

Автономная некоммерческая образовательная организация высшего  
образования «Сколковский институт науки и технологий»

На правах рукописи

Коротин Александр Андреевич

**Параметрические методы вычисления оптимальных  
транспортных отображений, расстояний и барицентров**

Специальность 1.2.2 —

«Математическое моделирование, численные методы и комплексы программ»

Диссертация на соискание учёной степени  
кандидата физико-математических наук

Научный руководитель:  
доктор физико-математических наук, доцент  
Бурнаев Евгений Владимирович

Москва — 2022

## Оглавление

	Стр.
<b>Введение . . . . .</b>	<b>5</b>
<b>Глава 1. Постановка задач исследования . . . . .</b>	<b>14</b>
1.1 Вычисление Васерштейн-2 оптимального отображения . . . . .	14
1.2 Вычисление Васерштейн-2 барицентра . . . . .	17
1.3 Вычисление Васерштейн-2 градиентного потока . . . . .	18
1.4 Построение эталонных пар для Васерштейн-2 транспорта . . . . .	20
<b>Глава 2. Метод для вычисления Васерштейн-2 оптимальных транспортных отображений . . . . .</b>	<b>21</b>
2.1 Введение . . . . .	21
2.2 Обзор существующих методов . . . . .	25
2.3 Предварительные сведения . . . . .	26
2.4 Одноэтапный неминимаксный алгоритм . . . . .	28
2.4.1 Алгоритм . . . . .	28
2.4.2 Аппроксимация . . . . .	31
2.5 Доказательства теоретических результатов . . . . .	33
2.6 Вычислительные эксперименты . . . . .	39
2.6.1 Общие детали экспериментов . . . . .	39
2.6.2 Синтетические эксперименты в 2D . . . . .	40
2.6.3 Оптимальный транспорт между гауссовскими распределениями . . . . .	41
2.6.4 Передача цветов . . . . .	43
2.6.5 Адаптация домена . . . . .	44
2.6.6 Оптимальный транспорт в латентном пространстве . . . . .	48
2.6.7 Перенос стилей между изображениями . . . . .	49
2.7 Нейронные сети. Выпуклые Архитектуры . . . . .	51
2.7.1 Общая архитектура с выпуклым входом . . . . .	51
2.7.2 Полносвязная нейронная сеть, выпуклая по входу . . . . .	53
2.7.3 Сверточная выпуклая по входу нейронная сеть . . . . .	55
2.8 Обсуждение . . . . .	55
<b>Глава 3. Метод для вычисления Васерштейн-2 барицентров . . . . .</b>	<b>56</b>

3.1	Введение . . . . .	56
3.2	Предварительные сведения . . . . .	57
3.3	Обзор существующих методов . . . . .	59
3.4	Одноэтапный неминимаксный алгоритм . . . . .	61
3.4.1	Вывод двойственной задачи . . . . .	61
3.4.2	Условие конгруэнтности . . . . .	62
3.4.3	Выполнение условия сопряженности для пар потенциалов	64
3.4.4	Алгоритм. Практическая реализация. Оптимизация . . .	66
3.5	Доказательства теоретических результатов . . . . .	68
3.6	Вычислительные эксперименты . . . . .	72
3.6.1	Линейно-разбросанные семейства в большой размерности	74
3.6.2	Агрегация апостериорных распределений на подмножествах	76
3.6.3	Усреднение цветовой палитры . . . . .	77
3.6.4	Технические детали экспериментов . . . . .	78
<b>Глава 4. Метод для моделирования Васерштейн-2 градиентных потоков функционалов . . . . .</b>		<b>80</b>
4.1	Введение . . . . .	80
4.2	Основные сведения о градиентных потоках Васерштейна . . . .	83
4.3	Вычисление градиентных потоков Васерштейна с использованием сетей ICNN . . . . .	86
4.3.1	Переформулировка JKO-схемы с использованием оптимальных операторов отображения . . . . .	86
4.3.2	Стохастическая оптимизация для JKO-схемы с использованием сетей ICNN . . . . .	87
4.3.3	Вычисление плотности диффузационного процесса . . . . .	89
4.4	Вычислительные эксперименты . . . . .	90
4.4.1	Сходимость к стационарному решению . . . . .	91
4.4.2	Моделирование процесса Орнштейна–Уленбека . . . . .	92
4.4.3	Выборка из ненормированных апостериорных распределений в байесовской логистической регрессии . .	93
4.5	Технические детали экспериментов . . . . .	94
4.5.1	Сходимость к стационарному распределению . . . . .	96
4.5.2	Моделирование процессов процесса Орнштейна–Уленбека	96

4.5.3	Получение выборки из апостериорного распределения . . . . .	97
4.6	Обсуждение . . . . .	97
<b>Глава 5. Методология построения Васерштейн-2 эталонных пар вероятностных распределений</b>		<b>102</b>
5.1	Введение . . . . .	102
5.2	Некоторые сведения из теории оптимального транспорта . . . . .	104
5.3	Параметрические двойственные решатели для задач оптимального транспорта с квадратичной функцией стоимости .	108
5.4	Тестирование решателей для задачи оптимального транспорта .	111
5.5	Детали сравнительных тестов и полученные результаты . . . . .	112
5.5.1	Наборы данных . . . . .	112
5.5.2	Метрики и базовые решатели . . . . .	115
5.5.3	Оценка решателей на многомерных эталонных парах . . . . .	115
5.5.4	Оценка решателей на эталонных парах изображений размера $64 \times 64$ из набора CelebA . . . . .	117
5.5.5	Тестирование решателей в генеративном моделировании изображений размера $64 \times 64$ . . . . .	119
5.6	Технические детали построения эталонных пар . . . . .	121
5.6.1	Эталонные пары большой размерности . . . . .	122
5.6.2	Эталонные пары изображений . . . . .	123
5.7	Технические детали экспериментов . . . . .	123
5.7.1	Архитектуры нейронной сети . . . . .	124
5.7.2	Гиперпараметры и детали имплементации . . . . .	126
5.8	Обсуждение . . . . .	128
<b>Заключение</b>		<b>130</b>
<b>Публикации автора по теме диссертации</b>		<b>131</b>
<b>Список литературы</b>		<b>132</b>
<b>Список рисунков</b>		<b>144</b>
<b>Список таблиц</b>		<b>147</b>

## Введение

Основной целью генеративного моделирования является построение математических моделей, описывающих сложную структуру многомерных данных, и дальнейшее использование этих моделей для симуляции новых данных со схожими статистиками или для изменения характеристик исходных данных [1]. За последние десять лет в области генеративного моделирования широкое распространение получили численные методы на основе математической теории **оптимального транспорта** [1; 2], одним из основоположников которой является советский математик и экономист Леонид Витальевич Канторович.

Оптимальный транспорт (ОТ) — это общее название обширного класса задач нахождения эффективного способа перемещения массы между вероятностными распределениями. В наши дни численные методы ОТ преимущественно используются для вычисления функций потерь при обучении генеративных моделей на основе глубоких нейросетей для синтеза искусственных данных [1], например, изображений. Другое важное применение численных методов ОТ — доменная адаптация [3], то есть адаптация модели (например, классификационной) для применения к данным с нового источника. Эта задача крайне важна, например, в медицине, где типичный размер выборок данных не превышает сотни объектов и необходимо комбинировать знания, полученные с нескольких выборок из разных источников. Для решения таких задач используются ОТ отображения, а также центры масс (ОТ барицентры) для агрегации выборок [4].

Наиболее популярна задача ОТ с квадратичной функцией стоимости, определяющая известную Васерштейн-2 метрику. Градиентные потоки функционалов на пространстве распределений, оснащенном этой метрикой, позволяют моделировать решения ряда дифференциальных уравнений, например, уравнения Фоккера-Планка, используемого для описания процессов в физике, финансах, демографии. Моделирование градиентных потоков можно осуществлять, используя численные методы ОТ [5].

Несмотря на значительные достижения в вычислительном ОТ за последнее десятилетие, существующие вычислительные методы решения задач ОТ по-прежнему обладают рядом критичных недостатков (описываемых далее), ограничивающих их эффективное применение в ряде вышеупомянутых задач моделирования распределений данных.

Большинство методов [6] в области вычислительного ОТ предназначены для данных, описываемых **дискретными** вероятностными распределениями. Такие методы способны аппроксимировать решения задачи ОТ для распределений с дискретным носителем из  $10^5 - 10^6$  точек. При больших количествах точек проблема неразрешима из-за большой вычислительной сложности. Таким образом, дискретные методы не могут обеспечить точную аппроксимацию решения задачи ОТ для данных высокой размерности, имеющих непрерывные распределения, поскольку требуются большие размеры выборок.

В то же время машинное обучение все чаще применяется для решения задач ОТ [2; 7]. Построение математических моделей на основе глубоких нейронных сетей для параметризации транспортных отображений позволяет масштабировать применения ОТ на высокоразмерные пространства и практические задачи, в которых дискретные методы работают неудовлетворительно [8]. В частности, численные методы ОТ на основе глубоких нейронных сетей позволяют получать решения задач ОТ для **непрерывных** распределений. Такие методы обычно называются непрерывными или **параметрическими**, а процесс вычисления ОТ – обучением модели.

Численные непрерывные методы ОТ условно делятся на две группы: методы на основе энтропийной регуляризации [2] и минимаксные методы [7]. Энтропийные методы из-за регуляризации находят решение задачи ОТ, которое смещено относительно истинного решения, что ведет к большой ошибке решения задачи ОТ и усложняет использование методов на практике. Минимаксные непрерывные методы не имеют этой проблемы, но требуют решения задачи поиска седловой точки некоторого функционала. На практике такая оптимизация может быть неустойчива, что усложняет процесс обучения с использованием этих методов.

Важно отметить, что хотя в области вычислительного ОТ активно ведется разработка новых непрерывных методов, вопрос оценки качества их функционирования все еще остается открытым. Связано это с тем, что существует лишь ограниченное число непрерывных задач ОТ, в которых аналитически известно истинное решение задачи, с которым можно сравнить решение вычисленное непрерывным методом.

*Таким образом, существует необходимость разработки новых численных методов, которые позволили бы находить решения непрерывных задач ОТ в больших размерностях без смещения и без решения минимаксной задачи*

*оптимизации. Для оценки точности таких методов требуется разработка количественной методики тестирования непрерывных методов вычисления ОТ.*

Основными математическими задачами является аналитический вывод двойственных формулировок задач ОТ, позволяющих осуществлять вычисление транспортных отображений с помощью нейросетей без необходимости использовать энтропийную регуляризацию или решать минимаксную задачу. Важной подзадачей является получение теоретических результатов, связывающих точность нахождения оптимума двойственного функционала (в смысле его значения) с ошибкой восстановленного решения прямой задачи ОТ. Построение надежных и теоретически обоснованных нейросетевых методов вычисления ОТ позволит усовершенствовать применения ОТ к ранее упомянутым практическим задачам.

**Целью** диссертационной работы является разработка математических моделей на основе нейросетей для решения задач непрерывного ОТ с квадратичной функцией стоимости (называемой Васерштейн-2 метрикой): вычисление ОТ отображения и расстояния между непрерывными вероятностными распределениями, нахождение барицентров семейств вероятностных распределений и вычисление градиентных потоков функционалов на пространстве вероятностных распределений. Для достижения цели в работе были поставлены и решены следующие **задачи**:

1. Разработать численный метод на основе двойственной формулировки ОТ для поиска Васерштейн-2 оптимальных отображений между непрерывными распределениями с помощью нейросетей; получить теоретические оценки того, как связаны ошибки аппроксимации решений предложенной двойственной и исходной задач.
2. Разработать численный метод на основе двойственной формулировки ОТ для поиска Васерштейн-2 барицентров семейства непрерывных распределений с помощью нейронных сетей; получить теоретические оценки того, как связаны ошибки аппроксимации решений предложенной двойственной и исходной задач.
3. Разработать численный метод на основе дискретизации времени и нейронных сетей для моделирования Васерштейн-2 градиентных потоков функционалов на пространстве распределений.
4. Разработать методологию для количественного сравнения непрерывных методов ОТ; предложить и математически обосновать метод создания

эталонных пар непрерывных распределений с аналитически известным Васерштейн-2 ОТ отображением.

### **Научная новизна:**

1. Впервые предложен масштабируемый метод вычисления Васерштейн-2 оптимальных отображений между непрерывными распределениями в пространствах большой размерности, который состоит из одного этапа обучения и не требует минимаксной оптимизации. В основе метода лежит новый подход на основе введения нового циклического регуляризатора в двойственную постановку задачи ОТ и использования нейросетей с выпуклой по входу архитектурой. В отличие от существующих регуляризаторов циклический регуляризатор не вносит смещение в решение задачи.
2. Впервые предложен масштабируемый метод вычисления Васерштейн-2 барицентров непрерывных вероятностных распределений, который состоит из одного этапа обучения и не требует минимаксной оптимизации. В основе подхода лежит комбинация циклического регуляризатора (см. выше), нового регуляризатора на конгруэнтность решений двойственной задачи и выпуклых по входу нейронных сетей. В отличие от существующих регуляризаторов, предложенная комбинация не вносит смещение в решение.
3. Впервые предложен масштабируемый численный метод для моделирования Васерштейн-2 градиентных потоков на пространстве высокоразмерных вероятностных распределений. Метод является практической реализацией теоретической схемы JKO [5], для которой ранее не были известны эффективные численные методы для её реализации в пространствах высокой размерности.
4. Предложена оригинальная методология синтеза пар многомерных непрерывных вероятностных распределений, для которых аналитически известно эталонное Васерштейн-2 транспортное отображение между ними. Разработанная методология и предложенные пары распределений позволяют сравнивать непрерывные численные методы ОТ в больших размерностях и устраняют недостатки существующих методик, связанные с узостью классов используемых тестовых распределений и не позволяющие объективно оценивать качество численных методов ОТ.

**Теоретическая значимость** работы заключается в том, что предложенные новые подходы на основе нейронных сетей и численных методов ОТ позволяют эффективно решать ряд задач математического моделирования, избавляясь от недостатков существующих подходов таких как использование численно нестабильной минимаксной оптимизации или наличие систематических ошибок в решении. Принципиальной особенностью предложенных подходов является то, что разработанные в работе решения для не требуют минимаксной оптимизации и не смещают оптимальное решение задачи, что приводит к более вычислительно эффективному и точному нахождению решения по сравнению с существующими численными методами ОТ на основе нейронных сетей. В частности, эти улучшения значительно упрощают теоретический анализ полученных решений, и, как следствие, позволяют выводить оценки ошибки оптимизации по ограниченному классу функций, например, представленному нейросетями.

**Практическая значимость.** Предложенные подходы математического моделирования на основе нейронных сетей и численных методов ОТ имеют потенциал для решения практически важных научно-исследовательских и промышленных задач. Например, разработанный метод, протестированный в задачах доменной адаптации и переноса стиля на искусственных и модельных данных, имеет ряд потенциальных приложений в задачах компьютерного зрения, таких как обучение моделей для улучшения разрешения изображений, устранения шума и пр. Другим возможным применением разработанных подходов является агрегация данных из нескольких источников, например, медицинских сканеров разных производителей, что позволит улучшить качество моделей классификации и сегментации медицинских изображений, построенных на них.

**Методология и методы исследования.** Для достижения поставленных в диссертационной работе целей используются:

1. Подходы теории оптимального транспорта для аналитического вывода двойственных формулировок задач ОТ и соотношений между прямыми/двойственными переменными, а также для построения методов регуляризации двойственных задач, позволяющих эффективно находить их решение.
2. Численные методы стохастической оптимизации для построения практически эффективных алгоритмов нахождения решений рассматриваемых задач ОТ с помощью нейронных сетей.

3. Подходы глубинного обучения для построения нейронных сетей с выпуклой по входу архитектурой для аппроксимации минимизаторов разработанных целевых функций.
4. Методы выпуклого анализа, линейной алгебры, математического и функционального анализа, теории вероятностей и теории меры для теоретического обоснования предложенных целевых функций оптимизации, анализа соотношений решений прямых и предложенных регуляризованных двойственных задач.
5. Способы количественной и качественной оценки результатов экспериментов, используемых в области вычислительного оптимального транспорта и генеративного машинного обучения.
6. Подходы объектно-ориентированного программирования с использованием языка программирования Python и фреймворка Pytorch глубинного обучения и численной оптимизации для практической реализации и тестирования предложенных методов.

**Основные положения, выносимые на защиту:**

1. Предложен алгоритм для вычисления Васерштейн-2 оптимального отображения и расстояния между непрерывными распределениями, базирующийся на нейросетях с выпуклой по входу архитектурой и применении разработанной циклической регуляризации в двойственной задаче. Получены теоретические оценки, связывающие ошибку минимизации разработанного целевого функционала с ошибкой решения исходной задачи нахождения ОТ отображения.
2. Предложен алгоритм для вычисления Васерштейн-2 барицентра семейства непрерывных распределений, базирующийся на нейросетях с выпуклой по входу архитектурой и использовании разработанных циклической регуляризации и регуляризации на конгруэнтность двойственных переменных задачи поиска барицентра. Получены теоретические оценки, связывающие ошибку минимизации разработанного целевого функционала с ошибкой решения исходной задачи нахождения барицентра.
3. Предложен численный метод моделирования Васерштейн-2 градиентных потоков функционалов на пространствах вероятностных распределений, основанный на дискретизации градиентного потока по времени и на использовании выпуклых по входу нейросетей.

4. Разработана методология на основе нейронных сетей с выпуклой по входу архитектурой для синтеза пар непрерывных распределений с аналитически известным ОТ отображением между ними. Данная методология позволяет устранить имеющийся пробел в тестировании непрерывных методов решения задачи ОТ и исследования математических моделей, на которых они основаны.

**Полученные результаты соответствуют следующим пунктам паспорта специальности 1.2.2** (математическое моделирование, численные методы и комплексы программ):

1. Результаты 1-3 соответствуют **п.1** «Разработка новых математических методов моделирования объектов и явлений», **п.3** «Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий» и **п.4** «Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительно-го эксперимента».
2. Результат 4 соотв. **п.2** «Развитие качественных и приближенных аналитических методов исследования математических моделей».

**Достоверность** полученных результатов обеспечивается корректностью применения апробированного в научной практике математического аппарата теории оптимального транспорта и других смежных теоретических областей, а также экспериментальной проверкой разработанных численных методов на большом количестве модельных и практических задач оптимального транспорта. Полученные теоретические результаты обосновываются математически строгими доказательствами, а для проведенных вычислительных экспериментов даются детальные описания, обеспечивающие их воспроизводимость. Результаты опубликованы в трудах ведущих рецензируемых международных конференций по машинному обучению и искусственному интеллекту.

**Публикации.** Список публикаций приведен в конце диссертации. Результаты диссертации изложены в **5** работах [A1-A5]. Все работы опубликованы в трудах **Core rank A\*** ведущих международных конференций по машинному обучению. Работы [A1, A2] индексируются системой **Scopus**.

**Апробация работы.** Результаты работы доложены на **3** ведущих международных научных конференциях ранга **A\*** (суммарно **5** докладов) по машинному обучению и искусственному интеллекту:

1. The 9th International Conference on Learning Representations (ICLR 2021, **Core A\***, онлайн) – **2** презентации по работам [A3, A4];
2. The 35th Conference on Neural Information Processing Systems (NeurIPS 2021, **Core A\***, онлайн) – **2** презентации по работам [A1, A2];
3. The 10th International Conference on Learning Representations (ICLR 2022, **Core A\***, онлайн) – **1** презентация по работам [A5];

Результаты работы были представлены на **5** следующих научных семинарах по машинному обучению и искусственному интеллекту:

- Huawei Machine Learning Workshop (2019, Сочи, Россия);
- Семинар по байесовским методам (2020, НИУ ВШЭ, онлайн);
- Geometry Data Processing Group Seminar (2020, MIT, онлайн);
- SMILES Machine Learning Summer School (2020, онлайн);
- Math of Machine Learning Summer School (2020, Сочи, Россия).

За научные результаты, включая результаты диссертации, автор в 2019 и 2021 годах становился лауреатом **премии** им. Ильи Сегаловича для поддержки молодых исследователей в области машинного обучения и искусственного интеллекта от компании **Яндекс**.

**Личный вклад.** Все положения, выносимые на защиту, получены автором лично. Все алгоритмы и теоретические результаты, приведенные в диссертации и работах [A1-A5], сформулированы и обоснованы непосредственно автором диссертации. В диссертации использованы результаты совместных исследований: экспериментальное тестирование метода для нахождения ОТ отображения [A3] проведено совместно с соавторами В. Егиазаряном, А. Асадулаевым и А. Сафина; тестирование метода для вычисления барицентров [A4] проведено совместно с соавтором Л. Ли; автор диссертации осуществил предварительную реализацию и тестирование разработанного метода для вычисления градиентных потоков [A2], адаптация кода для рассматриваемых в экспериментальной части практических задач сравнения с альтернативными подходами осуществлена соавтором П. Мокровым; все вычислительные эксперименты в работе [A1] осуществлены непосредственно автором диссертации, а в работе [A5] – соавтором Л. Роут.

**Примечание.** Диссертационная работа выполнена автором в автономной некоммерческой образовательной организации высшего образования «Сколковский институт науки и технологий».

**Объем и структура работы.** Диссертация состоит из введения, 5 глав и заключения. Полный объём диссертации составляет 148 страниц, включая 30 рисунков и 17 таблиц. Список литературы содержит 138 наименований.

## Глава 1. Постановка задач исследований

В данной главе приводится общее описание 4 основных задач, решаемых в диссертационной работе: вычисление Васерштейн-2 оптимальных транспортных отображений и расстояний (**задача 1**), вычисление Васерштейн-2 барицентров (**задача 2**), вычисление Васерштейн-2 градиентных потоков функционалов (**задача 3**) и построение эталонных пар вероятностных распределений для Васерштейн-2 оптимального транспорта (**задача 4**). Ниже мы приводим математические постановки вышеописанных задач, а также кратко приводим недостатки существующих подходов к их решению. Детальное обсуждение приводится в последующих главах работы.

Далее, через  $(\mathbb{R}^D, \|\cdot\|)$  обозначим  $D$ -мерное евклидово пространство с евклидовой метрикой  $\|\cdot\|$ . Через  $\mathcal{P}_2(\mathbb{R}^D)$  обозначим множество всех борелевских вероятностных распределений на  $\mathbb{R}^D$  с конечным вторым моментом. Также через  $\mathcal{P}_{2,\text{ac}}(\mathbb{R}^D) \subset \mathcal{P}_2(\mathbb{R}^D)$  обозначим подмножество всех абсолютно непрерывных распределений (относительно меры Лебега). Образ распределения  $\mathbb{P} \in \mathcal{P}_2(\mathbb{R}^D)$  под действием измеримого отображения  $T : \mathbb{R}^D \rightarrow \mathbb{R}^D$  обозначим через  $T\sharp\mathbb{P}$ .

### 1.1 Вычисление Васерштейн-2 оптимального отображения

**Задача 1** (Вычисление оптимального транспорта с квадратичной ценой). Для пары вероятностных распределений  $\mathbb{P}, \mathbb{Q} \in \mathcal{P}_2(\mathbb{R}^D)$  задача оптимального транспорта с квадратичной ставится следующим образом:

$$\mathbb{W}_2^2(\mathbb{P}, \mathbb{Q}) \stackrel{\text{def}}{=} \min_{\pi \in \Pi(\mathbb{P}, \mathbb{Q})} \int_{\mathbb{R}^D \times \mathbb{R}^D} \frac{\|x - y\|^2}{2} d\pi(x, y), \quad (1.1)$$

где  $\Pi(\mathbb{P}, \mathbb{Q})$  – множество транспортных планов, то есть вероятностных распределений на  $\mathbb{R}^D \times \mathbb{R}^D$ , чьи маргиналы суть  $\mathbb{P}, \mathbb{Q}$  соответственно.

Решение задачи (1.1) предполагает нахождение оптимального плана  $\pi^* \in \Pi(\mathbb{P}, \mathbb{Q})$ , доставляющего минимум. Соответствующие минимальное значение называется оптимальной транспортной стоимостью, а её корень называется расстоянием Васерштейн-2 и обозначается через  $\mathbb{W}_2$ .

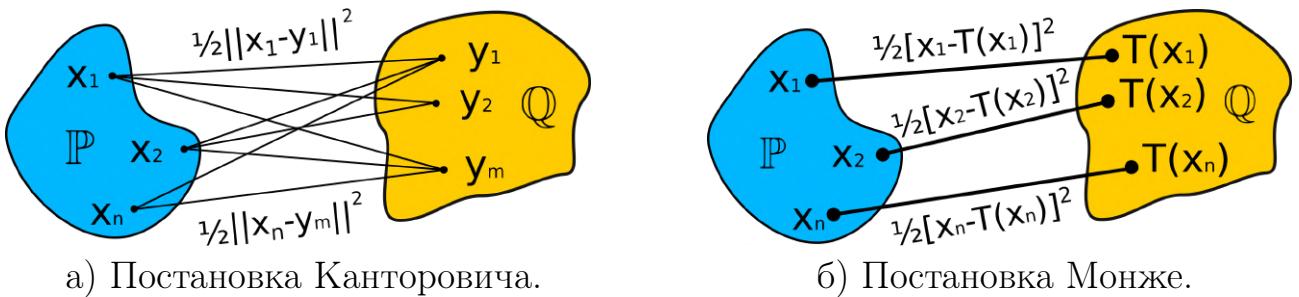


Рисунок 1.1 — Схематически изображенная задача 1 оптимального транспорта между распределениями  $\mathbb{P}, \mathbb{Q}$  с квадратичной функцией стоимости.

