<sup>a,b</sup>

<sup>a</sup>*Skolkovo Institute of Science and Technology, Moscow, Russia*

<sup>b</sup>*Artificial Intelligence Research Institute, Moscow, Russia*

<sup>c</sup>*CNRS, IMJ, Paris CitBre University, Paris, France*

<sup>d</sup>*ETH Zurich, Switzerland*

Presented by Academician of the RAS A.L. Semenov

We propose to study neural networks' loss surfaces by methods of topological data analysis. We suggest to apply barcodes of Morse complexes to explore topology of loss surfaces. An algorithm for calculations of the loss function's barcodes of local minima is described. We have conducted experiments for calculating barcodes of local minima for benchmark functions and for loss surfaces of small neural networks. Our experiments confirm our two principal observations for neural networks' loss surfaces. First, the barcodes of local minima are located in a small lower part of the range of values of neural networks' loss function. Secondly, increase of the neural network's depth and width lowers the barcodes of local minima. This has some natural implications for the neural network's learning and for its generalization properties.