Формулировка (1.1) задачи оптимального транспорта часто называется формулировкой **Канторовича** задачи ОТ (Рис. 1.1а) и представляет собой оптимизацию линейного функционала от  $\pi$  (интеграла по  $\pi$ ) на пространстве вероятностных мер на  $\mathbb{R}^D \times \mathbb{R}^D$  в линейных ограничениях  $\pi \in \Pi(\mathbb{P}, \mathbb{Q})$ . Однако аналитические решения (1.1) для пар  $\mathbb{P}, \mathbb{Q}$  общего вида неизвестны, а решать данную задачу оптимизации на практике вычислительно сложно.

В случае, когда  $\mathbb{P}, \mathbb{Q}$  – **дискретные** вероятностные распределения с конечным носителем, задача обычно решается численными методами линейного программирования или же используются дополнительные регуляризаторы, приводящие к итеративным матричным методам [6]. К сожалению, когда  $\mathbb{P}$  и  $\mathbb{Q}$  **непрерывные** распределения (доступные через конечные или бесконечные случайные выборки), вышеупомянутые дискретные методы не дают достаточного качества, так как необходимые точные дискретизации распределений вычислительно недостижимы.

Далее мы кратко рассмотрим основные существующие методы решения задачи (1.1) в случае непрерывных  $\mathbb{P}, \mathbb{Q}$  и обозначим их ключевые недостатки, которые устраняются в данной диссертации. Для начала отметим, что если  $\mathbb{P} \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$ , то ОТ план  $\pi$  единственным образом определяется в виде  $\pi^* = [\text{id}_{\mathbb{R}^D}, T^*] \sharp \mathbb{P}$ , где  $T^* : \mathbb{R}^D \rightarrow \mathbb{R}^D$  – единственный минимизатор задачи ОТ в форме **Монжа** (Рис. 1.1б) [9]:

$$\mathbb{W}_2^2(\mathbb{P}, \mathbb{Q}) = \min_{T \sharp \mathbb{P} = \mathbb{Q}} \int_{\mathbb{R}^D} \frac{\|x - T(x)\|^2}{2} d\mathbb{P}(x), \quad (1.2)$$

где минимум берется по измеримым отображениям  $T : \mathbb{R}^D \rightarrow \mathbb{R}^D$ , удовлетворяющим условию переноса распределения  $T \sharp \mathbb{P} = \mathbb{Q}$ . Таким образом, в непрерывном случае задача (1.1) сводится к *нахождению оптимального транспортного отображения*  $T^* : \mathbb{R}^D \rightarrow \mathbb{R}^D$ .

Существующие непрерывные методы основываются на решении двойственной задачи к (1.1) и (1.2) и дальнейшем восстановлении оптимального транспортного отображения  $T^*$  (решения прямой задачи) из найденных двойственных переменных. Для  $\mathbb{P}, \mathbb{Q} \in \mathcal{P}_2(\mathbb{R}^D)$  **двойственная формулировка** задачи квадратичного ОТ имеет следующий вид (см. [9]):

$$\mathbb{W}_2^2(\mathbb{P}, \mathbb{Q}) = \max_{f \oplus g \leq \frac{1}{2}\|\cdot\|^2} \left[ \int_{\mathbb{R}^D} f(x)d\mathbb{P}(x) + \int_{\mathbb{R}^D} g(y)d\mathbb{Q}(y) \right], \quad (1.3)$$

где максимум берется по всем  $\mathbb{P}$ -интегрируемым функциям  $f : \mathbb{R}^D \rightarrow \mathbb{R} \cup \{\infty\}$  и  $\mathbb{Q}$ -интегрируемым  $g : \mathbb{R}^D \rightarrow \mathbb{R} \cup \{\infty\}$  (потенциалам), для которых выполнено  $f(x) + g(y) \leq \frac{1}{2}\|x - y\|^2$  при всех  $x, y \in \mathbb{R}^D$ . Существуют альтернативные к (1.3) двойственные формы, например,

$$\mathbb{W}_2^2(\mathbb{P}, \mathbb{Q}) = \text{Const}(\mathbb{P}, \mathbb{Q}) - \min_{\psi \in \text{Convex}} \overbrace{\left[ \int_{\mathbb{R}^D} \psi(x)d\mathbb{P}(x) + \int_{\mathbb{R}^D} \bar{\psi}(y)d\mathbb{Q}(y) \right]}^{\text{Corr}(\mathbb{P}, \mathbb{Q}|\psi)}, \quad (1.4)$$

где минимум берется по выпуклым потенциалам  $\psi$ . При этом  $\bar{\psi}(y) \stackrel{\text{def}}{=} \max_{x \in \mathbb{R}^D} [\langle x, y \rangle - \psi(x)]$  – сопряженная функция к  $\psi$ , которая тоже является выпуклой функцией. Второе слагаемое  $\text{Corr}(\mathbb{P}, \mathbb{Q}|\psi)$  в (1.4) называется корреляцией, а функции  $\psi, \bar{\psi}$  – потенциалами Бренье.

В случае, когда  $\mathbb{P} \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$ ,  $\mathbb{P}$ -почти всюду выполнено  $T^*(x) = \nabla\psi^*(x)$ , то есть градиент оптимального потенциала Бренье является оптимальным транспортным отображением из  $\mathbb{P}$  в  $\mathbb{Q}$ . Этот факт позволяет для решения задач ОТ (1.1), (1.2) фокусироваться лишь на решении двойственной задачи, которая представляется более простой, так в задаче устраняется нетривиальное граниченое условие  $\pi \in \Pi(\mathbb{P}, \mathbb{Q})$  в (1.1) или  $T^*\sharp\mathbb{P} = \mathbb{Q}$  в (1.2). Однако решение (1.4), то есть нахождение двойственных потенциалов, все еще нетривиально, т.к. задача представляет собой оптимизацию по пространству *выпуклых* функций  $\psi$ . При этом вычисление сопряженной функции  $\bar{\psi}$  представляет собой внутреннюю задачу оптимизации.

В работе [7] авторы предлагают аппроксимировать  $\psi^*, \bar{\psi}^*$  с помощью *выпуклых по входу нейронных сетей* (англ. input-convex neural networks, сокращенно – ICNNs) и свести задачу к оптимизации параметров нейронных сетей. Однако для использования метода предполагается решение минимаксной задачи поиска седловой точки некоторого функционала. Полученная задача

представляется сложной на практике из-за численной нестабильности минимаксной оптимизации.

В работе [2] авторы рассматривают классическую двойственную форму (1.3) вместо (1.4) и ослабляют условие  $f \oplus g \leq c$  с помощью наложения регуляризации (энтропийной и пр.), которая мягко штрафует потенциалы  $f, g$  за нарушение этого условия. Из-за регуляризации решение задачи (оптимальный  $f^*$ ) смещается относительно истинного решения, что не позволяет из него точно находить решение исходной задачи ОТ (1.2).

В данной диссертации **предлагается численный метод** (глава 2 работы), который устраняет недостатки вышеописанных непрерывных методов, а именно: **(1)** оптимизируется неманимаксный критерий, **(2)** решением которого является истинный несмешенный потенциал  $\Psi^*$ , градиент  $\nabla\Psi^*$  которого является искомым ОТ отображением  $T^*$ .

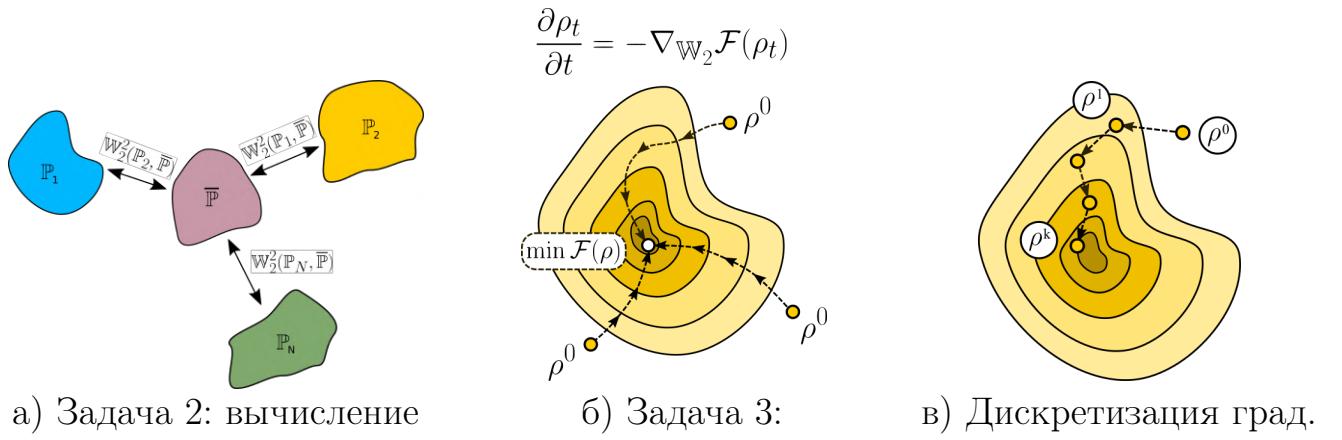
## 1.2 Вычисление Васерштейн-2 барицентра

**Задача 2** (Задача поиска Васерштейн-2 барицентра). *Рассмотрим распределения  $\mathbb{P}_1, \dots, \mathbb{P}_N \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$ . Тогда их барицентр  $\bar{\mathbb{P}}$  относительно весов  $\alpha_1, \dots, \alpha_N$  (при  $\alpha_n > 0$  и  $\sum_{n=1}^N \alpha_n = 1$ ) определяется через (рис. 1.2a)*

$$\bar{\mathbb{P}} \stackrel{\text{def}}{=} \arg \min_{\mathbb{P} \in \mathcal{P}_2(\mathbb{R}^D)} \mathcal{F}_{\text{Bar}}(\mathbb{P}) \stackrel{\text{def}}{=} \arg \min_{\mathbb{P} \in \mathcal{P}_2(\mathbb{R}^D)} \sum_{n=1}^N \alpha_n \mathbb{W}_2^2(\mathbb{P}_n, \mathbb{P}). \quad (1.5)$$

Здесь  $\mathcal{F}_{\text{Bar}}$  называется *барицентрическим функционалом*, и его минимальное значение обозначается через  $\mathcal{F}_{\text{Bar}}^* \stackrel{\text{def}}{=} \mathcal{F}_{\text{Bar}}(\bar{\mathbb{P}})$ .

Проблемы существующих методов вычисления барицентров аналогичны проблемам методов решения задачи ОТ (1.2). В случае, когда  $\mathbb{P}_1, \dots, \mathbb{P}_n$  являются многомерными непрерывными распределениями, дискретные методы не дают достаточного качества, так как необходимые точные дискретизации распределений вычислительно недостижимы. Непрерывные же методы для вычисления барицентров либо требуют решения сложной  $\min \max \min$  задачи оптимизации [10], либо используют энтропийную регуляризацию [8], ведущую к смещению решения.



а) Задача 2: вычисление барицентра  $\bar{P}$  семейства  $\{\mathbb{P}_n\}$ .

б) Задача 3: моделирование град. потока  $\rho_t$  функц.  $\mathcal{F}(\rho)$ .

в) Дискретизация град. потока по времени (схема JKO).

Рисунок 1.2 — Схематически изображенные задачи 2, 3 и схема JKO [5].

В данной диссертации **предлагается численный метод** (глава 3 работы), который лишен недостатков вышеописанных методов, а именно: **(1)** оптимизируется неманиакский критерий, **(2)** решение которого позволяет восстановить истинный несмещенный барицентр.

### 1.3 Вычисление Васерштейн-2 градиентного потока

**Задача 3** (Вычисление Васерштейн-2 градиентного потока функционала на пространстве распределений). *Рассмотрим функционал на пространстве вероятностных распределений  $\mathcal{F} : \mathcal{P}_2(\mathbb{R}^D) \rightarrow \mathbb{R}$ . Кривая, состоящая из распределений  $\{\rho_t\}_{t \in \mathbb{R}_+}$  и являющаяся решением уравнения*

$$\frac{\partial \rho_t}{\partial t} = \operatorname{div}(\rho_t \nabla_x \mathcal{F}'(\rho_t)), \quad s.t. \rho_0 = \rho^0, \quad (1.6)$$

называется *Васерштейн-2 градиентным потоком функционала  $\mathcal{F}$ , стартующим из точки  $\rho^0 \in \mathcal{P}_2(\mathbb{R}^D)$* , см. рис. 1.2б. Здесь  $\mathcal{F}'(\rho_t) : \mathbb{R}^D \rightarrow \mathbb{R}$  обозначает первую вариацию функционала  $\mathcal{F}$  в точке  $\rho_t$  [9] (см. определение в п. 4.2 далее).

Научно-практический интерес в вычислении градиентных потоков заключается в том, что многие диффузионные процессы, возникающие на практике, являются Васерштейн-2 градиентными потоками определенных функционалов. В данной диссертационной работе основной объект исследования представляют диффузии, являющиеся градиентными потоками функционала  $\mathcal{F}_{\text{FP}}$  свободной

энергии Фоккера-Планка ( $\beta > 0$ )

$$\mathcal{F}_{\text{FP}}(\rho) \stackrel{\text{def}}{=} \int_{\mathbb{R}^D} \Phi(x) d\rho(x) + \beta^{-1} \int_{\mathbb{R}^D} \log \frac{d\rho}{dx}(x) d\rho(x), \quad (1.7)$$

являющегося разностью потенциальной энергии и энтропии. Соответствующий градиентный поток удовлетворяет  $\frac{\partial \rho_t}{\partial t} = \text{div}(\nabla \Phi(x) \rho_t) + \beta^{-1} \Delta \rho_t$  при  $\rho_0 = \rho^0$ . Отметим, что при  $\beta = \infty$  выражение является классическим уравнением переноса, а в случае  $\Phi(x) \equiv 0$  – уравнением теплопроводности.

В общем случае задача вычисления градиентных потоков Васерштейна-2 является сложной. Обычно аналитическое решение получить не удается, и поэтому приходится применять методы численной аппроксимации. Джордан, Киндерлерер и Отто предложили подход (позже получивший сокращенное название “JKO-интегрирование”) для аппроксимации динамики  $\rho_t$  в (1.6), см. [5]. В этом методе осуществляется дискретное по времени итеративное уточнение непрерывного потока (рис. 1.2в):

$$\rho^{(k)} \leftarrow \arg \min_{\rho \in \mathcal{P}_2(\mathbb{R}^n)} \left[ \mathcal{F}(\rho) + \frac{1}{h} \mathbb{W}_2^2(\rho^{(k-1)}, \rho) \right], \quad (1.8)$$

где  $\rho^{(0)} = \rho^0$  – начальное условие и  $h > 0$  – величина шага дискретизации по времени. Дискретный временной градиентный поток сходится к непрерывному потоку при  $h \rightarrow 0$ , т.е.  $\rho^{(k)} \approx \rho_{kh}$ .

Стоит отметить, что выполнение JKO-итераций является сложной задачей из-за наличия минимизации по  $\mathbb{W}_2^2$ . Типичный подход к выполнению JKO-шага основан на дискретизации пространственного домена. При размере носителя  $\lesssim 10^6$  задача (1.8) может быть решена стандартными дискретными алгоритмами ОТ [6], упомянутыми в предыдущих разделах. В размерностях  $D \geq 3$  дискретные распределения обычно неточно аппроксимируют непрерывные распределения и, как следствие, динамику градиентных потоков. Чтобы обойти эту трудность, в [11] предложен параметрический метод для аппроксимации  $\rho_t$  на основе энтропийного ОТ, который приводит к смещению найденного решения.

В данной диссертации **предлагается численный метод** (глава 4 работы) для моделирования несмещенной динамики JKO с помощью замены оптимизации (1.8) по вероятностным распределениям на эквивалентную, но практически более обозримую оптимизацию по выпуклым функциям.

## 1.4 Построение эталонных пар для Васерштейн-2 транспорта

**Задача 4** (Построение эталонных пар непрерывных распределений для Васерштейн-2 оптимального транспорта). *Построить пары непрерывных вероятностных распределений  $(\mathbb{P}, \mathbb{Q})$ , для которых оптимальное решение  $T^*$  задачи 1 (ОТ отображение) аналитически известно. Подразумевается, что построенные распределения  $\mathbb{P}, \mathbb{Q}$  на практике должны быть доступны через случайные независимые выборки.*

Эталонные пары могут быть использованы для количественного тестирования численных методов для решения задачи 1 поиска ОТ плана или отображения. На данный момент возможности такого тестирования значительно ограничены. Этот пробел в научной области вычислительного ОТ не позволяет объективно оценить качество существующих численных методов для решения задач ОТ. В литературе для оценки качества существующих непрерывных методов используются два типа метрик.

В явных метриках вычисленное транспортное отображение  $\hat{T}$  сравнивается с истинным  $T^*$ . В настоящее время имеется сравнительно немного явных метрик, поскольку число примеров  $\mathbb{P}, \mathbb{Q}$  с известным эталоном  $T^*$  (т.е. решений задачи 4) достаточно мало:  $T^*$  может быть получено в дискретном, 1-мерном и линейно-разбросанном случаях.

Неявные метрики применяют численный метод ОТ как компоненту в многоэтапной схеме обучения. При этом качество конечного решения задачи используется как косвенная оценка качества самого численного метода. Например, в генеративном моделировании, при котором ОТ используется как функция потери нейронной сети генератора, качество генератора оценивают через популярные метрики генеративного моделирования [1]. Такие метрики не дают четкого понимания о качестве самого численного метода, так как зависят от компонентов модели, не связанных с ОТ.

Таким образом, возникает необходимость разработки нетривиальных пар непрерывных вероятностных распределений  $(\mathbb{P}, \mathbb{Q})$ , которые могли бы быть использованы для **явной** оценки качества численных методов ОТ, то есть имели бы аналитически известное ОТ отображение. В данной диссертационной работе **предлагается метод** (глава 5 работы) построения эталонных непрерывных пар  $\mathbb{P}, \mathbb{Q}$  с известным  $T^*$ .

## Глава 2. Метод для вычисления Васерштейн-2 оптимальных транспортных отображений

В данной главе предлагается новый одноэтапный неминимаксный алгоритм для обучения отображений оптимального транспорта относительно квадратичной функции стоимости (расстояния Васерштейна-2, задача 1). Предложенный алгоритм использует выпуклые по входу нейронные сети и циклически согласованную регуляризацию для аппроксимации расстояния Васерштейн-2. В отличие от популярных энтропийных и квадратичных регуляризаторов циклическая согласованность не приводит к появлению смещения и хорошо масштабируется на случай больших размерностей. Мы даем теоретические оценки свойств транспортного отображения, обученного нашим алгоритмом. С практической точки зрения наш алгоритм тестируется на большом числе задач: перенос цвета между изображениями, оптимальный транспорт в латентном пространстве, перенос стиля между изображениями, адаптация домена.

### 2.1 Введение

За последние пару лет, прошедших после введения генеративных состязательных сетей (GAN) в работе [12], они приобрели большую популярность. С их помощью можно определить стохастическую процедуру для получения выборок из заданного сложного вероятностного распределения  $\mathbb{Q}$  на пространстве  $\mathbb{R}^D$  (например, на пространстве изображений). Обычная схема работы генеративной сети включает в себя моделирование выборки из простого распределения  $\mathbb{P}$  и применение транспортного (генеративного) отображения  $T$ , которое отображает  $\mathbb{P}$  в требуемое распределение  $\mathbb{Q}$ .

Во многих случаях для вероятностных распределений  $\mathbb{P}, \mathbb{Q}$  имеются несколько различных транспортных отображений. Например, отображение на рис. 2.1б выглядит лучше, чем отображение на рис. 2.1а, так как оно хорошо структурировано и является обратимым.

В существующих подходах к генеративному обучению мало внимания уделяется структурным свойствам отображения. Например, в подходах, осно-

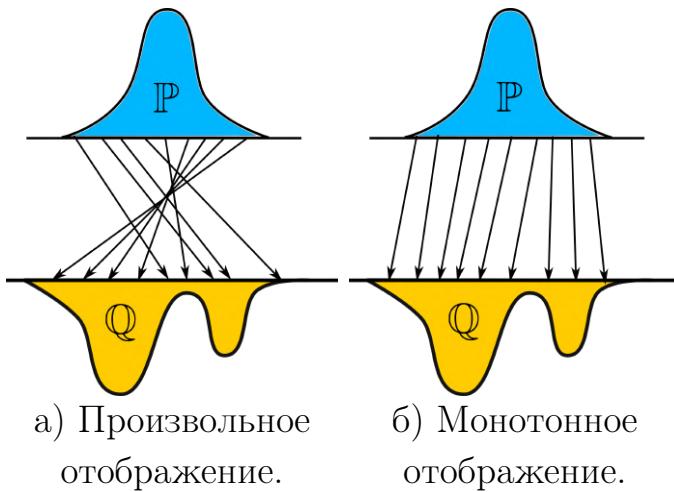


Рисунок 2.1 — Два возможных генеративных отображения, переводящих распределение  $\mathbb{P}$  в распределение  $\mathbb{Q}$ .

ванных на сетях GAN (таких, как  $f$ -GAN [13; 14], W-GAN [1] и других [15; 16]), транспортное отображение аппроксимируется нейронной сетью с архитектурой, зависящей от конкретной задачи.

Возникает естественный вопрос: как найти **хорошо структурированное** транспортное отображение  $T \sharp \mathbb{P} = \mathbb{Q}$ . Как правило, чем лучше структура отображения, тем такое отображение легче найти. Есть много способов определения того, что является хорошо структурированным отображением. Обычно такое отображение считается по определению непрерывным и, часто, обратимым. Можно отметить, что в случае, когда  $\mathbb{P}$  и  $\mathbb{Q}$  оба одномерны, единственный класс отображений  $T : \mathbb{R}^1 \rightarrow \mathbb{R}^1$  с такими свойствами составляют монотонные отображения,<sup>1</sup> т.е. отображения со свойством:  $(T(x) - T(x')) \cdot (x - x') > 0$  для любых  $x, x' \in \mathbb{R}^1$  ( $x \neq x'$ ). Случай одномерных отображений можно интуитивно легко распространить на случай пространств  $\mathbb{R}^D$ . В этом случае накладывается следующее условие:

$$\langle T(x) - T(x'), x - x' \rangle > 0 \quad \forall x, x' \in \mathbb{R}^D \quad (x \neq x'), \quad (2.1)$$

называемое условием монотонности. Отметим, что любая сюръективная функция, удовлетворяющая этому условию, является обратимой. В одномерном случае для любой пары абсолютно непрерывных  $\mathbb{P}, \mathbb{Q}$  с всюду положительной плотностью существует единственное монотонное транспортное отображение, задаваемое формулой  $T(x) = F_{\mathbb{Q}}^{-1}(F_{\mathbb{P}}(x))$  (см. [17]), где  $F_{(.)}$  – кумулятивная

<sup>1</sup>Мы рассматриваем только монотонные **возрастающие** отображения. Убывающие отображения имеют аналогичные свойства.

функция распределения для  $\mathbb{P}$  или  $\mathbb{Q}$ . Для  $D > 1$  монотонных транспортных отображений может быть много. Например, если  $\mathbb{P} = \mathbb{Q}$  – стандартные двумерные гауссовские распределения, то все вращения на углы из промежутка  $-\frac{\pi}{2} < \alpha < \frac{\pi}{2}$  являются монотонными и сохраняют распределение.

Единственность можно достичь путем рассмотрения только максимальных (см. [18]) монотонных отображений  $T : \mathbb{R}^D \rightarrow \mathbb{R}^D$ , удовлетворяющих условию

$$\sum_{n=1}^N \langle T(x_n), x_n - x_{n+1} \rangle > 0 \quad (2.2)$$

для любых  $N = 2, 3 \dots$  и  $N$  различных точек  $x_1, \dots, x_N \in \mathbb{R}^D$  ( $N + 1 \equiv 1$ ). Условие (2.2), известное под именем **циклической монотонности**, также влечет “обычную” монотонность (2.1).

Важно, что для почти всех пар непрерывных вероятностных распределений  $\mathbb{P}, \mathbb{Q}$  на  $\mathbb{R}^D$  существует единственное циклически монотонное отображение  $T : \mathbb{R}^D \rightarrow \mathbb{R}^D$  такое, что  $T \sharp \mathbb{P} = \mathbb{Q}$  (см. [17]). Таким образом, вместо поиска произвольного транспортного отображения можно существенно уменьшить аппроксимирующий класс отображений путем рассмотрения только циклически монотонных отображений.

Согласно работе [19] каждое циклически монотонное отображение  $T$  содержится в субградиенте некоторой выпуклой функции  $\psi : \mathbb{R}^D \rightarrow \mathbb{R}$ . Таким образом, по любому выпуклому классу функций можно строить циклически монотонные отображения (путем рассмотрения субградиентов функций). На практике в качестве класса выпуклых функций можно использовать глубокие **выпуклые по входу нейронные сети** (ICNN, см. [20]).

Формально для обучения циклического монотонного генеративного отображения можно применять существующие подходы, например, сети GAN (см. [12]), в которых множество генераторов сужается до градиентов сетей ICNN. Однако обычно при работе с сетями GAN требуется решение сложной задачи минимаксной оптимизации.

Оказывается, что циклические монотонные генераторы тесно связаны с расстоянием **Васерштейна-2** (задача 1). В работах [7; 21] используется двойственная форма  $\mathbb{W}_2$  для нахождения оптимального транспортного отображения, являющегося циклически монотонным. Прототипом обоих подходов является алгоритм градиентного спуска для вычисления расстояния  $\mathbb{W}_2$  из работы [22].

Недостатком всех таких методов является схожесть с GAN, т.е. целевая задача оптимизации является минимаксной.

Для циклически монотонных генераторов требуется, чтобы распределения  $\mathbb{P}, \mathbb{Q}$  были на пространствах одной и той же размерности. Это требование не представляет собой практического ограничения. Действительно, генеративное отображение оказывается возможным скомбинировать с декодером предварительно обученного автокодировщика, т.е. обучить генеративное отображение в латентном пространстве. Следует также отметить, что случай равных размерностей достаточно распространен в задачах компьютерного зрения. Типичным примером здесь является перенос стиля между изображениями. В этой задаче входные и выходные изображения имеют одинаковый размер и одинаковое число каналов. Среди других примеров упомянем перенос цвета между изображениями, задачу адаптации домена и т.п.

В настоящей главе мы развиваем подход, связанный с циклически монотонным генеративным обучением, то есть решением задачи 1. **Основной вклад главы** заключается в следующем:

1. Разработка одноэтапного неминимаксного алгоритма для обучения циклически монотонных транспортных отображений, т.е. оптимальных отображений для квадратичной стоимости Васерштейн-2 (п. 2.4).
2. Доказательство теоретической оценки для свойств аппроксимации транспортного отображения, обученного предложенным методом (п. 2.5).
3. Демонстрация работоспособности метода на модельных и практических задачах: ОТ между двумерными распределениями, ОТ между многомерными гауссовскими распределениями, ОТ для переноса цветов между изображениями, транспорта массы в латентных пространствах, переноса стиля между изображениями, адаптации домена (п. 2.6).
4. Исследование класса выпуклых по входу нейронных сетей, градиенты которых используются для приближения ОТ отображений (п. 2.7).

Наш алгоритм обобщает подход работы [7] и не использует минимальную оптимизацию, в нем применяется циклическая регуляризация и решается неминимаксная оптимационная задача. Как результат, наш алгоритм хорошо масштабируется на случай в больших размерностей и сходится до 10 раз быстрее, чем предыдущие аналогичные алгоритмы.

## 2.2 Обзор существующих методов

Современное генеративное обучение в основном использует **генеративные состязательные сети** (GAN) (см. [1; 12]). Основная модель сетей GAN состоит из двух конкурирующих сетей: генератора  $T$  и дискриминатора  $d$ . Генератор  $T$  получает входные выборки  $x$  из заданного распределения  $\mathbb{P}$  и пытается выдать реалистичные образцы из реального распределения данных  $\mathbb{Q}$ . Дискриминатор  $d$  пытается найти различие между сгенерированными и реальными распределениями  $T\sharp\mathbb{P}$  и  $\mathbb{Q}$  соответственно. Формально он аппроксимирует меру отличия между  $T\sharp\mathbb{P}$  и  $\mathbb{Q}$  (например,  $f$ -дивергенцию в смысле [13] или расстояние Васерштейна-1 [1]). Хотя для многих сетей GAN заявляется их превосходное качество (см. [23; 24]), обучение таких моделей всегда сопряжено с трудностями из-за минимаксной природы рассматриваемой оптимизационной задачи.

Другая важная ветвь генеративного обучения связана с теорией **оптимального транспорта** (ОТ) [6; 25]. В методах теории оптимального транспорта ищется генеративное отображение  $T : \mathbb{R}^D \rightarrow \mathbb{R}^D$ , оптимальное в смысле заданной стоимости транспорта  $c : \mathbb{R}^D \times \mathbb{R}^D \rightarrow \mathbb{R}$ :

$$\text{Cost}(\mathbb{P}, \mathbb{Q}) = \min_{T\sharp\mathbb{P}=\mathbb{Q}} \int_{\mathbb{R}^D} c(x, T(x)) d\mathbb{P}(x). \quad (2.3)$$

Задача (2.3) также известна как формулировка Монже задачи оптимального транспорта [25] для цены  $c$  общего вида.

Популярный метод оптимального транспорта [2] основан на оптимизации **регуляризованной двойственной к (2.3) формы** (1.3), упомянутой ранее. В этом методе обучаются два потенциала  $f, g$  (двойственные переменные задачи ОТ), после чего применяется метод барицентрических проекций для получения требуемой (третьей) генеративной сети  $T$ . Хотя в этом методе используется неминимаксная оптимизация, он состоит из двух последовательных шагов и находит смещенное решение задачи ОТ из-за энтропийной регуляризации.

В случае квадратичной стоимости транспорта  $c(x, y) = \frac{\|x-y\|^2}{2}$  величина (2.3) известна как квадрат расстояния **Васерштейна-2** (1.2). Это расстояние исследовалось в работах [17; 25—27] и обладает многими полезными свойствами, которые мы обсудим более детально в п. 2.3. Отметим, что оптимальное отображение для квадратичной функции стоимости является циклически монотонным. Для нахождения такого отображения имеется ряд алгоритмов: [7; 21; 28].

Недавно предложенный в работе [21] метод использует алгоритм градиентного спуска из работы [22] для вычисления расстояния  $\mathbb{W}_2$ . Ключевой идеей метода является аппроксимация оптимального потенциала  $\psi^*$  в (1.4) сетью ICNN (см. [20]) и извлечение оптимального генератора  $T^*$  из его градиента  $\nabla\psi^*$ . Однако этот подход является практически нецелесообразным из-за высокой вычислительной сложности: во время основного оптимизационного цикла в нем решается дополнительная оптимизационная подзадача, которая вычислительно затратна, хотя и выпукла. Это было отмечено в оригинальной статье и де-факто подтверждается отсутствием экспериментов со сложными распределениями. Уточнение этого подхода предлагается в [7]. В этой работе внутренняя оптимизационная подзадача отсутствует, и дополнительная нейросеть используется для аппроксимации ее решения. Это ускоряет вычисления, но при этом задача все равно остается минимаксной.

### 2.3 Предварительные сведения

В этом параграфе мы напомним свойства  $\mathbb{W}_2$ -расстояния (1.2) и укажем на его связь с монотонными отображениями. В настоящей работе мы предполагаем, что  $\mathbb{P}, \mathbb{Q} \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$ . Это условие гарантирует, что задача (2.3) корректно поставлена в том смысле, что оптимальное отображение  $T^*$  всегда существует. Из Теоремы 2.12 в [27] следует, что его ограничение на носитель  $\mathbb{P}$  единствено (с точностью до значений на малых множествах) и обратимо. Свойства симметрии также верны для его обратного отображения  $(T^*)^{-1}$ , что влечет симметрию в определении (2.3) квадратичной функции стоимости (1.2). Напомним, что согласно [27] двойственная форма (1.4) к (1.2) имеет вид

$$\begin{aligned} \mathbb{W}_2^2(\mathbb{P}, \mathbb{Q}) &= \underbrace{\int_{\mathbb{R}^D} \frac{\|x\|^2}{2} d\mathbb{P}(x) + \int_{\mathbb{R}^D} \frac{\|y\|^2}{2} d\mathbb{Q}(y)}_{\text{Const}(\mathbb{P}, \mathbb{Q})} - \\ &\quad \underbrace{\min_{\psi \in \text{Convex}} \left[ \int_{\mathbb{R}^D} \psi(x) d\mathbb{P}(x) + \int_{\mathbb{R}^D} \bar{\psi}(y) d\mathbb{Q}(y) \right]}_{\text{Corr}(\mathbb{P}, \mathbb{Q})}, \end{aligned} \tag{2.4}$$

где минимум берется по всем выпуклым функциям (потенциалам)  $\psi : \mathbb{R}^D \rightarrow \mathbb{R} \cup \{\infty\}$  и  $\bar{\psi}(y) = \max_{x \in \mathbb{R}^D} (\langle x, y \rangle - \psi(x))$  – выпуклая сопряженная функция [29] to  $\psi$ , также являющейся выпуклой функцией  $\bar{\psi} : \mathbb{R}^D \rightarrow \mathbb{R} \cup \{\infty\}$ .

Значение минимума в (1.3) мы будем называть циклически монотонными корреляциями и обозначать посредством  $\text{Corr}(\mathbb{P}, \mathbb{Q})$ . Приравнивая (1.2) и (2.4), мы приходим к формуле

$$\text{Corr}(\mathbb{P}, \mathbb{Q}) = \max_{T: \mathbb{P} = \mathbb{Q}} \int_{\mathbb{R}^D} \langle x, T(x) \rangle d\mathbb{P}(x). \quad (2.5)$$

Отметим, что  $(-\text{Corr}(\mathbb{P}, \mathbb{Q}))$  может рассматриваться как стоимость оптимального транспорта для билинейной функции стоимости  $c(x, y) = -\langle x, y \rangle$ , см. [17]. Таким образом, поиск отображения оптимального транспорта  $T^*$  для  $\mathbb{W}_2$  эквивалентен нахождению отображения, максимизирующего корреляции в (2.5).

Напомним также, что для расстояния  $\mathbb{W}_2$  градиент  $T^* = \nabla \psi^*$  оптимального потенциала  $\psi^*$  непосредственно минимизирует (1.1), см. [27]. Будучи градиентом выпуклой функции, он автоматически является циклически монотонным. В частности, обратное отображение может быть получено путем взятия градиента по входу сопряженного оптимального потенциала  $\bar{\psi}^*(y)$  (см. [17]), т.е.

$$(T^*)^{-1}(y) = (\nabla \psi^*)^{-1}(y) = \nabla \bar{\psi}^*(y). \quad (2.6)$$

На практике основной потенциал  $\psi$  можно аппроксимировать параметрическим классом  $\Theta$  входных выпуклых функций  $\psi_\theta$  и затем оптимизировать корреляции

$$\min_{\theta \in \Theta} \text{Corr}(\mathbb{P}, \mathbb{Q} | \psi_\theta) = \min_{\theta \in \Theta} \left[ \int_{\mathbb{R}^D} \psi_\theta(x) d\mathbb{P}(x) + \int_{\mathbb{R}^D} \bar{\psi}_\theta(y) d\mathbb{Q}(y) \right] \quad (2.7)$$

для нахождения приближения оптимального генератора  $\nabla \hat{\psi}$  из приближенного потенциала  $\hat{\psi}$ .

Для оптимизации (2.7) можно использовать стандартный метод стохастического градиентного спуска. При этом, можно получить аналитическую формулу для градиента (2.7) относительно параметра  $\theta$  с использованием только  $\psi_\theta$  (по поводу соответствующих выкладок см. [21; 22]):

$$\frac{\partial \text{Corr}(\mathbb{P}, \mathbb{Q} | \psi_\theta)}{\partial \theta} = \int_{\mathbb{R}^D} \frac{\partial \psi_\theta(x)}{\partial \theta} d\mathbb{P}(x) - \int_{\mathbb{R}^D} \frac{\partial \psi_\theta(\hat{x})}{\partial \theta} d\mathbb{Q}(y),$$

где производная  $\frac{\partial \psi_\theta}{\partial \theta}$  во втором интеграле вычисляется в точке  $\hat{x} = (\nabla \psi_\theta)^{-1}(y)$ , т.е.  $\hat{x}$  – прообраз  $y$  для  $\nabla \psi_\theta$ .

На практике оба интеграла заменяются их оценками по методу Монте-Карло по случайнм выборкам из  $\mathbb{P}$  и  $\mathbb{Q}$ . Тем не менее, для вычисления второго интеграла необходимо восстановить обратные значения текущего отображения  $\nabla\psi_\theta$  при всех  $y \sim \mathbb{Q}$  в выборке. Для этой цели следует решить следующую оптимизационную подзадачу:

$$\hat{x} = (\nabla\psi_\theta)^{-1}(y) \Leftrightarrow \hat{x} = \arg \max_{x \in \mathbb{R}^D} (\langle x, y \rangle - \psi_\theta(x)) \quad (2.8)$$

для каждого  $y \sim \mathbb{Q}$  в выборке. Оптимизационная задача (2.8) выпукла, но **сложна**, поскольку в ней требуется многократное вычисление градиента  $\psi_\theta$ . Она также вычислительно затратна, поскольку в общем  $\psi_\theta$  является большой нейронной сетью. Помимо этого каждый раз при выполнении итераций по  $\theta$  возникает новая независимая выборка. Это затрудняет использование информации о решении (2.8) с предыдущего шага метода градиентного спуска по  $\theta$  в (2.7).

## 2.4 Одноэтапный неминимаксный алгоритм

В п. 2.4.1 мы описываем наш новый одноэтапный алгоритм с неминимаксной оптимизацией для обучения циклически монотонных генеративных отображений. В п. 2.4.2 мы приводим наши основные теоретические результаты по свойствам аппроксимации предложенного алгоритма.

### 2.4.1 Алгоритм

Для упрощения внутренней оптимизационной процедуры обращения значений градиента  $\nabla\psi_\theta$  рассмотрим следующую вариационную аппроксимацию основной задачи:

$$\begin{aligned} \min_{\psi \in \text{Convex}} \text{Corr}(\mathbb{P}, \mathbb{Q} | \psi) &= \min_{\psi \in \text{Convex}} \left[ \int_{\mathbb{R}^D} \psi(x) d\mathbb{P}(x) + \overbrace{\int_{\mathbb{R}^D} \max_{x \in \mathbb{R}^D} [\langle x, y \rangle - \psi(x)] d\mathbb{Q}(y)}^{=\bar{\psi}(y)} \right] \\ &= \min_{\psi \in \text{Convex}} \left[ \int_{\mathbb{R}^D} \psi(x) d\mathbb{P}(x) + \max_{H: \mathbb{R}^D \rightarrow \mathbb{R}^D} \int_{\mathbb{R}^D} [\langle H(y), y \rangle - \psi(H(y))] d\mathbb{Q}(y) \right], \end{aligned} \quad (2.9)$$

где, рассматривая произвольные измеримые функции  $H : \mathbb{R}^D \rightarrow \mathbb{R}^D$ , мы получаем вариационную оценку снизу, которая согласуется с действительным значением при  $H = (\nabla\psi)^{-1}(y) = \nabla\bar{\psi}(y)$ . Таким образом, возможным подходом здесь является аппроксимация как основного  $\psi$ , так и двойственного  $\bar{\psi}$  потенциалов двумя различными сетями  $\psi_\theta$  и  $\varphi_\omega$  с последующим решением оптимизационной задачи по параметрам  $\theta, \omega$ , например, путем метода стохастического градиентного спуска/подъема (см. [7]). Однако такая задача все еще остается минимаксной и, как следствие, обладает недостатками, присущими таким задачам: сходимость к локальным седловым точкам и неустойчивость во время обучения. Также в ней обычно требуется осуществлять нетривиальный подбор гиперпараметров.

Мы предлагаем безминимаксный метод, основанный на введении дополнительной регуляризации. **Ключевой идеей** нашего метода является введение регуляризационного члена  $\mathcal{R}_2^\mathbb{Q}(\psi, \varphi)$ , при котором стимулируется **циклическая состоятельность** [30], т.е. в нем оптимизированные транспортные отображения  $\nabla\psi$  и  $\nabla\varphi$  должны быть взаимно обратны:

$$\mathcal{R}_2^\mathbb{Q}(\psi, \varphi) \stackrel{def}{=} \int_{\mathbb{R}^D} \|\nabla\psi(\nabla\varphi(y)) - y\|^2 d\mathbb{Q}(y). \quad (2.10)$$

Из предыдущих рассуждений и уравнения (2.6) мы получаем, что циклическая состоятельность является вполне естественным условием для задач, связанных с  $\mathbb{W}_2$ -расстоянием. Более конкретно, если  $\nabla\psi$  и  $\nabla\varphi$  являются в точности обратными друг к другу, то  $\varphi$  – выпуклая сопряженная функция к  $\psi$  с точностью до константы.

В отличие от регуляризации, использованной в работе [2], в предлагаемых штрафах используются не сами значения потенциалов  $\psi, \varphi$ , а значения их градиентов (генераторов). Это помогает стабилизировать значение регуляризационного члена, который в случае [2] может принимать экстремально большие значения вследствие того, что выпуклые потенциалы быстро растут по модулю<sup>2</sup>.

Предлагаемая нами регуляризация приводит к следующей задаче **неминимаксной оптимизации** ( $\lambda > 0$ ):

$$\begin{aligned} & \min_{\varphi, \psi \in \text{Convex}} \text{Corr}(\mathbb{P}, \mathbb{Q} \mid \psi, \varphi; \lambda) \stackrel{def}{=} \\ & \min_{\varphi, \psi \in \text{Convex}} \left( \int_{\mathbb{R}^D} \psi(x) d\mathbb{P}(x) + \int_{\mathbb{R}^D} [\langle \nabla\varphi(y), y \rangle - \psi(\nabla\varphi(y))] d\mathbb{Q}(y) \right) + \frac{\lambda}{2} \mathcal{R}_2^\mathbb{Q}(\varphi, \psi). \end{aligned} \quad (2.11)$$

<sup>2</sup>Например, в случае  $T(x) = \nabla\psi(x) = x$  мы имеем квадратичный рост:  $\psi(x) = \frac{\|x\|^2}{2} + c$ .

---

**Алгоритм 1:** Численная процедура оптимизации регуляризованных корреляций (2.11)

---

**Вход:** Распределения  $\mathbb{P}, \mathbb{Q}$  с выборочным доступом; коэффициент циклически согласованного регуляризатора  $\lambda > 0$ ; пара выпуклых по входу нейронных сетей  $\psi_\theta$  и  $\varphi_\omega$ ; размер случайных выборок  $K > 0$ ;

**for**  $t = 1, 2, \dots$  **do**

1. Получение выборки  $X \sim \mathbb{P}$  и  $Y \sim \mathbb{Q}$ ;
2. Вычисление оценок по методу Монте–Карло для корреляций:

$$\mathcal{L}_{\text{Corr}} = \frac{1}{K} \left[ \sum_{x \in X} \psi_\theta(x) + \sum_{y \in Y} [\langle \nabla \varphi_\omega(y), y \rangle - \psi_\theta(\nabla \varphi_\omega(y))] \right];$$

3. Вычисление оценки по методу Монте–Карло для циклически согласованного регуляризатора:

$$\mathcal{L}_{\text{Cycle}} := \frac{1}{K} \sum_{y \in Y} \|\nabla \psi_\theta(\nabla \varphi_\omega(y)) - y\|_2^2;$$

4. Вычисление общей суммы потерь  $\mathcal{L}_{\text{Total}} := \mathcal{L}_{\text{Corr}} + \frac{\lambda}{2} \cdot \mathcal{L}_{\text{Cycle}}$ ;
5. Выполнение шага метода градиентного спуска по  $\{\theta, \omega\}$  с использованием  $\frac{\partial \mathcal{L}_{\text{Total}}}{\partial \{\theta, \omega\}}$ ;

**end**

---

Метод практической численной оптимизации дается в алгоритме 1. Для аппроксимации  $\psi, \varphi$  мы используем две выпуклые по входу нейронные сети  $\psi_\theta, \varphi_\omega$  [20]. Все интегралы заменяются их оценками по методу Монте–Карло на выборках из  $\mathbb{P}$  и  $\mathbb{Q}$ . Оптимизация осуществляется применением стохастического градиентного спуска по параметрам  $\theta, \omega$  основного  $\psi_\theta$  и сопряженного  $\varphi_\omega$  потенциалов.

Для оценки градиентов  $\nabla \varphi_\omega, \nabla \psi_\theta$  и градиентов в (2.11) по параметрам  $\theta, \omega$  мы используем процедуру автоматического дифференцирования. Время для вычисления градиента в (2.11) по  $\theta, \omega$  отличается по порядку (на мультиплексивную константу) от времени для вычисления значения  $\psi_\theta(x)$ . Эмпирически найденное значение этой константы равно 8–12 (это значение зависит от конкретной архитектуры сети ICNN  $\psi_\theta(x)$ ). В п. 2.6.3 мы показываем, что наш неминимаксный подход **сходится до 10 раз быстрее**, чем минимаксные методы из работ [7] и [21].

## 2.4.2 Аппроксимация

Наш подход, основанный на методе градиентного спуска, представлен в п. 2.4.1. В нем вычисляется  $\text{Corr}(\mathbb{P}, \mathbb{Q})$  путем аппроксимации с ограниченным множеством выпуклых потенциалов. Пусть  $(\hat{\psi}, \hat{\phi})$  – пара потенциалов, полученных путем оптимизации корреляции. Обученные отображения  $\nabla \hat{\psi}$  (прямое) и  $\nabla \hat{\phi}$  (обратное) формально являются побочными продуктами оптимизации (2.11). Мы показываем, что на самом деле выполнено  $\nabla \hat{\psi} \approx \nabla \psi^*$  и  $\nabla \hat{\phi} \approx \nabla \bar{\psi}^*$ .

**Теорема 2.4.1** (транспортное свойство для приближений регуляризованных корреляций). *Пусть  $\mathbb{P}, \mathbb{Q} \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$ . Пусть  $\psi^* : \mathbb{R}^D \rightarrow \mathbb{R}$  – оптимальный выпуклый потенциал Бренье:*

$$\psi^* = \arg \min_{\psi \in \text{Convex}} \text{Corr}(\mathbb{P}, \mathbb{Q} | \psi) = \arg \min_{\psi \in \text{Convex}} \left[ \int_{\mathbb{R}^D} \psi(x) d\mathbb{P}(x) + \int_{\mathbb{R}^D} \bar{\psi}(y) d\mathbb{Q}(y) \right]. \quad (2.12)$$

Пусть дифференцируемые выпуклые функции  $\hat{\psi} : \mathbb{R}^D \rightarrow \mathbb{R}$  и  $\hat{\phi} : \mathbb{R}^D \rightarrow \mathbb{R}$  таковы, что, при некотором  $\varepsilon \in \mathbb{R}$ ,

$$\begin{aligned} \text{Corr}(\mathbb{P}, \mathbb{Q} | \hat{\psi}, \hat{\phi}; \lambda) &\leqslant \left[ \int_{\mathbb{R}^D} \psi^*(x) d\mathbb{P}(x) + \int_{\mathbb{R}^D} \bar{\psi}^*(y) d\mathbb{Q}(y) \right] + \varepsilon = \\ &= \text{Corr}(\mathbb{P}, \mathbb{Q} | \psi^*) + \varepsilon. \end{aligned} \quad (2.13)$$

Предположим, что  $\hat{\psi}$  является  $\beta$ -сильно выпуклой ( $\beta > \frac{1}{\lambda} > 0$ ) и  $\mathcal{B}$ -гладкой ( $\mathcal{B} \geqslant \beta$ ). Предположим также, что  $\hat{\phi}$  имеет биективный градиент  $\nabla \hat{\phi}$ . Тогда имеют место следующие неравенства:

1. *Оценка сверху для корреляции* (регуляризованные корреляции мажорируют истинные)

$$\text{Corr}(\mathbb{P}, \mathbb{Q} | \hat{\psi}, \hat{\phi}; \lambda) \geqslant \text{Corr}(\mathbb{P}, \mathbb{Q} | \psi^*) \quad (\text{т.е. } \varepsilon \geqslant 0);$$

2. *Прямое транспортное свойство* (отображение  $\nabla \hat{\psi}$  является  $O(\varepsilon)$ -близким к оптимальному отображению  $\nabla \psi^*$ )

$$\frac{1}{2} \int_{\mathbb{R}^D} \|\nabla \hat{\psi}(x) - \nabla \psi^*(x)\|^2 d\mathbb{P}(x) \leqslant \frac{(\mathcal{B})^2 \cdot \varepsilon}{\lambda \beta - 1} \cdot \left[ \frac{1}{\sqrt{\beta}} + \sqrt{\lambda} \right]^2 = O(\varepsilon);$$

3. *Обратное транспортное свойство* (отображение  $\nabla \hat{\phi}$  является  $O(\varepsilon)$ -близким к обратному оптимальному  $(\nabla \psi^*)^{-1}$ )

$$\frac{1}{2} \int_{\mathbb{R}^D} \|\nabla \hat{\phi}(y) - \nabla \bar{\psi}^*(y)\|^2 d\mathbb{Q}(y) \leqslant \frac{\varepsilon}{\beta - \frac{1}{\lambda}} = O(\varepsilon).$$

Неформально говоря, теорема 2.4.1 утверждает, что чем лучше мы аппроксимируем корреляции между  $\mathbb{P}$  и  $\mathbb{Q}$  при помощи потенциалов  $\hat{\psi}, \hat{\phi}$ , тем более близкими к  $\nabla\psi^*$  и  $\nabla\bar{\psi}^*$  ожидаются быть полученные  $\nabla\hat{\psi}$  и  $\nabla\hat{\phi}$ , соответственно (в смысле квадрата  $\mathcal{L}_2$  расстояния между функциями). В следующей секции мы докажем теорему 2.4.1 и обсудим вопросы, связанные с гладкостью и сильной выпуклостью.

В теореме 2.14 демонстрируется, что в действительности корреляции можно аппроксимировать требуемым образом, если аппроксимирующий класс функций для потенциалов является достаточно большим.

**Теорема 2.4.2** (аппроксимируемость корреляций). *Пусть  $\mathbb{P}, \mathbb{Q} \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$ . Пусть  $\psi^* : \mathbb{R}^D \rightarrow \mathbb{R}$  – оптимальный выпуклый потенциал. Далее, пусть  $\Psi, \Phi$  – классы дифференцируемых выпуклых функций  $\mathbb{R}^D \rightarrow \mathbb{R}$  соответственно. Предположим, что*

1.  $\exists \psi_1 \in \Psi$  с  $\varepsilon_1$ -близким градиентом прямого отображения  $\nabla\psi^*$  в смысле  $\mathcal{L}^2(\mathbb{R}^D \rightarrow \mathbb{R}^D, \mathbb{P})$ :

$$\|\nabla\psi_1 - \nabla\psi^*\|_{\mathbb{P}}^2 \stackrel{\text{def}}{=} \int_{\mathbb{R}^D} \|\nabla\psi_1(x) - \nabla\psi^*(x)\|^2 d\mathbb{P}(x) \leq \varepsilon_1,$$

где  $\psi_1$  является  $\mathcal{B}$ -гладкой;

2.  $\exists \varphi_2 \in \Phi$  с  $\varepsilon_2$ -близким градиентом к обратному отображению  $\nabla\bar{\psi}^*$  в смысле  $\mathcal{L}^2(\mathbb{R}^D \rightarrow \mathbb{R}^D, \mathbb{Q})$ :

$$\|\nabla\varphi_2 - \nabla\bar{\psi}^*\|_{\mathbb{Q}}^2 \stackrel{\text{def}}{=} \int_{\mathbb{R}^D} \|\nabla\varphi_2(y) - \nabla\bar{\psi}^*(y)\|^2 d\mathbb{Q}(y) \leq \varepsilon_2.$$

Пусть  $(\hat{\psi}, \hat{\phi})$  – минимизаторы регуляризованных корреляций в  $\Psi \times \Phi$ :

$$(\hat{\psi}, \hat{\phi}) = \arg \min_{\psi \in \Psi, \varphi \in \Phi} \text{Corr}(\mathbb{P}, \mathbb{Q} | \psi, \varphi; \lambda). \quad (2.14)$$

Тогда регуляризованные корреляции для  $(\hat{\psi}, \hat{\phi})$  удовлетворяют следующему неравенству:

$$\begin{aligned} & \text{Corr}(\mathbb{P}, \mathbb{Q} | \hat{\psi}, \hat{\phi}; \lambda) \\ & \leq \text{Corr}(\mathbb{P}, \mathbb{Q}) + \left[ \frac{\lambda}{2} (\mathcal{B}\sqrt{\varepsilon_2} + \sqrt{\varepsilon_1})^2 + (\mathcal{B}\sqrt{\varepsilon_2} + \sqrt{\varepsilon_1}) \cdot (\sqrt{\varepsilon_2}) + \frac{\mathcal{B}}{2} \varepsilon_2 \right], \end{aligned} \quad (2.15)$$

т.е. регуляризованные корреляции не превышают истинных плюс  $O(\varepsilon_1 + \varepsilon_2)$ .

Применяя теоремы 2.4.2 и 2.4.1, мы заключаем, что приближенные решения  $\hat{\psi}, \hat{\phi}$  задачи (2.14) являются  $O(\varepsilon_1 + \varepsilon_2)$ -близкими к  $\nabla\psi^*$  и  $\nabla\bar{\psi}^*$  соответственно. На практике разумно использовать выпуклые по входу нейронные сети в качестве классов функций  $\Psi, \Phi$ . Полносвязные сети ICNN удовлетворяют универсальному аппроксимационному свойству, см. работу [31].

## 2.5 Доказательства теоретических результатов

В данной главе мы докажем наши основные теоремы 2.4.1 и 2.4.2. Далее в работе через  $\mathcal{L}^2(\mathbb{R}^D \rightarrow \mathbb{R}^D, \mathbb{P})$  мы будем обозначать гильбертово пространство квадратично интегрируемых функций  $f : \mathbb{R}^D \rightarrow \mathbb{R}^D$  относительно вероятностной меры  $\mathbb{P}$ . Соответствующее скалярное произведение для  $f_1, f_2 \in \mathcal{L}^2(\mathbb{R}^D \rightarrow \mathbb{R}^D, \mathbb{P})$  обозначается посредством

$$\langle f_1, f_2 \rangle_{\mathbb{P}} \stackrel{\text{def}}{=} \int_{\mathbb{R}^D} \langle f_1(x), f_2(x) \rangle d\mathbb{P}(x).$$

Пусть  $\|\cdot\|_{\mathbb{P}} = \sqrt{\langle \cdot, \cdot \rangle_{\mathbb{P}}}$  – норма, порожденная скалярным произведением. Для доказательства основных результатов нам понадобится следующая лемма.

**Lemma 2.5.1** ( $\mathcal{L}^2$  неравенство для расстояния Васерштейна-2). *Пусть  $\mathbb{P}$  – вероятностное распределение на  $\mathbb{R}^D$ . Пусть  $T_1, T_2 \in \mathcal{L}^2(\mathbb{R}^D \rightarrow \mathbb{R}^D, \mathbb{P})$ . Тогда выполнено следующее неравенство:*

$$\frac{1}{2} \|T_1(x) - T_2(x)\|_{\mathbb{P}}^2 \geq \mathbb{W}_2^2(T_1 \sharp \mathbb{P}, T_2 \sharp \mathbb{P}).$$

*Доказательство.* Определим транспортный план  $\mu = [T_1, T_2] \sharp \mathbb{P}$  между  $T_1 \sharp \mathbb{P}$  и  $T_2 \sharp \mathbb{P}$  и воспользуемся фактом, что его стоимость не меньше, чем оптимальная стоимость, то есть  $\mathbb{W}_2^2(T_1 \sharp \mathbb{P}, T_2 \sharp \mathbb{P})$ .  $\square$

Сначала мы докажем нашу основную теорему 2.4.1. Далее мы докажем основную теорему 2.4.2. В конце параграфа мы обсудим константы, возникающие в теоремах: параметры сильной выпуклости и гладкости.

*Доказательство теоремы 2.4.1.* Мы разобьем доказательство на три части.

### Часть 1. Оценка сверху для корреляций.

Для начала мы получим оценку снизу для регуляризованных корреляций  $\text{Corr}(\mathbb{P}, \mathbb{Q} | \hat{\psi}, \hat{\phi}; \lambda)$ , опуская регуляризационный член:

$$\begin{aligned} & \int_{\mathbb{R}^D} \hat{\psi}(x) d\mathbb{P}(x) + \int_{\mathbb{R}^D} [\langle y, \nabla \hat{\phi}(y) \rangle - \hat{\psi}(\nabla \hat{\phi}(y))] d\mathbb{Q}(y) = \\ & \int_{\mathbb{R}^D} \hat{\psi}(\nabla \overline{\psi^*}(y)) d\mathbb{Q}(y) + \int_{\mathbb{R}^D} [\langle y, \nabla \hat{\phi}(y) \rangle - \hat{\psi}(\nabla \hat{\phi}(y))] d\mathbb{Q}(y) = (2.16) \\ & \int_{\mathbb{R}^D} [\hat{\psi}(\nabla \overline{\psi^*}(y)) - \hat{\psi}(\nabla \hat{\phi}(y))] d\mathbb{Q}(y) + \int_{\mathbb{R}^D} [\langle y, \nabla \hat{\phi}(y) \rangle] d\mathbb{Q}(y) \geqslant \end{aligned}$$

$$\int_{\mathbb{R}^D} [\langle \nabla \hat{\psi}(\nabla \hat{\phi}(y)), \nabla \overline{\psi^*}(y) - \nabla \hat{\phi}(y) \rangle + \frac{\beta}{2} \|\nabla \overline{\psi^*}(y) - \nabla \hat{\phi}(y)\|^2] d\mathbb{Q}(y) + (2.17)$$

$$\int_{\mathbb{R}^D} \langle y, \nabla \hat{\phi}(y) \rangle d\mathbb{Q}(y) + \left[ \underbrace{\int_{\mathbb{R}^D} \langle y, \nabla \overline{\psi^*}(y) \rangle d\mathbb{Q}(y)}_{\text{Corr}(\mathbb{P}, \mathbb{Q})} - \underbrace{\int_{\mathbb{R}^D} \langle y, \nabla \overline{\psi^*}(x) \rangle d\mathbb{Q}(y)}_{\text{Corr}(\mathbb{P}, \mathbb{Q})} \right] = (2.18)$$

$$\langle \nabla \hat{\psi} \circ \nabla \hat{\phi}, \nabla \overline{\psi^*} - \nabla \hat{\phi} \rangle_{\mathbb{Q}} + \frac{\beta}{2} \|\nabla \overline{\psi^*} - \nabla \hat{\phi}\|_{\mathbb{Q}}^2 - \langle \text{id}_{\mathbb{R}^D}, \nabla \overline{\psi^*} - \nabla \hat{\phi} \rangle_{\mathbb{Q}} + \text{Corr}(\mathbb{P}, \mathbb{Q}) = (2.19)$$

$$\begin{aligned} & \langle \nabla \hat{\psi} \circ \nabla \hat{\phi} - \text{id}_{\mathbb{R}^D}, \nabla \overline{\psi^*} - \nabla \hat{\phi} \rangle_{\mathbb{Q}} + \frac{\beta}{2} \|\nabla \overline{\psi^*} - \nabla \hat{\phi}\|_{\mathbb{Q}}^2 + \text{Corr}(\mathbb{P}, \mathbb{Q}) = \\ & \frac{1}{2\beta} \|\nabla \hat{\psi} \circ \nabla \hat{\phi} - \text{id}_{\mathbb{R}^D}\|_{\mathbb{Q}} + \langle \nabla \hat{\psi} \circ \nabla \hat{\phi} - \text{id}_{\mathbb{R}^D}, \nabla \overline{\psi^*} - \nabla \hat{\phi} \rangle_{\mathbb{Q}} + \\ & \frac{\beta}{2} \|\nabla \overline{\psi^*} - \nabla \hat{\phi}\|_{\mathbb{Q}}^2 + \text{Corr}(\mathbb{Q}, \mathbb{P}) - \frac{1}{2\beta} \|\nabla \hat{\psi} \circ \nabla \hat{\phi} - \text{id}_{\mathbb{R}^D}\|_{\mathbb{Q}}^2 = \\ & \text{Corr}(\mathbb{Q}, \mathbb{P}) + \frac{1}{2} \left\| \frac{1}{\sqrt{\beta}} [\nabla \hat{\psi} \circ \nabla \hat{\phi} - \text{id}_{\mathbb{R}^D}] + \sqrt{\beta} [\nabla \overline{\psi^*} - \nabla \hat{\phi}] \right\|_{\mathbb{Q}}^2 - \\ & \frac{1}{2\beta} \|\nabla \hat{\psi} \circ \nabla \hat{\phi} - \text{id}_{\mathbb{R}^D}\|_{\mathbb{Q}}^2. (2.20) \end{aligned}$$

Здесь при переходе к (2.16) мы воспользовались заменой переменных  $\mathbb{P} = \nabla \overline{\psi^*} \sharp \mathbb{Q}$ , а при выводе (2.17) использовали  $\beta$ -сильную выпуклость функции  $\hat{\psi}$  и потом добавили нулевой член в (2.18). Далее, начиная с (2.19), для простоты мы заменили интегральные обозначения на обозначения с  $\mathcal{L}^2(\mathbb{R}^D \rightarrow \mathbb{R}^D, \mathbb{Q})$ .

Добавляя опущенный ранее регуляризационный член в (2.20), мы получаем следующий результат:

$$\begin{aligned} \text{Corr}(\mathbb{P}, \mathbb{Q} | \hat{\psi}, \hat{\phi}; \lambda) & \geqslant \text{Corr}(\mathbb{Q}, \mathbb{P}) + \frac{1}{2} (\lambda - \frac{1}{\beta}) \cdot \|\nabla \hat{\psi} \circ \nabla \hat{\phi} - \text{id}_{\mathbb{R}^D}\|_{\mathbb{Q}}^2 + \\ & \frac{1}{2} \left\| \frac{1}{\sqrt{\beta}} [\nabla \hat{\psi} \circ \nabla \hat{\phi} - \text{id}_{\mathbb{R}^D}] + \sqrt{\beta} [\nabla \overline{\psi^*} - \nabla \hat{\phi}] \right\|_{\mathbb{Q}}^2. (2.21) \end{aligned}$$

Так как  $\lambda > \frac{1}{\beta}$ , то из полученного неравенства вытекает, что истинные корреляции  $\text{Corr}(\mathbb{P}, \mathbb{Q})$  дают оценки сверху для регуляризованных корреляций  $\text{Corr}(\mathbb{P}, \mathbb{Q} | \hat{\psi}, \hat{\phi}; \lambda)$ . Отметим, что если оптимальное отображение  $\nabla\psi^*$  является  $\geq \beta$ -сильно выпуклым, то оценка (2.21) является точной. Действительно, эта оценка превращается в равенство при подстановке  $\nabla\hat{\phi} = (\nabla\hat{\psi})^{-1} = \nabla\overline{\psi^*}$ .

### Часть 2. Обратное генеративное свойство.

Продолжим вывод свойства из части 1. Пусть  $u = \nabla\hat{\psi} \circ \nabla\hat{\phi} - \text{id}_{\mathbb{R}^D}$  и  $v = \nabla\overline{\psi^*} - \nabla\hat{\phi}$ . Сопоставляя (2.21) с (2.13), получаем

$$\varepsilon \geq \frac{1}{2}\left(\lambda - \frac{1}{\beta}\right)\|u\|_{\mathbb{Q}}^2 + \frac{1}{2}\left\|\frac{1}{\sqrt{\beta}}u + \sqrt{\beta}v\right\|_{\mathbb{Q}}^2. \quad (2.22)$$

Теперь получим оценку сверху для  $\|v\|_{\mathbb{Q}}^2$ . При фиксированном  $u$  имеем

$$\left\|\frac{1}{\sqrt{\beta}}u + \sqrt{\beta}v\right\|_{\mathbb{Q}}^2 \leq 2\varepsilon - \left(\lambda - \frac{1}{\beta}\right)\|u\|_{\mathbb{Q}}^2.$$

Далее, применяя неравенство треугольника, получаем, что

$$\|\sqrt{\beta}v\|_{\mathbb{Q}} \leq \left\|\frac{1}{\sqrt{\beta}}u + \sqrt{\beta}v\right\|_{\mathbb{Q}} + \left\|\frac{1}{\sqrt{\beta}}u\right\|_{\mathbb{Q}} \leq \sqrt{2\varepsilon - \left(\lambda - \frac{1}{\beta}\right)\|u\|_{\mathbb{Q}}^2} + \left\|\frac{1}{\sqrt{\beta}}u\right\|_{\mathbb{Q}}. \quad (2.23)$$

Выражение в правой части (2.23) принимает свое максимальное значение  $\sqrt{\frac{2\varepsilon}{1 - \frac{1}{\lambda\beta}}}$

в  $\|u\|_{\mathbb{Q}} = \sqrt{\frac{2\varepsilon}{\lambda^2\beta - \lambda}}$ . Как следствие,

$$\|\nabla\overline{\psi^*} - \nabla\hat{\phi}\|_{\mathbb{Q}}^2 = \|v\|_{\mathbb{Q}}^2 \leq \frac{2\varepsilon}{\beta - \frac{1}{\lambda}}.$$

Применяя  $\mathcal{L}^2$ -неравенство из леммы 2.5.1 к распределению  $\mathbb{Q}$ , отображениям  $\nabla\overline{\psi^*}$  и  $\nabla\hat{\phi}$ , получим  $\mathbb{W}_2^2(\nabla\hat{\phi} \# \mathbb{P}, \mathbb{Q}) \leq \frac{\varepsilon}{\beta - \frac{1}{\lambda}}$ , что дает оценку сверху для расстояния между сгенерированным и целевым распределением.

### Часть 3. Прямое генеративное свойство.

Нам потребуется оценка (2.22). Поскольку в ней все слагаемые неотрицательны, имеем

$$\|u\|_{\mathbb{Q}}^2 \leq \frac{2\varepsilon}{\lambda - \frac{1}{\beta}}. \quad (2.24)$$

Для оценки сверху  $\|\nabla\psi^* - \nabla\hat{\psi}\|_{\mathbb{P}}$  воспользуемся (2.24).

Для начала отметим, что так как функция  $\hat{\psi}$   $\beta$ -сильно выпуклая, то выпуклая сопряженная к ней функция  $\hat{\psi}$  является  $\frac{1}{\beta}$ -гладкой. Таким образом,

градиент  $\nabla \bar{\hat{\Psi}}$  является  $\frac{1}{\beta}$ -липшицевым. Мы заключаем, что при всех  $x, x' \in \mathbb{R}^D$ :

$$\|\nabla \bar{\hat{\Psi}}(x) - \nabla \bar{\hat{\Psi}}(x')\| \leq \frac{1}{\beta} \|x - x'\|. \quad (2.25)$$

Возводя обе части (2.25) в квадрат, подставляя  $x = y$  и  $x' = \nabla \hat{\Psi}(\nabla \hat{\phi}(y))$  и интегрируя по  $\mathbb{R}^D$  по распределению  $\mathbb{Q}$ , мы приходим к следующему неравенству:

$$\begin{aligned} & \int_{\mathbb{R}^D} \|\nabla \bar{\hat{\Psi}}(y) - \nabla \hat{\phi}(y)\|^2 d\mathbb{Q}(y) \leq \\ & \frac{1}{(\beta)^2} \int_{\mathbb{R}^D} \|y - \nabla \hat{\Psi}(\nabla \hat{\phi}(y))\|^2 d\mathbb{Q}(y) \end{aligned} \quad (2.26)$$

Далее, имеем

$$\begin{aligned} \|\nabla \bar{\hat{\Psi}} - \nabla \hat{\phi}\|_{\mathbb{Q}}^2 &= \int_{\mathbb{R}^D} \|\nabla \bar{\hat{\Psi}}(y) - \nabla \hat{\phi}(y)\|^2 d\mathbb{Q}(y) \leq \\ & \int_{\mathbb{R}^D} \frac{1}{(\beta)^2} \underbrace{\|y - \nabla \hat{\Psi}(\nabla \hat{\phi}(y))\|}_{=-u(y)}^2 d\mathbb{Q}(y) = \frac{\|u\|_{\mathbb{Q}}^2}{(\beta)^2}. \end{aligned} \quad (2.27)$$

Здесь при переходе к (2.27) мы воспользовались неравенством (2.26).

Далее, используя неравенство треугольника для  $\|\cdot\|_{\mathbb{Q}}$ , получаем

$$\begin{aligned} \|\nabla \bar{\hat{\Psi}} - \nabla \bar{\Psi}^*\|_{\mathbb{Q}} &\leq \|\nabla \bar{\hat{\Psi}} - \nabla \hat{\phi}\|_{\mathbb{Q}} + \underbrace{\|\nabla \hat{\phi} - \nabla \bar{\Psi}^*\|_{\mathbb{Q}}}_{=v} \leq \\ & \sqrt{\frac{2\varepsilon}{\lambda - \frac{1}{\beta}}} \cdot \frac{1}{\beta} + \sqrt{\frac{2\varepsilon}{\beta - \frac{1}{\lambda}}} = \sqrt{\frac{2\varepsilon}{\lambda - \frac{1}{\beta}}} \cdot \left( \frac{1}{\beta} + \sqrt{\frac{\lambda}{\beta}} \right) = \\ & \sqrt{\frac{2\varepsilon}{\lambda\beta - 1}} \cdot \left( \frac{1}{\sqrt{\beta}} + \sqrt{\lambda} \right) \end{aligned} \quad (2.28)$$

Далее, мы получим оценку снизу левой части (2.28), используя  $\mathcal{B}$ -гладкость функции  $\hat{\Psi}$ . При всех  $x, x' \in \mathbb{R}^D$  имеем

$$\|\nabla \hat{\Psi}(x) - \nabla \hat{\Psi}(x')\| \leq \mathcal{B} \|x - x'\|. \quad (2.29)$$

Возводя в квадрат обе части (2.29), подставляя  $x = \nabla \bar{\hat{\Psi}}(y)$  и  $x' = \nabla \bar{\Psi}^*(y)$  и интегрируя по  $\mathbb{Q}$ , получим

$$\int_{\mathbb{R}^D} \|\nabla \hat{\Psi}(\nabla \bar{\hat{\Psi}}(y)) - \nabla \hat{\Psi}(\nabla \bar{\Psi}^*(y))\|^2 d\mathbb{Q}(y) \leq \mathcal{B}^2 \int_{\mathbb{R}^D} \|\nabla \bar{\hat{\Psi}}(y) - \nabla \bar{\Psi}^*(y)\|^2 d\mathbb{Q}(y)$$

Далее находим, что

$$\|\nabla \bar{\hat{\Psi}} - \nabla \bar{\Psi}^*\|_{\mathbb{Q}}^2 = \int_{\mathbb{R}^D} \|\nabla \bar{\hat{\Psi}}(y) - \nabla \bar{\Psi}^*(y)\|^2 d\mathbb{Q}(y) \geq$$

$$\begin{aligned} \int_{\mathbb{R}^D} \frac{1}{(\mathcal{B})^2} \left\| \underbrace{\nabla \hat{\psi}(\nabla \bar{\psi}(y))}_{=y} - \nabla \hat{\psi}(\nabla \bar{\psi}^*(y)) \right\|^2 d\mathbb{Q}(y) = \\ \frac{1}{(\mathcal{B})^2} \int_{\mathbb{R}^D} \|\nabla \psi^*(x) - \nabla \hat{\psi}(x)\|^2 d\mathbb{P}(x) \geq \frac{2}{(\mathcal{B})^2} \mathbb{W}_2^2(\nabla \hat{\psi} \sharp \mathbb{P}, \mathbb{Q}). \end{aligned} \quad (2.30)$$

Здесь при выводе (2.30) мы воспользовались  $\mathcal{L}^2$ -свойством расстояния Васерштейна-2 (лемма 2.5.1). Как следствие,

$$\mathbb{W}_2^2(\nabla \hat{\psi} \sharp \mathbb{P}, \mathbb{Q}) \leq \frac{1}{2} \int_{\mathbb{R}^D} \|\nabla \psi^*(x) - \nabla \hat{\psi}(x)\|^2 d\mathbb{P}(x) \leq \frac{(\mathcal{B})^2 \cdot \varepsilon}{\lambda \beta - 1} \cdot \left( \frac{1}{\sqrt{\beta}} + \sqrt{\lambda} \right)^2,$$

что завершает доказательство.  $\square$

*Доказательство теоремы 2.4.2.* Мы будем рассуждать, как при выводе первой части теоремы 2.4.1, но вместо свойства сильной выпуклости функции  $\psi_1$  для получения оценки сверху для регуляризованных корреляций мы воспользуемся  $\mathcal{B}$ -гладкостью для получения оценки снизу. В результате аналог (2.31) оценки (2.21) имеет вид:

$$\begin{aligned} \text{Corr}(\mathbb{P}, \mathbb{Q} | \psi_1, \varphi_2, \lambda) - \text{Corr}(\mathbb{P}, \mathbb{Q}) \leq \\ \frac{1}{2} (\lambda - \frac{1}{\mathcal{B}}) \cdot \|\nabla \psi_1 \circ \nabla \varphi_2 - \text{id}_{\mathbb{R}^D}\|_{\mathbb{Q}}^2 + \\ \frac{1}{2} \left\| \frac{1}{\sqrt{\mathcal{B}}} [\nabla \psi_1 \circ \nabla \varphi_2 - \text{id}_{\mathbb{R}^D}] + \sqrt{\mathcal{B}} [\nabla \bar{\psi}^* - \nabla \varphi_2] \right\|_{\mathbb{Q}}^2 \leq \end{aligned} \quad (2.31)$$

$$\begin{aligned} \frac{1}{2} (\lambda - \frac{1}{\mathcal{B}}) \cdot \|\nabla \psi_1 \circ \nabla \varphi_2 - \text{id}_{\mathbb{R}^D}\|_{\mathbb{Q}}^2 + \\ \frac{1}{2} \left[ \frac{1}{\sqrt{\mathcal{B}}} \|\nabla \psi_1 \circ \nabla \varphi_2 - \text{id}_{\mathbb{R}^D}\|_{\mathbb{Q}} + \sqrt{\mathcal{B}} \|\nabla \bar{\psi}^* - \nabla \varphi_2\|_{\mathbb{Q}} \right]^2 = \end{aligned} \quad (2.32)$$

$$\begin{aligned} \frac{\lambda}{2} \cdot \|\nabla \psi_1 \circ \nabla \varphi_2 - \text{id}_{\mathbb{R}^D}\|_{\mathbb{Q}}^2 + \|\nabla \psi_1 \circ \nabla \varphi_2 - \text{id}_{\mathbb{R}^D}\|_{\mathbb{Q}} \cdot \|\nabla \bar{\psi}^* - \nabla \varphi_2\|_{\mathbb{Q}} + \\ \frac{\mathcal{B}}{2} \|\nabla \bar{\psi}^* - \nabla \varphi_2\|_{\mathbb{Q}}^2 \end{aligned} \quad (2.33)$$

Здесь при переходе от (2.31) к (2.32) мы воспользовались неравенством треугольника. Для каждого  $y \in \mathbb{R}^D$  имеем

$$\|\nabla \psi_1(\nabla \varphi_2(y)) - \nabla \psi_1(\nabla \bar{\psi}^*(y))\| \leq \mathcal{B} \cdot \|\nabla \varphi_2(y) - \nabla \bar{\psi}^*(y)\|.$$

Возводя обе части неравенства в квадрат и интегрируя по  $\mathbb{Q}$ , мы получаем

$$\|\nabla \psi_1 \circ \nabla \varphi_2 - \nabla \psi_1 \circ \nabla \bar{\psi}^*\|_{\mathbb{Q}}^2 \leq (\mathcal{B})^2 \cdot \|\nabla \varphi_2 - \nabla \bar{\psi}^*\|_{\mathbb{Q}}^2 \leq (\mathcal{B})^2 \cdot \varepsilon_2. \quad (2.34)$$

Далее, учитывая  $\nabla \bar{\Psi}^* \sharp \mathbb{Q} = \mathbb{P}$ , имеем

$$\|\nabla \psi_1 \circ \nabla \bar{\Psi}^* - \underbrace{\nabla \psi^* \circ \nabla \bar{\Psi}^*}_{\text{id}_{\mathbb{R}^D}}\|_{\mathbb{Q}}^2 = \|\nabla \psi_1 - \nabla \psi^*\|_{\mathbb{P}}^2 \leq \varepsilon_1. \quad (2.35)$$

Применяя (2.34), (2.35) и неравенство треугольника, мы получаем оценку

$$\begin{aligned} & \|\nabla \psi_1 \circ \nabla \varphi_2 - \text{id}_{\mathbb{R}^D}\|_{\mathbb{Q}} \leq \\ & \|\nabla \psi_1 \circ \nabla \varphi_2 - \nabla \psi_1 \circ \nabla \bar{\Psi}^*\|_{\mathbb{Q}} + \|\nabla \psi_1 \circ \nabla \bar{\Psi}^* - \underbrace{\nabla \psi^* \circ \nabla \bar{\Psi}^*}_{\text{id}_{\mathbb{R}^D}}\|_{\mathbb{Q}} \leq \\ & \mathcal{B} \sqrt{\varepsilon_2} + \sqrt{\varepsilon_1}. \end{aligned} \quad (2.36)$$

Подставляя эти оценки в (2.31), получаем

$$\begin{aligned} & \text{Corr}(\mathbb{P}, \mathbb{Q} | \psi_1, \varphi_2, \lambda) - \text{Corr}(\mathbb{P}, \mathbb{Q}) \leq \\ & \frac{\lambda}{2} (\mathcal{B} \sqrt{\varepsilon_2} + \sqrt{\varepsilon_1})^2 + (\mathcal{B} \sqrt{\varepsilon_2} + \sqrt{\varepsilon_1}) \cdot (\sqrt{\varepsilon_2}) + \frac{\mathcal{B}}{2} \varepsilon_2. \end{aligned} \quad (2.37)$$

Используя неравенство

$$\text{Corr}(\mathbb{P}, \mathbb{Q} | \hat{\psi}, \hat{\varphi}, \lambda) \leq \text{Corr}(\mathbb{P}, \mathbb{Q} | \psi_1, \varphi_2, \lambda),$$

которое следует из определения  $\hat{\psi}, \hat{\varphi}$ , мы завершаем доказательство.  $\square$

Можно сформулировать и доказать аналогичный результат для основного оптимизационного метода с одним потенциалом (2.7). Однако мы не приводим его здесь ввиду наличия в литературе соответствующего результата (см. [21]).

Во всех наших теоретических результатах требовалась гладкость или сильная выпуклость потенциалов. Следует отметить, что аналогичное предположение о гладкости и сильной выпуклости также имеется в других работах о оптимальном транспорте для расстояния Васерштейна–2 (см., например, [32]).

Свойство  $\mathcal{B}$ -гладкости выпуклой функции  $\psi$  означает, что константа Липшица ее градиента  $\nabla \psi$  ограничена величиной  $\mathcal{B}$ . В нашем случае, константа  $\mathcal{B}$  служит в качестве адекватной меры сложности обученного отображения  $\nabla \psi$ : она оценивает то, насколько сильно изображение может искажать пространство.

Свойство сильной выпуклости является двойственным к гладкости в том смысле, что выпуклая сопряженная функция  $\bar{\Psi}$  к  $\beta$ -сильно выпуклой функции  $\psi$  является  $\frac{1}{\beta}$ -гладкой (и наоборот) [33]. В нашем случае,  $\beta$ -сильная выпуклость потенциала означает, что обратный к нему градиент  $(\nabla \psi)^{-1} = \nabla \bar{\Psi}$  существенно не деформирует пространство, т.е. его константа Липшица ограничена величиной  $\frac{1}{\beta}$ .

Воспользуемся нашей теоремой 2.4.2. Предположим, что отображение оптимального транспорта  $\nabla\psi^*$  между  $\mathbb{P}$  и  $\mathbb{Q}$  является градиентом  $\beta$ -сильно выпуклой ( $\beta > 0$ ) и  $\mathcal{B}$ -гладкой ( $\mathcal{B} < \infty$ ) функции. В этом случае, рассматривая классы  $\Psi = \Phi$  всех  $\min(\beta, \frac{1}{\beta})$ -сильно выпуклых и  $\max(\mathcal{B}, \frac{1}{\beta})$ -гладких функций и используя наш метод (при любом  $\lambda > \frac{1}{\beta}$ ), мы сможем точно вычислить корреляции и найти оптимальное  $\nabla\psi^*$ .

## 2.6 Вычислительные эксперименты

В этом параграфе мы экспериментально тестируем предложенный численный метод вычисления Васерштейн-2 ОТ отображений. Программный код, реализованный на фреймворке **PyTorch**, доступен публично по адресу

<https://github.com/iamalexkorotin/Wasserstein2GenerativeNetworks>.

Сети были обучены с использованием одной видеокарты GTX 1080Ti.

Описание обучения приводится в первом параграфе. В каждом последующем параграфе рассматривается конкретная задача, представляются результаты экспериментов и детали обучения: синтетические эксперименты в п. 2.6.2, оптимальный транспорт между гауссианами 2.6.3, оптимальный транспорт в латентном пространстве в п. 2.6.6, перенос цветов между изображениями в 2.6.4, адаптация домена в п. 2.6.5, непарный перенос стиля между изображениями в п. 2.6.7.

Архитектуры используемых входных выпуклых сетей (Dense/Conv ICNN) описываются в п. 2.7. Целью экспериментов является демонстрация возможности технической реализации нашего подхода и архитектур.

### 2.6.1 Общие детали экспериментов

В каждом из экспериментов оба потенциала (основной  $\psi_\theta$  и сопряженный  $\varphi_\omega$ ) имеют одну и ту же нейросетевую архитектуру. Минимизация (2.11) осуществляется с использованием стохастического метода градиентного спуска

с обрезкой весов (за исключением свободных членов) на слоях в СР-блоке на  $[0, +\infty)$ , см. п. 2.7. Был использован оптимизатор Adam [34].

Для каждой конкретной задачи мы предварительно обучаем потенциальную сеть  $\Psi_\theta$  путем минимизации квадрата ошибки для выполнения приближенного равенства  $\nabla\Psi_\theta(x) \approx x$  и копируем веса на  $\varphi_\omega$ . Таким образом, мы получаем хорошую инициализацию для основного обучения, т.е.  $\nabla\Psi_\theta$  и  $\nabla\varphi_\omega$  взаимно обратны.

В экспериментах с многомерными данными (оптимальный транспорт в латентном пространстве, адаптация домена, перенос стилей) добавляется следующий дополнительный регуляризатор к основной целевой функции (2.11):

$$\mathcal{R}_2^{\mathbb{P}}(\psi, \varphi) = \int_{\mathbb{R}^D} \|\nabla\varphi(\nabla\psi(x)) - x\|^2 d\mathbb{P}(x). \quad (2.38)$$

Член (2.38) аналогичен члену  $\mathcal{R}_2^{\mathbb{Q}}$  из (2.10). Он также сохраняет взаимную обратность прямого и обратного генеративных отображений. С теоретической точки зрения не представляет труда обобщить результаты теорем 2.4.1 и 2.4.2, но для оптимизации с двумя членами:  $\mathcal{R}_2^{\mathbb{P}}$  и  $\mathcal{R}_2^{\mathbb{Q}}$ . Однако мы не рассматриваем ситуацию с членом  $\mathcal{R}_2^{\mathbb{P}}$  для большей прозрачности доказательств.

## 2.6.2 Синтетические эксперименты в 2D

В этом параграфе мы тестируем наш алгоритм на 2D синтетических распределениях из [2; 35]. Во всех экспериментах распределение  $\mathbb{P}$  – стандартный гауссов шум и  $\mathbb{Q}$  – смеси гауссовских распределений или “швейцарский рулет”. Примеры полученных результатов показаны на рис. 2.3а, 2.3б и 2.2.

Как основные, так и сопряженные потенциалы  $\Psi_\theta$  и  $\varphi_\omega$  обладают архитектурой DenseICNN [2; 128; 128, 64]. Каждая сеть имеет приблизительно 25000 обучаемых параметров. Некоторые из них зануляются при обучении из-за обрезания весов (weight clipping). Для каждой конкретной задачи сеть была обучена 30000 итераций с размером случайных выборок 1024. Использовался оптимизатор Adam [34] с  $lr = 10^{-3}$ . В нашей циклической регуляризации мы полагали  $\lambda = 1$  и использовали дополнительную  $10^{-10} \mathcal{L}^1$  регуляризацию.

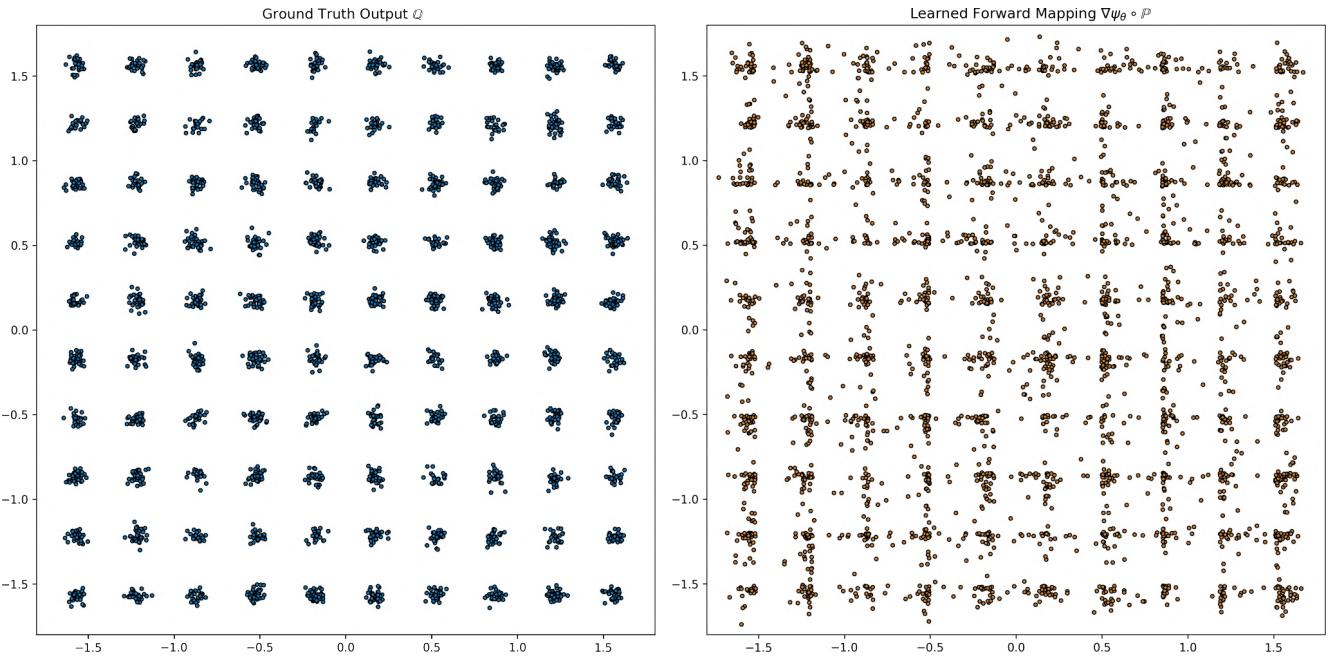


Рисунок 2.2 — Смесь 100 гауссовых распределений  $\mathbb{Q}$  и распределение  $\nabla\psi_\theta \# \mathbb{P}$ , вычисленное нашим алгоритмом.

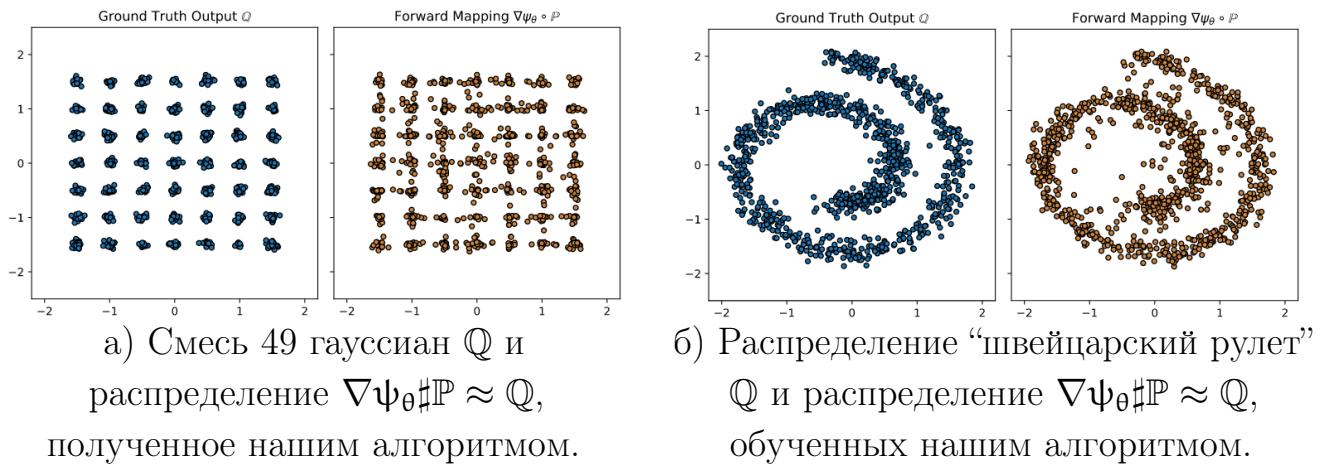


Рисунок 2.3 — Синтетические распределения, обученные нашим алгоритмом.

### 2.6.3 Оптимальный транспорт между гауссовскими распределениями

Мы рассматриваем гауссовскую постановку  $\mathbb{P}, \mathbb{Q} = \mathcal{N}(0, \Sigma_{\mathbb{P}}), \mathcal{N}(0, \Sigma_{\mathbb{Q}})$ , для которой эталонное решение задачи оптимального транспорта известно аналитически, см. Теорему 2.3 в [36]. Отметим, что здесь рассмотрение нецентрированных  $\mathbb{P}, \mathbb{Q}$  не является необходимым, поскольку  $\mathbb{W}_2^2(\mathbb{P}, \mathbb{Q}) = \frac{1}{2}\|\mu_{\mathbb{P}} - \mu_{\mathbb{Q}}\|^2 + \mathbb{W}_2^2(\mathbb{P}_0, \mathbb{Q}_0)$ , где  $\mathbb{P}_0, \mathbb{Q}_0$  — центрированные копии  $\mathbb{P}, \mathbb{Q}$ . В  $D$ -мерном пространстве,  $\sqrt{\Sigma_{\mathbb{P}}} (\sqrt{\Sigma_{\mathbb{Q}}} -$  аналогично) инициализируется как  $S_{\mathbb{P}}^T \Lambda S_{\mathbb{P}}$ , где  $S_{\mathbb{P}} \in O_D$  — случайное вращение,  $\Lambda$

– диагональная матрица с собственными значениями  $[\frac{1}{2}, \dots, \frac{1}{2}b^k, \dots, 2]$ ,  $b = \sqrt[D-1]{4}$ . Эталонное отображение оптимального транспорта из  $\mathbb{P}$  в  $\mathbb{Q}$  линейно и имеет вид

$$\nabla\psi^*(x) = \Sigma_{\mathbb{P}}^{-\frac{1}{2}} \left( \Sigma_{\mathbb{P}}^{\frac{1}{2}} \Sigma_{\mathbb{Q}} \Sigma_{\mathbb{P}}^{\frac{1}{2}} \right)^{\frac{1}{2}} \Sigma_{\mathbb{P}}^{-\frac{1}{2}} x.$$

Мы сравниваем наш подход с методом работы [2] [LSOT] и минимаксными методами из [21] [MM-1] и [7] [MM-2]. Положим  $\lambda = \min(D, 50)$  для нашего метода и  $\epsilon = 0.01$  (см. [2]) для метода LSOT (эти величины выбраны эмпирически). Во всех методах мы используем DenseICNN[1;  $D, D, \frac{D}{2}$ ] из п. 2.7.2. В методе LSOT мы не овыпукляем сети (веса не обрезаются), т.е. рассматриваются "обычные" нейронные сети без ограничений (эмпирически этот подход работает лучше), что и делается в работе по LSOT.

Для оценки качества восстановленного транспортного отображения  $\nabla\hat{\psi}$  мы рассматриваем **процент необъясненной дисперсии** (unexplained variance percentage):  $\mathcal{L}^2\text{-UVP}(\nabla\hat{\psi}) = 100 \cdot [\|\nabla\hat{\psi} - \nabla\psi^*\|_{\mathbb{P}}^2 / \text{Var}(\mathbb{Q})]\%$ . Здесь  $\nabla\psi^*$  – отображение оптимального транспорта. Для значений  $\approx 0\%$   $\nabla\hat{\psi}$  является хорошей аппроксимацией отображения оптимального транспорта. Для значений  $\geq 100\%$  отображение  $\nabla\hat{\psi}$  практически бесполезно. Действительно, тривиальный ответ  $\nabla\psi^0(x) = \mathbb{E}_{\mathbb{Q}}[y]$  дает  $\mathcal{L}^2\text{-UVP}(\nabla\psi^0) = 100\%$ .

Таблица 1 показывает, что метод LSOT приводит к большим ошибкам, которые сильно растут с увеличением размерности. Методы W2GN (наш), MM-1 and MM-2 функционируют практически одинаково в терминах метрики. Это, впрочем, и ожидалось, поскольку они все оптимизируют аналогичные функционалы. Эти методы вычисляют отображения оптимального транспорта с маленькой ошибкой ( $\mathcal{L}^2\text{-UVP} < 3\%$  даже в  $\mathbb{R}^{4096}$ ). Однако, как видно из рис. 2.4 (сходимость), наш подход **сходится в несколько раз быстрее**: это естественно следует из того, что методы типа MM содержат внутренний оптимизационный цикл.

<i>Dim</i>	2	4	8	16	32	64	128	256	512	1024	2048	4096
LSOT	<1	3.7	7.5	14.3	23	34.7	46.9				>50	
MM-1	<1	<1	<1	<1	<1	1.2	1.4	1.3	1.5	1.6	1.8	2.7
MM-2	<1	<1	<1	<1	<1	<1	1	1.1	1.2	1.3	1.5	2.1
W2GN	<1	<1	<1	<1	<1	<1	1	1.1	1.3	1.3	1.8	1.5

Таблица 1 — Сравнение  $\mathcal{L}^2$ -UVP (%) для методов LSOT, MM-1, MM-2 и (нашего) метода W2GN в размерностях  $D = 2, 4, \dots, 2^{12}$ .

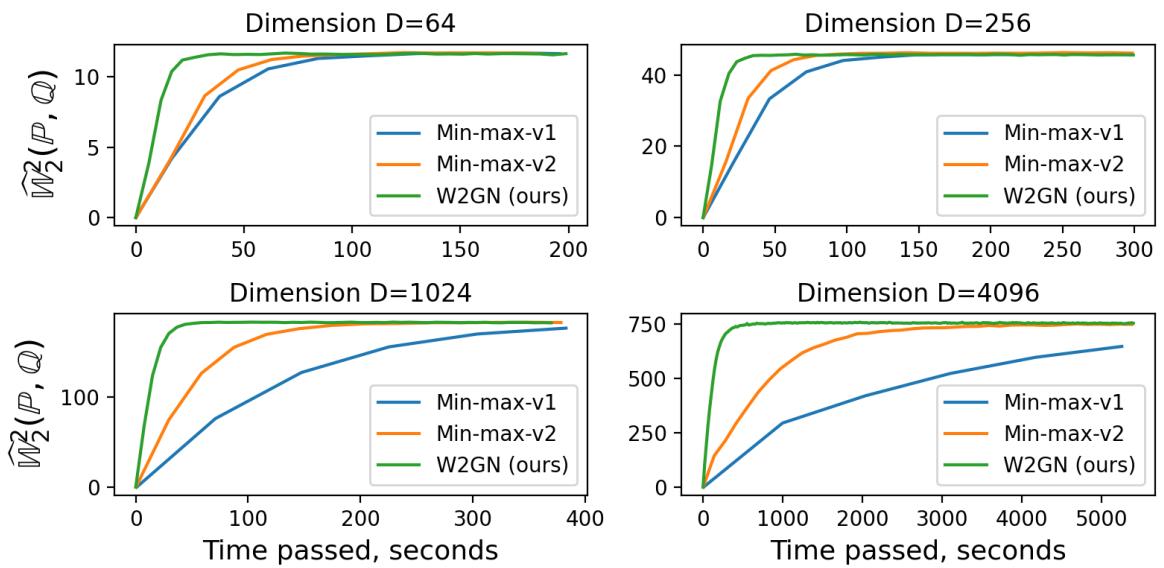


Рисунок 2.4 — Сравнение скорости сходимости для методов W2GN, MM-1 и MM-2 для размерностей  $D = 64, 256, 1024, 4096$ .

#### 2.6.4 Передача цветов

Задача передачи цветов между изображениями<sup>3</sup> состоит в отображении **цветовой палитры** одного изображения в палитру другого для того, чтобы изображение выглядело и “ощущалось” как оригинал.

Задача оптимального транспорта может быть применена и к задаче передачи цветов, но она является чувствительной к шуму и выбросам. Чтобы избежать этих сложностей, были предложены релаксации [32; 37]. При таких подходах решается дискретный вариант задачи оптимального транспорта Ва-

<sup>3</sup>Изображения могут иметь неравные размеры. Тем не менее, предполагается, что у них равное число каналов (например, каналов RGB).

серштейна-2. Вычисление стоимости оптимального транспорта при больших изображениях вряд ли технически осуществимо или даже вовсе неосуществимо из-за чрезвычайно большого размера цветовой палитры. Таким образом, обычно выполняется сжатие пиксельной цветовой палитры при помощи метода  $k$ -средних для того, чтобы сделать технически возможным вычисления в задаче оптимального транспорта. И все же при таком способе редукции может быть потеряна информация о цвете.

Наш алгоритм использует стохастическую оптимизацию на случайных выборках. Таким образом, в нем нет ограничений по размеру цветовой палитры. При обучении мы последовательно подаем выборки из пикселей изображений ( $\in \mathbb{R}^3$ ) в потенциальные сети с архитектурой DenseICNN с параметрами [3; 128, 128, 64]. Сеть была обучена за 5000 итераций с 1024 пикселями в случайной выборке. Использовался оптимизатор Adam с  $lr = 10^{-3}$ . В качестве параметра циклической регуляризации мы брали  $\lambda = 3$ . Мы налагаем дополнительный  $10^{-10} \mathcal{L}^1$ -штраф на веса.

Результаты передачи цветов для изображений размером  $\approx 10$  мегапикселей представлены на рис. 2.5а. Соответственные цветовые палитры приведены на рис. 2.5б. Ещё один пример передачи цветов представлен на рис. 2.6.

## 2.6.5 Адаптация домена

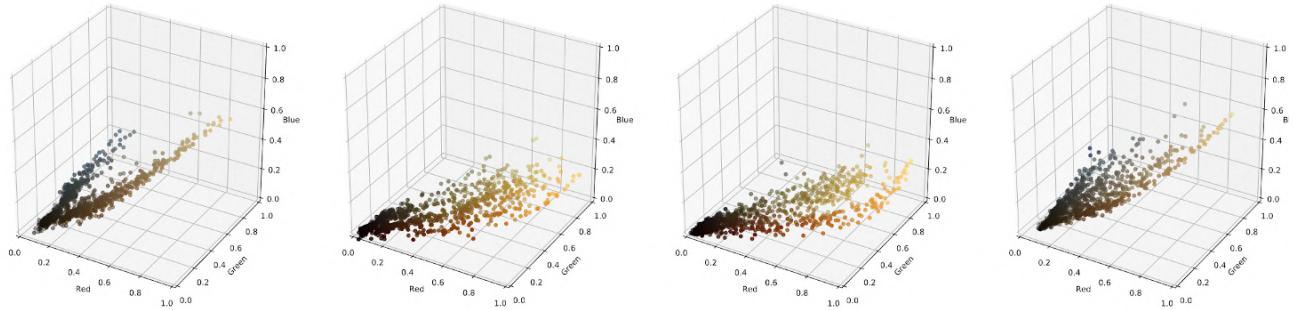
Задача адаптации домена заключается в обучении модели  $f$  (например, классификатора) на исходном распределении  $\mathbb{Q}$ . Модель должна хорошо функционировать на другом (но близком) распределении  $\mathbb{P}$ .

Большинство методов, основанных на теории оптимального транспорта, явно решают задачу адаптации домена путем преобразования распределения  $\mathbb{P}$  в  $\mathbb{Q}$  с последующим применением модели  $f$  для генерации образцов. В некоторых случаях отображение  $T : \mathbb{R}^D \rightarrow \mathbb{R}^D$  (преобразующее  $\mathbb{P}$  в  $\mathbb{Q}$ ) получается путем решения дискретной задачи оптимального транспорта [3; 38; 39], в то время как при некоторых подходах используется адаптация нейронных сетей для оценки отображения  $T$  (см. [2; 40]).

Мы рассмотрим задачу адаптации домена без учителя – эта задача представляет собой наиболее сложный вариант. В этой постановке метки класса



а) Исходные изображения (слева) и изображения, полученные путем передачи цветов (справа). Изображения имеют размеры:  $3300 \times 4856$  (первое) и  $2835 \times 4289$  (второе).



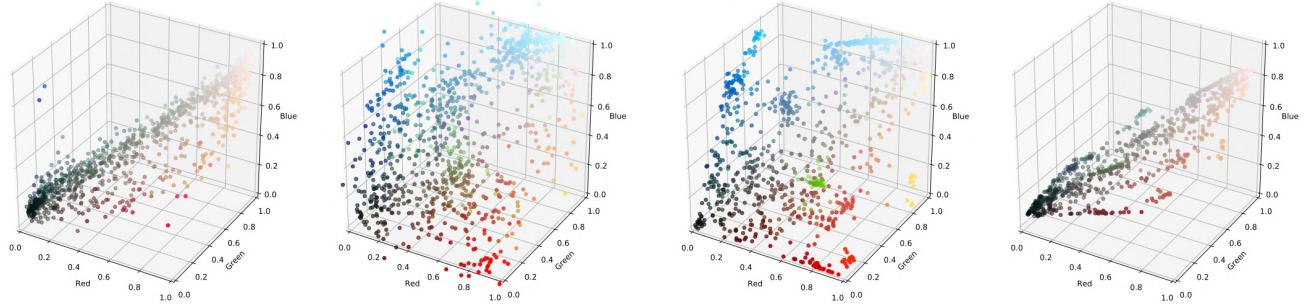
б) Цветовые палитры (3000 случайных пикселей) для оригинальных изображений (слева) и для изображений с перенесенными цветами (справа). Рисунок 2.5 — Результаты передачи цветов между изображениями с высоким разрешением (размера  $\approx 10$  мегапикселей) при помощи циклически монотонных отображений.

имеются в исходном домене. При обучении данная информация не используется. Наш метод обучает отображение  $T$  как градиент выпуклой функции. После обучения его можно применить ко вновь поступающим образцам, не представленным в обучающей выборке.

Наша модель тестируется на наборах данных рукописных цифр MNIST ( $\approx 60000$  изображений;  $28 \times 28$ ) и USPS ( $\approx 10000$  изображений, масштабированных к размеру  $28 \times 28$ ). Далее выполняется адаптация домена **USPS**  $\rightarrow$  **MNIST**. Для этой цели мы обучаем классификатор LeNet с  $\geq 99\%$ -процентной точностью на наборе данных MNIST. Далее мы применяем  $h$  к обоим наборам данных и выделяем 84 признаков последнего слоя. После этого формируются распределения  $\mathbb{Q}$  (признаки для набора данных MNIST) и  $\mathbb{P}$  (признаки для набора данных USPS).



а) Исходные изображения (слева) и изображения, полученные путем передачи цветов (справа).



б) Цветовые палитры (3000 случайных пикселей) для оригинальных изображений (слева) и для изображений с перенесенными цветами (справа).

Рисунок 2.6 — Результаты передачи цветов между изображениями с помощью циклически монотонного отображения.

Чтобы обучить циклически монотонное отображение для адаптации домена, мы используем потенциалы DenseICNN с параметрами [32; 128; 128, 128]. Наша модель обучалась на случайных выборках размера 64 с 10000 итерациями с циклической регуляризацией  $\lambda = 1000$ . Использовался оптимизатор Adam с  $lr = 10^{-4}$  и накладывался  $L^1$ -штраф  $10^{-7}$  на веса сетей.

Аналогично [2] мы сравниваем точность метода MNIST 1-NN (1 ближайший сосед) классификатора  $f$ , примененного к признакам  $x \sim \mathbb{P}$  of USPS с тем же классификатором, примененным к отображенным признакам  $\hat{T}(x)$ . Метод 1-NN выбирался в качестве модели классификации для исключения любого влияния модели базовой классификации на адаптацию домена.

Результаты эксперимента представлены в табл. 2. Поскольку качество адаптации домена зависит в высокой степени от качества извлеченных признаков, мы повторили эксперимент 3 раза, т.е. мы обучали 3 классификатора LeNet для эталонного набора данных MNIST для извлечения признаков. Результаты представлены с учетом дисперсии. Для сравнения с лучшими методами мы также добавляли показатели классификатора 1-NN, примененного к признакам исходного набора данных USPS, отображенных в признаки MNIST через дис-

	Повтор 1	Повтор 2	Повтор 3	Среднее ( $\mu \pm \sigma$ )
Target features	75.7%	77%	75.4%	76 $\pm$ 0.8%
<b>Отображенные признаки (W2GN, наш)</b>	80.6%	80.3%	82.7%	81.2 $\pm$ 1%
Отображенные признаки (дискретный ОТ)	76%	75.7%	76.1%	75.9 $\pm$ 0.4%
Отображенные признаки [2]	-	-	-	77.92%

Таблица 2 — Точность классификации (больше = лучше) 1-NN на наборах данных USPS  $\rightarrow$  MNIST в задаче адаптации домена.

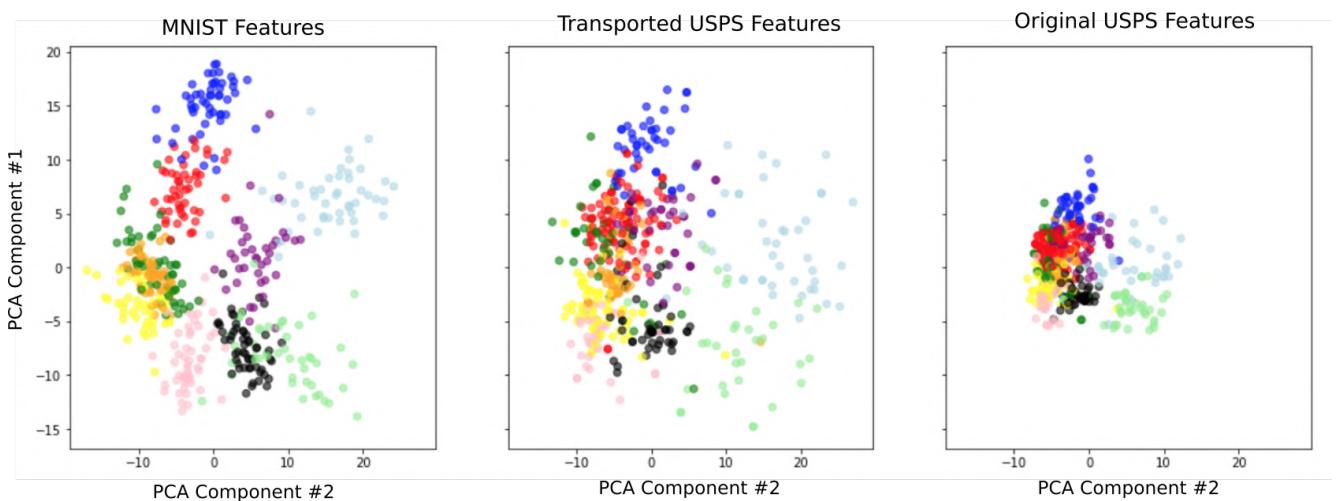


Рисунок 2.7 — Пара главных компонент пространства признаков. Слева направо: пространство признаков MNIST; признаки USPS, отраженные W2GN; оригинальные признаки USPS. Главные компоненты получены из признаков MNIST.

Цвета соответствуют различным классам цифр 0 – 9.

крайний оптимальный транспорт. Это может рассматриваться как “наиболее прямое” отображение оптимального транспорта<sup>4</sup>.

Результаты для нашего метода сравнимы с результатами из [2]. Однако мы не повторяли эксперименты из [2], поскольку в этой работе не был представлен исходный код для адаптации домена. Таким образом, мы отсылаем читателя напрямую к представленным результатам этой работы (табл. 1 [2], первый столбец с полученными значениями).

Для наглядности мы графически изображаем две основные компоненты метода анализа главных компонент (PCA) пространства признаков (использован-

<sup>4</sup>В отличие от нашего метода этот подход не может быть напрямую применен к примерам вне обучающей выборки (out-of-train-sample examples). Более того, сравнение оказывается технически неосуществимым для больших наборов данных.

Method	FID
AE: $Dec(Enc(X))$	7.5
AE Raw Decode: $Dec(Z)$	31.81
<b>W2GN+AE: <math>Dec(\hat{T}(Z))</math></b>	17.21
WGAN-QC : $Gen(Z)$	14.41

Таблица 3 — Значения метрики FID для сгенерированных изображений.

ного в проведенных экспериментах) на рис. 2.7: признаки MNIST, отображенные признаки USPS по нашему методу, оригинальные признаки USPS.

## 2.6.6 Оптимальный транспорт в латентном пространстве

Мы тестируем наш алгоритм на генерации изображений **CelebA** [41] ( $64 \times 64$ ). Для начала мы строим распределение над латентным пространством путем использования невариационного сверточного автокодировщика для кодирования изображений в латентные вектора размерности 128. Далее мы используем пару сетей DenseICNN для обучения циклически монотонного отображения для преобразования стандартного нормального шума в распределение над латентным пространством (рис. 2.8).

Мы используем сеть DenseICNN с параметрами [4; 256; 256; 128; 64] для обучения циклически монотонного генеративного отображения для преобразования стандартного нормального шума в распределение над латентным пространством. Для каждой задачи сеть была обучена для 100000 итераций с размером случайных выборок 128. Был использован оптимизатор Adam с  $lr = 3 \times 10^{-4}$ . В качестве параметра циклической регуляризации возьмем  $\lambda = 100$ .

На рис. 2.9 представлены изображения, сгенерированные непосредственно путем взятия выборки из стандартного нормального шума до (1-я строка) и после (2-я строка) применения нашего транспортного отображения. Наше генеративное отображение не приводит к кардинальным изменениям изображения, но разница до и после его применения хорошо видна визуально, а также подтверждается улучшением расстояния Fréchet Inception Distance (FID, см. [42]), см. табл. 3. Для сравнения мы также указываем полученные результаты для недавно введенной сети WGAN-QC в [43].

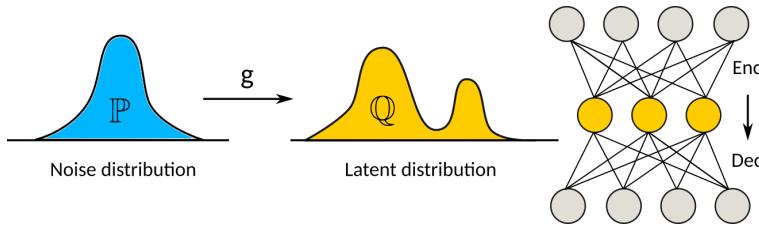


Рисунок 2.8 — Схема транспорта массы в латентном пространстве.



Рисунок 2.9 — Изображения, декодированные из стандартного латентного гауссовского шума (1-я строка) и из того же шума, перенесенного нашим циклическим монотонным отображением (2-я строка).

### 2.6.7 Перенос стилей между изображениями

В проблеме непарного переноса стилей алгоритм получает два набора изображений (каждый со своими атрибутами), например, каждый набор данных состоит из ландшафтов, относящихся к какому-либо сезону года. Целью является обучение отображения, которое должно быть способно к решению задачи **переноса атрибутов** из одного набора данных в другой, например, замене зимнего ландшафта на соответствующий летний ландшафт.

Мы проводили наши эксперименты с потенциалами ConvICNN на публично доступных<sup>5</sup> наборах данных **Winter2Summer** и **Photo2Cezanne** ( $256 \times 256$  pix). Наша модель обучалась на случайных выборках из 8 случайно обрезанных частей RGB-изображений размером  $128 \times 128$  пикселей. В качестве дополнительной аугментации мы используем случайные вращения ( $\pm \frac{\pi}{18}$ ), случайные отражения по горизонтали и небольшой гауссовский шум ( $\sigma = 0.01$ ). Сети были обучены 20000 итераций с циклической регуляризацией  $\lambda = 35000$ . Был использован оптимизатор Adam, вводился дополнительный  $10^{-1} \mathcal{L}^1$ -штраф для весов сетей. Наша схема перенесения стилей между наборами данных представлена на рис. 2.10. На рис. 2.11 и 2.12 представлены результаты для наборов данных Winter2Summer and Photo2Cezanne.

<sup>5</sup><https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix>

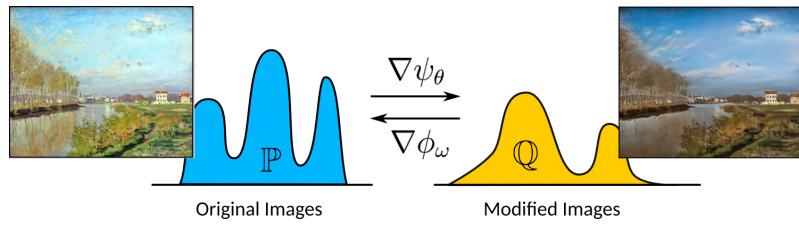


Рисунок 2.10 — Схема переноса стилей между наборами изображений с помощью пары ConvICNN обученных нашим методом.

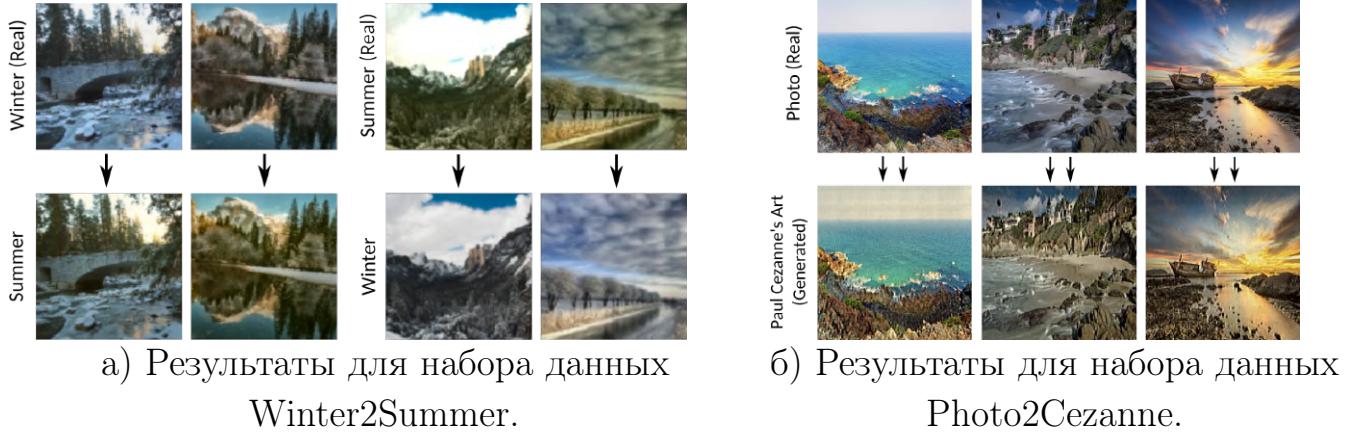
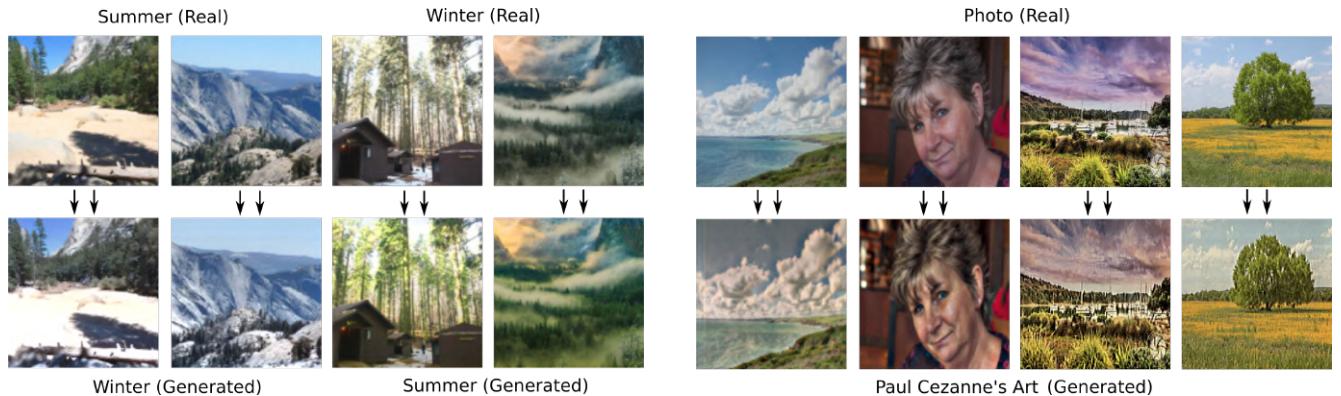


Рисунок 2.11 — Результаты переноса стиля между изображениями, полученными при использовании сети ConvICNN для изображений  $128 \times 128$  пикселей.

Наша генеративная модель обучает циклическое монотонное отображение. Однако требуемый перенос стилей может не быть циклически монотонным. Таким образом, наша модель может передавать только некоторые из требуемых атрибутов. Например, для передачи стиля “зима–лето” наша модель обучилась закрашивать деревья в зеленый цвет. Однако проблемным для нашей модели оказалось заполнение снежных покровов зеленой травой.

В работе [2] отмечается, что задача оптимального транспорта является перестановочно инвариантной. В ней не учитывается связь между размерностями, например, не учитываются окрестности пикселей или каналы в одном пикселе. Таким образом, теория оптимального транспорта старается обучить оптимальное генеративное отображение с использованием сверточной архитектуры (предназначенной для сохранения локальной структуры изображения).

Для того, чтобы обойти эту проблему, можно рассмотреть задачу оптимального транспорта для квадратичной функции стоимости, определенной на гауссовской пирамиде изображения (см. [44]) или аналогично перцепционным потерям, используемым для достижения сверхвысокого разрешения [45], рас-



а) Дополнительные результаты для набора данных Winter2Summer.

б) Дополнительные результаты для набора данных Photo2Cezanne.

Рисунок 2.12 — Дополнительные результаты переноса стиля между изображениями для наборов данных Winter2Summer and Photo2Monet.

смотреть перцептуальную квадратичную функцию стоимости. Это замечание указывает на одну из потенциальных тем дальнейших исследований.

## 2.7 Нейронные сети. Выпуклые Архитектуры

В п. 2.7.1, мы описываем общую архитектуру выпуклых по входу сетей. Затем мы описываем частные случаи общей архитектуры, используемой в дальнейших экспериментах: **DenseICNN** в п. 2.7.2 и **ConvICNN** в п. 2.7.3.

### 2.7.1 Общая архитектура с выпуклым входом

Мы аппроксимируем выпуклые потенциалы входными выпуклыми нейронными сетями [20]. Соответствующая общая архитектура схематично представлена на рис. 2.13.

Выпуклая по входу сеть состоит из двух основных блоков:

1. Линейный (L) блок состоит из линейных слоев. Функции активации и оператор пулинга в блоке также являются линейными, например, тождественная активация или пулинг средним значением.

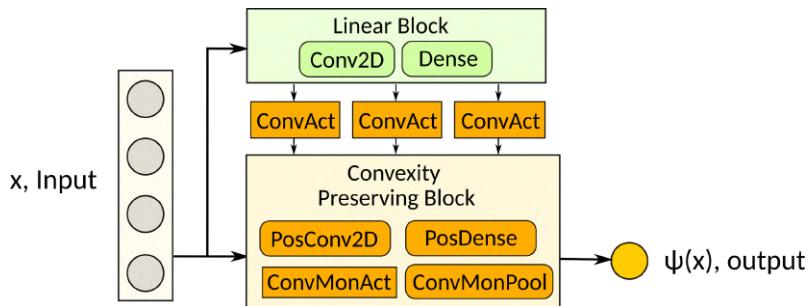


Рисунок 2.13 – Общая архитектура выпуклой по входу нейронной сети.

2. Блок сохранения выпуклости (CP) состоит из линейных слоев с неотрицательными весами (за исключением свободных членов). Операции активации и пулинг в этом блоке являются выпуклыми и монотонными. Внутри блоков возможно использовать произвольные обходные связи (skip connections), удовлетворяющие указанным правилам. Нейроны L-блока могут быть произвольно связаны с нейронами CP-блока путем применения выпуклой активации<sup>6</sup> и добавления положительного веса к полученному результату. Согласно правилам действий над выпуклыми функциями каждый нейрон (включая и выходной) в архитектуре, показанной на рис. 2.13, является выпуклой функцией входных данных<sup>7</sup>.

В нашем случае, мы ожидаем, что сеть будет способна легко обучать тождественное генеративное отображение

$$T(x) = \nabla \psi(x) = x,$$

т.е.  $\psi(x) = \frac{1}{2}\|x\|^2 + c$  – квадратичная функция. Таким образом, мы в основном используем квадратичные активации между L- и CP-блоками. Этот подход отличается от подхода [20], в котором такая активация не использовалась. Градиенты входных квадратичных функций соответствуют линейным деформациям входных функций и интуитивно весьма полезны как строительные блоки (в частности, для обучения тождественного отображения).

Мы используем специфичные архитектуры для обучения общей схемы, показанной на рис. 2.13. Потенциал **ConvICNN** используется для задач, свя-

<sup>6</sup>В отличие от активации в блоках, сохраняющими выпуклость, выпуклая активация между L- и CP-блоками может не быть монотонной, например, в качестве активации можно рассмотреть  $\sigma(x) = x^2$ .

<sup>7</sup>Можно использовать нормализацию батча и дропаут для L- и CP- блоков, а также между ними. Введение таких слоев не влияет на выпуклость, поскольку их можно рассматривать (при тестировании) как линейные слои с неотрицательными весами.

занных с обработкой изображений, а **DenseICNN** – в других целях. Более подробно эти архитектуры будут описаны в последующих разделах.

Функция CELU будет использована как выпуклая и монотонная активация (в СР-блоке) во всех сетях. Мы также рассмотрели SoftPlus среди некоторых других непрерывных и дифференцируемых функций, однако это дало отрицательный эффект на функционирование. Возможно также использование функции ReLU, но в этом случае градиент потенциала будет разрывным. Таким образом, он не будет липшицевым, поэтому наши теоремы 2.4.1 и 2.4.2 не могут быть применимы.

В качестве выпуклого и монотонного пулинга (в СР-блоке) возможно использование пулингов Average и LogSumExp (сглаженных max-пулингов). При этом следует избегать использования чистого пулинга Max по тем же причинам, что и активации ReLU. Однако в архитектуре ConvICNN мы используем свертки со страйдом (stride) вместо пулинга, см. п. 2.7.3.

Для того, чтобы воспользоваться теоремой 2.4.1, мы предполагаем, что потенциалы являются сильно выпуклыми и гладкими. Как отмечается в предыдущей секции,  $\mathcal{B}$ -гладкость выпуклой функции равносильна  $\frac{1}{\beta}$  сильной выпуклости ее сопряженной функции (и наоборот). Таким образом, мы считаем обе сети  $\psi_\theta, \varphi_\omega$   $\beta := \frac{1}{\beta}$  сильно выпуклыми, при этом циклическая регуляризация доставляет  $\lesssim \frac{1}{\beta} = \mathcal{B}$  гладкость для  $\nabla \psi_\theta \approx (\nabla \varphi_\omega)^{-1}$  и  $\nabla \varphi_\omega \approx (\nabla \psi_\theta)^{-1}$ . На практике мы достигаем сильной выпуклости путем добавления дополнительного значения  $\frac{\beta}{2} \|x\|^2$  к выходу финального нейрона сети. Во всех экспериментах мы полагаем  $\beta^{-1} = 1000000$ . Накладываемые условия гладкости и сильной выпуклости можно рассматривать в качестве регуляризации отображения: оно не дает слишком больших или малых деформаций входных данных. См., например, [32]. В дополнение к сглаживанию сильная выпуклость гарантирует, что  $\nabla \psi_\theta$  и  $\nabla \varphi_\omega$  биективны, что использовано в теоремах 2.4.1, 2.4.2.

### 2.7.2 Полносвязная нейронная сеть, выпуклая по входу

Для сети DenseICNN мы реализуем слой **Convex Quadratic**, в котором каждый выходной нейрон является выпуклой квадратичной функцией входных данных. Более точно, для каждого входа  $x \in \mathbb{R}^{N_{\text{in}}}$  его выходы

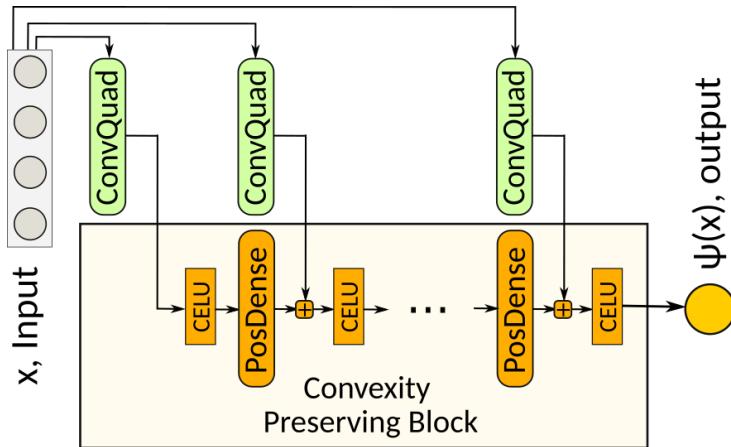


Рисунок 2.14 — Полносвязная выпуклая по входу нейронная сеть.

$$(cq_1(x), \dots, cq_{N_{\text{out}}}(x)) \in \mathbb{R}^{N_{\text{out}}} \text{ с}$$

$$cq_n(x) = \langle x, A_n x \rangle + \langle b_n, x \rangle + c_n$$

для неотрицательно определенной квадратичной формы  $A \in \mathbb{R}^{N_{\text{in}} \times N_{\text{in}}}$ , вектора  $b \in \mathbb{R}^{N_{\text{in}}}$  и константы  $c \in \mathbb{R}$ . Отметим, что при больших  $N_{\text{in}}$  размер такого слоя быстро возрастает, т.е.  $\geq O(N_{\text{in}}^2 \cdot N_{\text{out}})$ . Для того, чтобы обойти эту сложность, мы представим каждую квадратную матрицу как произведение  $A_n = F_n^T F_n$ , где  $F \in \mathbb{R}^{r \times N_{\text{in}}}$  — матрица ранга  $\leq r$ . Это помогает ограничить оптимизацию только положительными квадратичными формами (в частности, симметричными) и свести число весов в квадратичной части до порядка  $O(r \cdot N_{\text{in}} \cdot N_{\text{out}})$ . Полученные квадратичные формы  $A_n$  будут иметь ранг  $\leq r$ .

Соответствующая архитектура указана на рис. 2.14. Мы используем выпуклые квадратичные слои (Convex Quadratic Layers) для потенциалов DenseICNN для организации прямого соединения входа с промежуточными слоями полносвязной сети. Отметим, что такие слои (даже полного ранга) не сильно увеличивают размер сети при небольшой размерности входных данных (например, в задаче о передаче цветов).

Гиперпараметрами для DenseICNN являются ширины слоев и ранги квадратичных выпуклых по входу слоев. Для простоты мы используем тот же ранг  $r$  для всех слоев. Через  $h_0$  мы обозначим ширину первого выпуклого квадратичного слоя и ширину  $k+1$ -го выпуклого квадратичного слоя и  $k$ -го линейного слоя по  $h_k$ . Полный набор гиперпараметров сети имеет вид  $[r; h_0; h_1, \dots, h_K]$ .

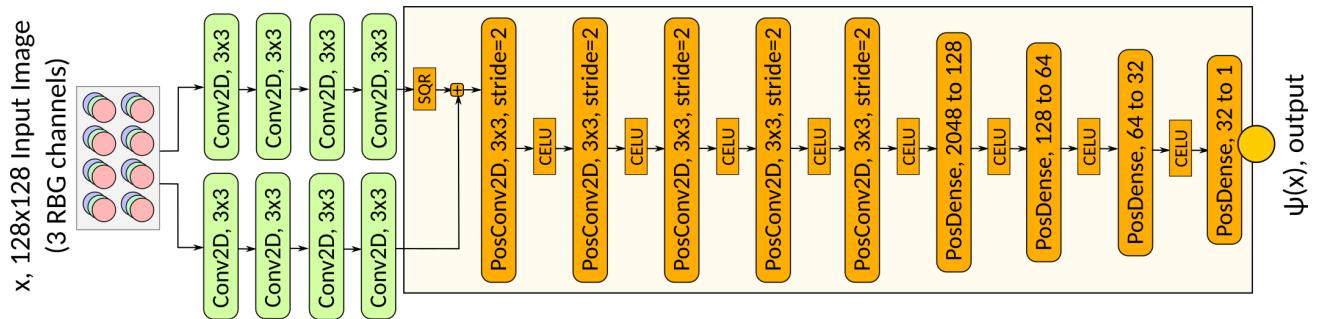


Рисунок 2.15 — Выпуклая нейронная сеть со сверточным входом. Все сверточные слои имеют 128 каналов.

### 2.7.3 Сверточная выпуклая по входу нейронная сеть

Мы применяем сверточные сети в проблеме непарного переноса стиля между изображениями. Архитектура сети ConvICNN показана на рис. 2.15. В качестве входного в сеть поступает изображение размером ( $128 \times 128$  с 3 каналами RGB), на выходе выдается одно значение. Градиент сети по входу рассматривается как генератор в нашем алгоритме.

В сети последовательно чередуются линейные блоки и блоки, сохраняющие выпуклость. При этом, связи с пропуском слоя не используются. Блок L состоит из двух отдельных частей с расположеннымными друг над другом свертками без промежуточной активации. Квадрат второй части добавляется к первой части и используется как вход для CR-блока. Все сверточные слои сети имеют 128 каналов (используется дополнение нулями со смещением = 1).

## 2.8 Обсуждение

В настоящей главе разработан одноэтапный алгоритм с неминимаксной целевой функцией для обучения циклически монотонных генеративных отображений, т.е. отображений оптимального транспорта для квадратичной функции стоимости. Дополнительно мы даем теоретическое обоснование нашего метода с точки зрения аппроксимации. Результаты вычислительных экспериментов подтверждают потенциал нашего алгоритма для различных практических задач: транспорт массы в латентном пространстве, перенос цветов и стилей между изображениями, адаптация домена.

## Глава 3. Метод для вычисления Васерштейн-2 барицентров

Геометрически барицентры Васерштейна суть взвешенные средние вероятностных мер в задаче об оптимальном транспорте. В настоящей главе мы представляем масштабируемый алгоритм для вычисления барицентров Васерштейна-2 по заданной выборке входных мер, которые не обязательно являются дискретными. В то время, как предыдущие методы для решения этой задачи использовали энтропийные или квадратичные регуляризации, мы используем выпуклые по входу нейронные сети и циклически согласованную регуляризацию, чтобы избежать введение смещения. Как результат наш подход не включает в себя минимаксную оптимизацию. Мы представляем теоретические результаты по оценкам ошибок, а также даем эмпирическое подтверждение эффективности предложенного подхода в случае маломерных качественных сценариев и для многомерных количественных экспериментов.

### 3.1 Введение

Барицентры Васерштейна приобрели свою популярность из-за способности к представлению усреднений вероятностных мер в геометрически содержательном смысле. Методы нахождения барицентров Васерштейна с успехом применяются во многих вычислительных задачах. В задачах обработки изображений барицентры Васерштейна используются для раскраски изображений и переноса стиля [37; 46], а также для синтеза текстур [47]. В задачах геометрической обработки изображений задача интерполяции форм может быть решена путем вычисления барицентров [48]. В задачах машинного обучения в реальном времени барицентры используются для агрегации вероятностных прогнозов экспертов [49]. В контексте байесовского вывода барицентр апостериорных распределений по подмножествам сходится к апостериорному распределению всего набора данных. На основе этого факта были получены эффективные вычислительные методы, основанные на нахождении барицентров (см. [50; 51]).

В настоящее время имеются алгоритмы быстрого и точного нахождения барицентра для дискретных распределений (см. обзор [6]). Однако для непре-

рывных распределений ситуация более сложна и продвижения в этой задаче появились лишь недавно (см. [8; 52; 53]). Стоит отметить, что дискретные методы обладают плохой способностью к масштабированию при увеличении количества точек носителя барицентра и не могут использоваться для хорошей аппроксимации непрерывных барицентров, особенно в случае больших размерностей.

В настоящей главе мы представляем метод нахождения барицентров Васерштейна-2 для *непрерывных* распределений. Наш метод основан на новой регуляризованной двойственной постановке, в которой выпуклые потенциалы параметризуются выпуклыми по входу нейронными сетями (см. [20]). Наш алгоритм является одноэтапным, не порождает смещения результатов (см., например, [8]) и не задействует минимаксную оптимизацию (см., например, [52]). Это достигается за счет комбинации нового *регуляризатора на конгруэнтность* с *циклически согласованной регуляризацией* (из главы 2). Мы показываем, что благодаря свойствам расстояния Васерштейна-2, градиенты получающихся выпуклых потенциалов “отображают” входные распределения близко к истинному барицентру, что позволяет строить хорошую аппроксимацию барицентра.

### 3.2 Предварительные сведения

Для полноты изложения в этой главе, в данной секции мы вновь напомним читателю основные свойства расстояния Васерштейн-2. Затем определим Васерштейн-2 барицентры, исследуемые в данной главе.

Напомним, что через  $\mathcal{P}_2(\mathbb{R}^D)$  мы обозначаем множество всех борелевских вероятностных мер на  $\mathbb{R}^D$  с конечным вторым моментом, а через  $\mathcal{P}_{2,\text{ac}}(\mathbb{R}^D) \subset \mathcal{P}_2(\mathbb{R}^D)$  – подмножество всех абсолютно непрерывных мер (относительно Лебеговой меры).

Для  $\mathbb{P}, \mathbb{Q} \in \mathcal{P}_2(\mathbb{R}^D)$  **расстояние Васерштейна-2** определяется следующим образом:

$$\mathbb{W}_2^2(\mathbb{P}, \mathbb{Q}) \stackrel{\text{def}}{=} \min_{\pi \in \Pi(\mathbb{P}, \mathbb{Q})} \int_{\mathbb{R}^D \times \mathbb{R}^D} \frac{\|x - y\|^2}{2} d\pi(x, y), \quad (3.1)$$

где  $\Pi(\mathbb{P}, \mathbb{Q})$  – множество вероятностных мер на  $\mathbb{R}^D \times \mathbb{R}^D$ , чьи маргиналы суть  $\mathbb{P}, \mathbb{Q}$  соответственно. (см. [17; 26]). Для дальнейших шагов нам снова понадобится

**двойственная постановка** (см. [27]):

$$\begin{aligned} \mathbb{W}_2^2(\mathbb{P}, \mathbb{Q}) &= \int_{\mathbb{R}^D} \frac{\|x\|^2}{2} d\mathbb{P}(x) + \int_{\mathbb{R}^D} \frac{\|y\|^2}{2} d\mathbb{Q}(y) - \\ &\min_{\psi \in \text{Convex}} \left[ \int_{\mathbb{R}^D} \psi(x) d\mathbb{P}(x) + \int_{\mathbb{R}^D} \bar{\psi}(y) d\mathbb{Q}(y) \right], \end{aligned} \quad (3.2)$$

где минимум берется по всем выпуклым функциям (потенциалам)  $\psi : \mathbb{R}^D \rightarrow \mathbb{R} \cup \{\infty\}$  и  $\bar{\psi}(y) = \max_{x \in \mathbb{R}^D} (\langle x, y \rangle - \psi(x)) : \mathbb{R}^D \rightarrow \mathbb{R} \cup \{\infty\}$  – **выпуклая сопряженная функция** к  $\psi$  [29], которая также является выпуклой функцией. Здесь оптимальный потенциал  $\psi^*$  определяется с точностью до аддитивной константы.

В работе [26] показано, что если  $\mathbb{P} \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$ , то оптимальный план  $\pi$  единственным образом определяется в виде  $\pi^* = [\text{id}_{\mathbb{R}^D}, T^*] \sharp \mathbb{P}$ , где  $T^* : \mathbb{R}^D \rightarrow \mathbb{R}^D$  – единственное решение задачи Монже:

$$T^* = \arg \min_{T \sharp \mathbb{P} = \mathbb{Q}} \int_{\mathbb{R}^D} \frac{\|x - T(x)\|^2}{2} d\mathbb{P}(x). \quad (3.3)$$

Между  $T^*$  и двойственной задачей (3.2) имеется следующая связь:  $T^* = \nabla \psi^*$ , где  $\psi^*$  – оптимальное решение задачи (3.2). Дополнительно, если  $\mathbb{Q} \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$ , то отображение  $T^*$  обратимо и

$$T^*(x) = \nabla \psi^*(x) = (\nabla \bar{\psi}^*)^{-1}(x), \quad (T^*)^{-1}(y) = \nabla \bar{\psi}^*(y) = (\nabla \psi^*)^{-1}(y).$$

**Барицентр Васерштейна–2.** Пусть  $\mathbb{P}_1, \dots, \mathbb{P}_N \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$ . Тогда их барицентр относительно весов  $\alpha_1, \dots, \alpha_N$  ( $\alpha_n > 0$  и  $\sum_{n=1}^N \alpha_n = 1$ ) определяется следующим образом:

$$\overline{\mathbb{P}} \stackrel{\text{def}}{=} \arg \min_{\mathbb{P} \in \mathcal{P}_2(\mathbb{R}^D)} \sum_{n=1}^N \alpha_n \mathbb{W}_2^2(\mathbb{P}_n, \mathbb{P}). \quad (3.4)$$

На протяжении настоящей главы мы предполагаем, что по крайней мере одно из распределений  $\mathbb{P}_1, \dots, \mathbb{P}_N \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$  имеет ограниченную плотность. При таком предположении распределение  $\overline{\mathbb{P}}$  единственно и абсолютно непрерывно, т.е.  $\overline{\mathbb{P}} \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$  и имеет ограниченную плотность (см. определение 3.6 и теорему 5.1 в [4]).

Для  $n \in \{1, 2, \dots, N\}$  пусть  $(\psi_n^*, \bar{\psi}_n^*)$  – оптимальная пара (взаимно) сопряженных потенциалов, которые отображают  $\mathbb{P}_n$  в  $\overline{\mathbb{P}}$ , т.е.  $\nabla \psi_n^* \sharp \mathbb{P}_n = \overline{\mathbb{P}}$  и  $\nabla \bar{\psi}_n^* \sharp \overline{\mathbb{P}} = \mathbb{P}_n$ . Тогда  $\{\bar{\psi}_n^*\}$  удовлетворяют уравнениям

$$\sum_{n=1}^N \alpha_n \nabla \bar{\psi}_n^*(x) = x \quad \text{и} \quad \sum_{n=1}^N \alpha_n \bar{\psi}_n^*(x) = \frac{\|x\|^2}{2} + c \quad (3.5)$$

при всех  $x \in \mathbb{R}^D$  [4; 36]. Так как оптимальные потенциалы определены с точностью до константы, для удобства мы полагаем  $c = 0$ . Условие (3.5) является базисным для нашего алгоритма вычисления барицентров Васерштейна-2. Мы говорим, что потенциалы  $\psi_1, \dots, \psi_N$  **конгруэнтны** относительно весов  $\alpha_1, \dots, \alpha_n$ , если их сопряженные потенциалы удовлетворяют условию (3.5), т.е. если  $\sum_{n=1}^D \alpha_n \overline{\psi_n}(x) = \frac{\|x\|^2}{2}$  при всех  $x \in \mathbb{R}^D$ .

### 3.3 Обзор существующих методов

Большинство вычислительных алгоритмов оптимальной транспортировки разработаны для дискретных постановок задач, в которых входные распределения имеют конечный носитель (см. недавний обзор [6]). Особенno популярным является семейство алгоритмов, основанных на энтропийной регуляризации. На их основе был разработан известный итерационный метод Синкхорна (Sinkhorn) [54; 55]. Использование таких методов обычно ограничивается носителем порядка  $10^5 - 10^6$  точек, поскольку при большем числе точек задача вычислительно не решается. Аналогично, дискретные методы нахождения барицентра [55] и особенно методы, основанные на фиксированном носителе барицентра [56; 57], не могут выдавать точную аппроксимацию непрерывных барицентров для больших размерностей, поскольку в этом случае требуется очень большой дискретный носитель (см., например, экспериментальные исследования в п. 4.3 работы [52]). Таким образом, мы дадим краткий обзор литературы только для непрерывного случая.

**Вычисление расстояния Васерштейна-2 и отображений.** В работе [58] демонстрируется возможность вычисления расстояний Васерштейна при наличии только выборочного доступа к распределениям путем параметризации двойственных потенциалов с помощью ядерных разложений. В работе [2] предлагается аналогичный метод с использованием нейронных сетей для параметризации потенциалов и использования энтропийной или  $\mathcal{L}^2$  регуляризации относительно  $\mathbb{P} \times \mathbb{Q}$  для сохранения приближенной сопряженности потенциалов. При этом транспортное отображение восстанавливается из оптимизированных потенциалов с помощью барицентрической проекции.

Как было отмечено в п. 3.2, расстояние  $\mathbb{W}_2$  обладает рядом полезных теоретических свойств. Например, оптимальный потенциал  $\psi^*$  является выпуклым. При этом соответствующее отображение оптимального транспорта задается в виде  $\nabla\psi^*$ . В работе [7] с использованием этих свойств был предложен алгоритм минимаксной оптимизации для восстановления транспортных отображений для выпуклых по входу нейронных сетей (ICNN) [20] в задаче аппроксимации потенциалов.

Альтернативой энтропийной регуляризации служит **циклически согласованная** регуляризация, предложенная в главе 2 настоящей диссертационной работы. Напомним, что в этом методе используется то свойство, что градиенты оптимальных двойственных потенциалов являются взаимно обратными. В таком регуляризаторе требуется интегрирование только по маргинальным мерам  $\mathbb{P}$  и  $\mathbb{Q}$ , а не по  $\mathbb{P} \times \mathbb{Q}$ , как требуется в альтернативах, основанных на энтропии. Такой метод сходится быстрее, чем минимаксный метод, поскольку в нем не задействуется внутренний оптимизационный цикл.

В работе [59] предлагается использование двух генеративных моделей, с общим латентным пространством для неявного вычисления оптимального транспортного соответствия между  $\mathbb{P}$  и  $\mathbb{Q}$ . Основываясь на полученном соответствии, авторы вычисляют оптимальное транспортное расстояние между распределениями.

**Вычисление барицентров Васерштейна-2.** В ряде недавних работ рассматривается задача нахождения барицентра (3.4) с использованием непрерывных (а не дискретных) аппроксимаций барицентра:

- **Оптимизация, основанная на мерах (генеративная оптимизация).** Задача (3.4) оптимизируется по вероятностным мерам. Для этой цели можно использовать общий алгоритм из работы [53] с применением генеративных сетей для вычисления барицентров. Этому подходу присущи обычные проблемы генеративных моделей, например, коллапс мод. Применение этого метода к задаче о  $\mathbb{W}_2$ -барицентрах требует оценки  $\mathbb{W}_2^2(\mathbb{P}_n, \mathbb{P})$ . В работе [52] этот подход был протестирован с использованием минимаксного метода из [7]. Однако в этой работе авторы получили труднорешаемую *min-max-min*-задачу.
- **Оптимизация, основанная на потенциалах:** в работе [8] оптимальные потенциалы восстанавливаются  $\{\psi_n^*\}$  через решение неминимаксной

регуляризованной задачи. При этом генеративная модель не требуется: барицентр восстанавливается через отображения прямых мер с использованием градиентов потенциалов или барицентрической проекции.

### 3.4 Одноэтапный неминимаксный алгоритм

Следуя [8], мы используем подход, основанный на потенциалах и восстанавливаем барицентр с использованием градиентов потенциалов как транспортных отображений. Основные отличия нашего подхода: (1) мы рассматриваем только выпуклые потенциалы; (2) конгруэнтность достигается через слагаемое регуляризации; (3) в нашей постановке задачи не вводится смещение, т.е. оптимальное решение нашей задачи дает истинный барицентр.

#### 3.4.1 Вывод двойственной задачи

Пусть  $\bar{\mathbb{P}}$  – истинный барицентр. Нашей целью является восстановление оптимальных потенциалов  $\{\psi_n^*, \bar{\psi}_n^*\}$ , отображающих входные меры  $\mathbb{P}_n$  в  $\bar{\mathbb{P}}$ .

Для начала мы выразим задачу поиска барицентра (3.4) через двойственную задачу (3.2):

$$\sum_{n=1}^N \alpha_n \mathbb{W}_2^2(\mathbb{P}_n, \bar{\mathbb{P}}) = \left[ \sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} \frac{\|x\|^2}{2} d\mathbb{P}_n(x) \right] + \int_{\mathbb{R}^D} \frac{\|y\|^2}{2} d\bar{\mathbb{P}}(y) - \\ \min_{\{\psi_n\} \in \text{Convex}} \left[ \sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} \psi_n(x) d\mathbb{P}_n(x) + \sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} \bar{\psi}_n(y) d\bar{\mathbb{P}}(y) \right] \quad (3.6)$$

Здесь минимум достигается не только среди выпуклых потенциалов  $\{\psi_n\}$ , но и среди конгруэнтных потенциалов (см. рассуждения после формулы (3.5)); таким образом, мы можем добавить ограничение, что  $\{\psi_n\}$  конгруэнтны (3.6). Как следствие имеем

$$\sum_{n=1}^N \alpha_n \mathbb{W}_2^2(\mathbb{P}_n, \bar{\mathbb{P}}) = \left[ \sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} \frac{\|x\|^2}{2} d\mathbb{P}_n(x) \right] -$$

$$\min_{\{\psi_n\} \text{ конгруэнтны}} \left[ \underbrace{\sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} \psi_n(y) d\mathbb{P}_n(y)}_{\text{MultiCorr}(\{\alpha_n, \mathbb{P}_n\} | \{\psi_n\})} \right]. \quad (3.7)$$

Для перехода от (3.6) к (3.7) мы использовали тот факт, что для конгруэнтных  $\{\psi_n\}$  выполнено  $\sum_{n=1}^N \alpha_n \overline{\psi_n}(x) = \frac{\|x\|^2}{2}$ , откуда  $\sum_{n=1}^N \int_{\mathbb{R}^D} \alpha_n \overline{\psi_n}(y) d\bar{\mathbb{P}}(y) = \int_{\mathbb{R}^D} \frac{\|y\|^2}{2} d\bar{\mathbb{P}}(y)$ .

Величину под знаком минимума в (3.7) мы будем называть **множественной корреляцией**  $\{\mathbb{P}_n\}$  с весами  $\{\alpha_n\}$  относительно потенциалов  $\{\psi_n\}$ . Отметим, что истинный барицентр  $\bar{\mathbb{P}}$  нигде не появляется в правой части (3.7). Таким образом, оптимальные потенциалы  $\{\psi_n^*\}$  могут быть восстановлены путем решения следующей задачи:

$$\min_{\{\psi_n\} \text{ конгруэнтны}} \text{MultiCorr}(\{\alpha_n, \mathbb{P}_n\} | \{\psi_n\}) = \min_{\{\psi_n\} \text{ конгруэнтны}} \left[ \sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} \psi_n(y) d\mathbb{P}_n(y) \right]. \quad (3.8)$$

### 3.4.2 Условие конгруэнтности

Как можно наложить условие конгруэнтности на потенциалы на практике? Наша следующая теорема дает оценку близости множества выпуклых потенциалов  $\{\psi_n\}$  к  $\{\psi_n^*\}$  в терминах отличия множественной корреляции.

**Теорема 3.4.1.** Пусть  $\bar{\mathbb{P}} \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$  – барицентр  $\mathbb{P}_1, \dots, \mathbb{P}_N \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$  с весами  $\alpha_1, \dots, \alpha_N$  и пусть  $\{\psi_n^*\}$  – оптимальные конгруэнтные потенциалы задачи нахождения барицентра. Предположим, что нам заданы  $\mathcal{B}$ -гладкие<sup>1</sup> выпуклые потенциалы  $\{\psi_n\}$  при некотором  $\mathcal{B} \in (0, +\infty)$ . Положим

$$\Delta = \text{MultiCorr}(\{\alpha_n, \mathbb{P}_n\} | \{\psi_n\}) - \text{MultiCorr}(\{\alpha_n, \mathbb{P}_n\} | \{\psi_n^*\}).$$

---

<sup>1</sup>Напомним, что дифференцируемая функция  $f : \mathbb{R}^D \rightarrow \mathbb{R}$  является  $\mathcal{B}$ -гладкой, если ее градиент  $\nabla f$  является  $\mathcal{B}$ -липшицевым.

Тогда выполняется

$$\Delta + \underbrace{\int_{\mathbb{R}^D} \sum_{n=1}^N [\alpha_n \overline{\psi}_n(y) - \frac{\|y\|^2}{2}] d\bar{\mathbb{P}}(y)}_{\text{Дисбаланс конгруэнтности}} \geq \frac{1}{2\mathcal{B}} \sum_{n=1}^N \alpha_n \|\nabla \psi_n^*(x) - \nabla \psi_n(x)\|_{\mathbb{P}_n}^2. \quad (3.9)$$

Здесь  $\|\cdot\|_\mu$  – норма, индуцированная скалярным произведением в гильбертовом пространстве  $\mathcal{L}^2(\mathbb{R}^D \rightarrow \mathbb{R}^D, \mu)$ .

Второе слагаемое в левой части (3.9) мы будем называть **дисбалансом конгруэнтности**. Мы докажем этот результат в п. 3.5. Отметим, что если дисбаланс конгруэнтности неположителен, то

$$\Delta \geq \frac{1}{2\mathcal{B}} \sum_{n=1}^N \alpha_n \|\nabla \psi_n^*(x) - \nabla \psi_n(x)\|_{\mathbb{P}_n}^2 \geq \frac{1}{\mathcal{B}} \sum_{n=1}^N \alpha_n \mathbb{W}_2^2(\nabla \psi_n \# \mathbb{P}_n, \bar{\mathbb{P}}), \quad (3.10)$$

где последнее неравенство в (3.10) обеспечивается леммой 2.5.1. Из (3.10) мы заключаем, что  $\mathbb{W}_2^2(\nabla \psi_n \# \mathbb{P}_n, \bar{\mathbb{P}}) \leq \frac{\mathcal{B}\Delta}{\alpha_n}$  при всех  $n \in \{1, \dots, N\}$ . Это показывает, что если дисбаланс конгруэнтности неположителен, то  $\Delta$  (различие во множественных корреляциях) дает оценку сверху для расстояния Васерштейна-2 между истинным барицентром и отображением  $\nabla \psi_n \# \mathbb{P}_n$  для всех  $n$ . Это делает правомерным использование  $\nabla \psi_n \# \mathbb{P}_n$  для восстановления барицентра. Отметим, что для оптимальных потенциалов дисбаланс конгруэнтности равен нулю.

Таким образом, для штрафа положительного дисбаланса конгруэнтности мы введем следующее слагаемое регуляризации:

$$\mathcal{R}_1^{\bar{\mathbb{P}}}(\{\alpha_n\}, \{\overline{\psi}_n\}) \stackrel{\text{def}}{=} \int_{\mathbb{R}^D} \left[ \sum_{n=1}^N \alpha_n \overline{\psi}_n(y) - \frac{\|y\|^2}{2} \right]_+ d\bar{\mathbb{P}}(y). \quad (3.11)$$

Так как мы берем положительную часть от подынтегрального выражения в (3.9) для получения (3.11) и так как правая часть в (3.9) неотрицательна, имеем:

$$[\text{MultiCorr}(\{\alpha_n, \mathbb{P}_n\} | \{\psi_n\}) + 1 \cdot \mathcal{R}_1^{\bar{\mathbb{P}}}(\{\alpha_n\}, \{\overline{\psi}_n\})] - \text{MultiCorr}(\{\alpha_n, \mathbb{P}_n\} | \{\psi_n^*\}) \geq 0$$

при всех выпуклых потенциалах  $\{\psi_n\}$ . С другой стороны, для оптимальных потенциалов  $\{\psi_n\} = \{\psi_n^*\}$  неравенство должно превращаться в равенство. Это влечет то, что добавление слагаемого регуляризации  $1 \cdot \mathcal{R}_1^{\bar{\mathbb{P}}}(\{\alpha_n\}, \{\overline{\psi}_n\})$  в (3.8) не приводит к появлению смещенностии – оптимальное решение по-прежнему задается через  $\{\psi_n^*\}$ .

Однако для точного вычисления (3.11) требуется априорное знание истинного барицентра  $\bar{\mathbb{P}}$ . Для преодоления этой трудности мы можем заменить  $\bar{\mathbb{P}}$  на другую абсолютно непрерывную меру  $\tau \cdot \hat{\mathbb{P}}$  ( $\tau \geq 1$  и  $\hat{\mathbb{P}}$  – вероятностная мера), чья плотность оценивает плотность  $\bar{\mathbb{P}}$  сверху почти всюду. В этом случае имеем

$$\tau \cdot \mathcal{R}_1^{\hat{\mathbb{P}}}(\{\alpha_n\}, \{\bar{\Psi}_n\}) = \tau \cdot \int_{\mathbb{R}^D} \left[ \sum_{n=1}^N \alpha_n \bar{\Psi}_n(y) - \frac{\|y\|^2}{2} \right]_+ d\hat{\mathbb{P}} \geq \mathcal{R}_1^{\bar{\mathbb{P}}}(\{\alpha_n\}, \{\bar{\Psi}_n\}). \quad (3.12)$$

Как следствие мы получаем следующий регуляризованный вариант (3.8) (где  $\{\Psi_n^*\}$  – оптимальное решение задачи):

$$\min_{\{\Psi_n\} \in \text{Convex}} [\text{MultiCorr}(\{\alpha_n, \mathbb{P}_n\} | \{\Psi_n\}) + \tau \cdot \mathcal{R}_1^{\hat{\mathbb{P}}}(\{\alpha_n\}, \{\bar{\Psi}_n\})]. \quad (3.13)$$

Однако выбор меры  $\tau \cdot \hat{\mathbb{P}}$  не столь очевиден. Рассмотрим случай, когда носители  $\{\mathbb{P}_n\}$  содержатся в компактных множествах  $\mathcal{X}_1, \dots, \mathcal{X}_N \subset \mathbb{R}^D$  и плотность для  $\mathbb{P}_1$  оценивается сверху величиной  $h < \infty$ . В этом случае плотность барицентра оценивается сверху величиной  $h \cdot \alpha_1^{-D}$  (см. замечание 3.2 работы [36]). Таким образом, мера  $\tau \cdot \hat{\mathbb{P}}$  с носителем  $\text{ConvexHull}(\mathcal{X}_1, \dots, \mathcal{X}_N)$  с такой плотностью дает оценку сверху на  $\bar{\mathbb{P}}$ . Вопрос корректного выбора  $\tau, \hat{\mathbb{P}}$  в практических задачах будет рассмотрен ниже в п. 3.4.4.

### 3.4.3 Выполнение условия сопряженности для пар потенциалов

В этом параграфе мы будем предполагать известной конечную меру  $\tau \cdot \hat{\mathbb{P}}$ , дающую оценку сверху для  $\bar{\mathbb{P}}$ . Оптимизационная задача (3.13) включает в себя не только потенциалы  $\{\Psi_n\}$ , но также и их сопряженные величины  $\{\bar{\Psi}_n\}$ . Это приводит к сложностям в ее практической реализации, поскольку задача вычисления сопряженных потенциалов является трудной (см. главу 2).

Мы рассмотрим потенциалы  $\Psi_n$  и  $\bar{\Psi}_n$ , как две отдельные выпуклые функции и обозначим их через  $\psi_n$  и  $\varphi_n$  соответственно. Следуя главе 2, для наложения условия сопряжения на  $\Psi_n$  и  $\varphi_n$  добавим дополнительный циклически согласованный регуляризатор

$$\mathcal{R}_2^{\mathbb{P}_n}(\Psi_n, \varphi_n) \stackrel{\text{def}}{=} \int_{\mathbb{R}^D} \|\nabla \varphi_n(\nabla \Psi_n(x)) - x\|_2^2 d\mathbb{P}_n(x) = \|\nabla \varphi_n \circ \nabla \Psi_n - \text{id}_{\mathbb{R}^D}\|_{\mathbb{P}_n}^2.$$

Отметим, что  $\mathcal{R}_2^{\mathbb{P}^n}(\psi_n, \varphi_n) = 0$  является необходимым условием для того, чтобы  $\psi_n$  и  $\varphi_n$  были взаимно сопряженными. Также это условие является достаточным условием для того, чтобы выпуклые функции были сопряженными с точностью до аддитивной константы.

Мы будем использовать одностороннюю регуляризацию. В нашем случае вычисление регуляризатора другого направления  $\|\nabla\psi_n \circ \nabla\varphi_n - \text{id}_{\mathbb{R}^D}\|_{\mathbb{P}}^2$  технически невозможно, поскольку мы не знаем  $\bar{\mathbb{P}}$ . В главе 2 мы наблюдали, что такое одностороннее условие является достаточным.

На этом пути мы используем  $2N$  выпуклых функций для  $\{\psi_n, \varphi_n\}$ . Путем добавления нового циклически состоятельного регуляризатора в (3.13), получаем нашу окончательную целевую задачу:

$$\min_{\{\psi_n, \varphi_n\}} \underbrace{\sum_{n=1}^N \left[ \alpha_n \int_{\mathbb{R}^D} [\underbrace{\langle x, \nabla\psi_n(x) \rangle - \varphi_n(\nabla\psi_n(x))}_{\approx \bar{\varphi}_n(x)}] d\mathbb{P}_n(x) \right]}_{\text{Приближенная множественная корреляция}} + \underbrace{\tau \cdot \mathcal{R}_1^{\hat{\mathbb{P}}}(\{\varphi_n\})}_{\text{Конгр. рег..}} + \underbrace{\lambda \sum_{n=1}^N \alpha_n \mathcal{R}_2^{\mathbb{P}^n}(\psi_n, \varphi_n)}_{\text{Циклический регуляризатор}}. \quad (3.14)$$

Отметим, что мы выражаем приближенную множественную корреляцию через оба потенциала  $\{\psi_n\}$  и  $\{\varphi_n\}$ . За счет этого мы избавляемся от произвольности в зависимости аддитивной константы от  $\{\psi_n\}$  (этого нельзя достичь циклической регуляризацией, см. главу 2). Обозначим через  $\text{MultiCorr}(\{\mathbb{P}_n\} | \{\psi\}, \{\varphi\}; \tau, \hat{\mathbb{P}}, \lambda)$  всю целевую функцию. Аналогично теореме 3.4.1 имеет место следующий результат, который показывает, что эта новая целевая функция имеет те же свойства, как и ее нерегуляризованный вариант из (3.8).

**Теорема 3.4.2.** *Пусть  $\bar{\mathbb{P}} \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$  – барицентр для  $\mathbb{P}_1, \dots, \mathbb{P}_N \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$  с весами  $\alpha_1, \dots, \alpha_N$  и пусть  $\{\psi_n^*\}$  – оптимальные конгруэнтные потенциалы для задачи о барицентре. Предположим, что  $\tau, \hat{\mathbb{P}}$  таковы, что  $\tau \geq 1$  и  $\tau \cdot \hat{\mathbb{P}} \geq \bar{\mathbb{P}}$ . Пусть нам также даны выпуклые потенциалы  $\{\hat{\psi}_n\}$  и  $\beta$ -сильно выпуклые и  $\mathcal{B}$ -гладкие выпуклые потенциалы  $\{\hat{\varphi}_n\}$  с  $0 < \beta \leq \mathcal{B} < \infty$  и  $\lambda > \frac{\beta}{2\beta^2}$ . Тогда*

$$\text{MultiCorr}(\{\alpha_n, \mathbb{P}_n\} | \{\hat{\psi}_n\}, \{\hat{\varphi}_n\}; \tau, \hat{\mathbb{P}}, \lambda) \geq \text{MultiCorr}(\{\alpha_n, \mathbb{P}_n\} | \{\psi_n^*\}). \quad (3.15)$$

Положим

$$\Delta = \text{MultiCorr}(\{\alpha_n, \mathbb{P}_n\} | \{\hat{\psi}_n\}, \{\hat{\varphi}_n\}; \tau, \hat{\mathbb{P}}, \lambda) - \text{MultiCorr}(\{\alpha_n, \mathbb{P}_n\} | \{\psi_n^*\}).$$

Тогда  $\Delta \geq 0$  и при всех  $n \in \{1, \dots, N\}$  имеем

$$\mathbb{W}_2^2(\nabla \hat{\psi}_n \sharp \mathbb{P}_n, \bar{\mathbb{P}}) \leq \frac{\Delta}{\alpha_n} \cdot \left( \sqrt{\frac{1}{\beta}} + \sqrt{\frac{1}{\lambda \beta^2 - \frac{\mathcal{B}}{2}}} \right)^2 = O(\Delta). \quad (3.16)$$

Неформально говоря, теорема 3.4.2 утверждает, что чем лучше мы решаем регуляризованную двойственную задачу (3.14), тем более близким можно ожидать любое  $\nabla \hat{\psi}_n \sharp \mathbb{P}_n$  к истинному барицентру  $\bar{\mathbb{P}}$  в  $\mathbb{W}_2$ . Из (3.15) следует, что наша окончательная целевая функция (3.14) является *несмешенной*, т.е. оптимальное решение получается при  $\{\Psi_n^*, \bar{\Psi}_n^*\}$ .

### 3.4.4 Алгоритм. Практическая реализация. Оптимизация

Для параметризации потенциалов  $\{\psi_n, \varphi_n\}$  мы используем сети DenseICNN  $\{\psi_{\theta_n}, \varphi_{\omega_n}\}$  (полносвязные выпуклые по входу нейросети) с квадратичными обходными связями, см. п. 2.7. В качестве начального шага мы предварительно обучаем потенциалы так, чтобы выполнялись следующие условия:

$$\psi_{\theta_n}(x) \approx \frac{\|x\|^2}{2} \quad \text{и} \quad \varphi_{\omega_n}(y) \approx \frac{\|y\|^2}{2}.$$

Такое предварительное обучение дает хороший старт для сетей: каждый потенциал  $\psi_{\theta_n}$  является приближенно сопряженным к соответствующему  $\varphi_{\omega_n}$ . С другой стороны, начальные сети  $\{\psi_{\theta_n}\}$  приближенно конгруэнтны в силу (3.5).

Мы используем стохастический метод градиентного спуска на случайной выборке для решения (3.14), где интегрирование выполняется при помощи выборок по методу Монте–Карло по входным мерам  $\{\mathbb{P}_n\}$  и мере регуляризации  $\hat{\mathbb{P}}$  аналогично [8]. Мы даем подробное описание оптимизационной процедуры (алгоритм 2).

Оказывается, что на практике, даже если условие  $\tau \cdot \hat{\mathbb{P}} \geq \bar{\mathbb{P}}$  не удовлетворяется при выборе  $\tau, \hat{\mathbb{P}}$ , часто получается, что меры отображения  $\nabla \psi_{\theta_n} \sharp \mathbb{P}_n$  сходятся к  $\bar{\mathbb{P}}$ . Для того, чтобы частично устраниТЬ разрыв между теорией и практикой, мы *динамически* корректируем меру  $\hat{\mathbb{P}}$  таким образом, что после

---

**Алгоритм 2:** Численная процедура для оптимизации множественных корреляций (3.14)

---

**Вход:** Распределения  $\mathbb{P}_1, \dots, \mathbb{P}_N$  с выборочным доступом;

Веса  $\alpha_1, \dots, \alpha_N \geq 0$  с  $\sum_{n=1}^N \alpha_n = 1$ ;

Распределение регуляризации  $\widehat{\mathbb{P}}$  с выборочным доступом;

Коэффициент конгруэнтности на согласованность  $\tau \geq 1$ ;

Балансировочный коэффициент  $\gamma \in [0, 1]$ ;

Коэффициент циклической согласованности  $\lambda > 0$ ;

$2N$  ICNN  $\{\psi_{\theta_n}, \varphi_{\omega_n}\}$ ; Размер случайных выборок  $K > 0$ ;

**for**  $t = 1, 2, \dots$  **do**

1. Получение выборки  $X_n \sim \mathbb{P}_n$  при всех  $n = 1, \dots, N$ ;
2. Вычисление отображений  $Y_n = \nabla \psi_{\theta_n} \# X_n$  при всех  $n = 1, \dots, N$ ;
3. Получение выборки  $Y_0 \sim \widehat{\mathbb{P}}$ ;
4. Вычисление оценки по методу Монте–Карло для регуляризатора конгруэнтности:

$$\mathcal{L}_{\text{Congruence}} := \frac{1}{K} \cdot \sum_{n=1}^N \gamma_n \sum_{y \in Y_n} \left[ \sum_{n'=1}^N \alpha_{n'} \varphi_{\omega_{n'}}(y) - \frac{\|y\|^2}{2} \right]_+,$$

где  $\gamma_0 = \gamma$  и  $\gamma_n = \alpha_n \cdot (1 - \gamma)$  для  $n = 1, 2, \dots, N$ ;

5. Вычисление оценки по методу Монте–Карло для циклически согласованного регуляризатора:

$$\mathcal{L}_{\text{Cycle}} := \frac{1}{K} \sum_{n=1}^N \alpha_n \left[ \sum_{x \in X_n} \|\nabla \varphi_{\omega_n}(\nabla \psi_{\theta_n}(x)) - x\|_2^2 \right];$$

6. Вычисление оценки по методу Монте–Карло для множественных корреляций:

$$\mathcal{L}_{\text{MultiCorr}} := \sum_{n=1}^N \left[ \alpha_n \cdot \frac{1}{K} \sum_{x \in X_n} [\langle x, \nabla \psi_{\theta_n}(x) \rangle - \varphi_{\omega_n}(\nabla \psi_{\theta_n}(x))] \right];$$

7. Вычисление полных потерь:

$$\mathcal{L}_{\text{Total}} := \mathcal{L}_{\text{MultiCorr}} + \lambda \cdot \mathcal{L}_{\text{Cycle}} + \tau \cdot \mathcal{L}_{\text{Congruence}};$$

8. Выполнение градиентного шага по  $\{\theta_n, \omega_n\}$  с использованием  $\frac{\partial \mathcal{L}_{\text{Total}}}{\partial \{\theta_n, \omega_n\}}$ ;

**end**

---

каждого оптимизационного шага мы полагаем (при  $\gamma \in [0,1]$ )

$$\widehat{\mathbb{P}}' := \gamma \cdot \widehat{\mathbb{P}} + (1 - \gamma) \cdot \sum_{n=1}^N \alpha_n \cdot [\nabla \psi_{\theta_n} \sharp \mathbb{P}_n],$$

т.е. вероятностная мера  $\widehat{\mathbb{P}}'$  является смесью заданной начальной меры  $\widehat{\mathbb{P}}$  и текущей оценки барицентра  $\{\nabla \psi_{\theta_n} \sharp \mathbb{P}_n\}$ . Для начального  $\widehat{\mathbb{P}}$  можно использовать барицентр  $\{\mathcal{N}(\mu_{\mathbb{P}_n}, \Sigma_{\mathbb{P}_n})\}$ . Его можно эффективно вычислить, используя итерационный алгоритм, основанный на поиске неподвижной точки (см. [36; 60]). При оптимизации эти оценки становятся ближе к истинному барицентру, что, в свою очередь, может улучшить регуляризатор конгруэнтности (3.12).

### 3.5 Доказательства теоретических результатов

В этом параграфе мы доказываем наши основные теоремы 3.4.1 и 3.4.2. Напомним, что через  $\mathcal{L}^2(\mathbb{R}^D \rightarrow \mathbb{R}^D, \mu)$  мы обозначаем **гильбертово пространство** функций  $f : \mathbb{R}^D \rightarrow \mathbb{R}^D$ , интегрируемых с квадратом по вероятностной мере  $\mu$ . Соответствующее скалярное произведение для  $f_1, f_2 \in \mathcal{L}^2(\mathbb{R}^D \rightarrow \mathbb{R}^D, \mu)$  определяется через  $\langle f_1, f_2 \rangle_\mu \stackrel{\text{def}}{=} \int_{\mathbb{R}^D} \langle f_1(x), f_2(x) \rangle d\mu(x)$ , где  $\langle f_1(x), f_2(x) \rangle$  – евклидово скалярное произведение. Через  $\|\cdot\|_\mu = \sqrt{\langle \cdot, \cdot \rangle_\mu}$  обозначим норму, индуцированную скалярным произведением  $\mathcal{L}^2(\mathbb{R}^D \rightarrow \mathbb{R}^D, \mu)$ .

Напомним следующее полезное свойство полуунпрерывных снизу выпуклых функций  $\psi : \mathbb{R}^D \rightarrow \mathbb{R}$ :

$$\nabla \psi(x) = \arg \max_{y \in \mathbb{R}^D} [\langle y, x \rangle - \bar{\psi}(y)]. \quad (3.17)$$

Это свойство вытекает из того, что

$$\hat{y} = \arg \max_{y \in \mathbb{R}^D} [\langle y, x \rangle - \bar{\psi}(y)] \iff x - \nabla \bar{\psi}(\hat{y}) = 0.$$

*Доказательство теоремы 3.4.1.* Мы рассматриваем разницу между оцененными и истинными корреляциями:

$$\Delta = \sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} \psi_n(x) d\mathbb{P}_n(x) - \sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} \psi_n^*(x) d\mathbb{P}_n(x) =$$

$$\begin{aligned} & \sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} [\langle \nabla \psi_n(x), x \rangle - \overline{\psi_n}(\nabla \psi_n(x))] d\mathbb{P}_n(x) - \\ & \sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} [\langle \nabla \psi_n^*(x), x \rangle - \overline{\psi_n^*}(\nabla \psi_n^*(x))] d\mathbb{P}_n(x). \end{aligned} \quad (3.18)$$

Здесь мы дважды воспользовались (3.17) для  $f = \psi_n$  и  $f = \psi_n^*$ . Отметим, что

$$\begin{aligned} \sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} \langle \nabla \psi_n^*(x), x \rangle d\mathbb{P}_n(x) &= \sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} \langle y, \nabla \overline{\psi_n^*}(y) \rangle d\bar{\mathbb{P}}(y) = \\ \int_{\mathbb{R}^D} \langle y, \sum_{n=1}^N \alpha_n \nabla \overline{\psi_n^*}(y) \rangle d\bar{\mathbb{P}}(y) &= \int_{\mathbb{R}^D} \langle y, y \rangle d\bar{\mathbb{P}}(y) = \|\text{id}_{\mathbb{R}^D}\|_{\bar{\mathbb{P}}}^2, \end{aligned} \quad (3.19)$$

где мы используем формулу замены переменных для  $\nabla \psi_n^* \# \mathbb{P}_n = \bar{\mathbb{P}}$  и (3.5). Аналогично

$$\begin{aligned} \sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} \overline{\psi_n^*}(\nabla \psi_n^*(x)) d\mathbb{P}_n(x) &= \sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} \overline{\psi_n^*}(y) d\bar{\mathbb{P}}(y) = \\ \int_{\mathbb{R}^D} \sum_{n=1}^N \alpha_n \overline{\psi_n^*}(y) d\bar{\mathbb{P}}(y) &= \int_{\mathbb{R}^D} \frac{\|y\|^2}{2} d\bar{\mathbb{P}}(y) = \frac{1}{2} \|\text{id}_{\mathbb{R}^D}\|_{\bar{\mathbb{P}}}^2. \end{aligned} \quad (3.20)$$

Так как каждый потенциал  $\psi_n$  является  $\mathcal{B}$ -гладким, то  $\overline{\psi_n}$  является  $\frac{1}{\mathcal{B}}$ -сильно выпуклым (см. [33]). Таким образом, имеем

$$\begin{aligned} & \overline{\psi_n}(\nabla \psi_n^*(x)) \geqslant \\ & \overline{\psi_n}(\nabla \psi_n(x)) + \underbrace{\langle \nabla \overline{\psi_n}(\nabla \psi_n(x)), \nabla \psi_n^*(x) - \nabla \psi_n(x) \rangle}_{=x} + \\ & \frac{1}{2\mathcal{B}} \|\nabla \psi_n^*(x) - \nabla \psi_n(x)\|^2 = \\ & \overline{\psi_n}(\nabla \psi_n(x)) + \langle x, \nabla \psi_n^*(x) - \nabla \psi_n(x) \rangle + \\ & \frac{1}{2\mathcal{B}} \|\nabla \psi_n^*(x) - \nabla \psi_n(x)\|^2, \end{aligned} \quad (3.21)$$

или эквивалентно

$$\begin{aligned} & -\overline{\psi_n}(\nabla \psi_n(x)) \geqslant \\ & -\overline{\psi_n}(\nabla \psi_n^*(x)) + \langle x, \nabla \psi_n^*(x) - \nabla \psi_n(x) \rangle + \frac{1}{2\mathcal{B}} \|\nabla \psi_n^*(x) - \nabla \psi_n(x)\|^2. \end{aligned} \quad (3.22)$$

Интегрируя (3.22) по  $\mathbb{P}_n$  и суммируя по  $n = 1, 2, \dots, N$  с весами  $\alpha_n$ , получаем

$$-\sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} \overline{\psi_n}(\nabla \psi_n(x)) d\mathbb{P}_n(x) \geqslant$$

$$\begin{aligned}
& - \sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} \overline{\Psi_n}(\nabla \Psi_n^*(x)) d\mathbb{P}_n(x) + \sum_{n=1}^N \alpha_n \langle x, \nabla \Psi_n^*(x) \rangle_{\mathbb{P}_n} - \\
& \sum_{n=1}^N \alpha_n \langle x, \nabla \Psi_n(x) \rangle_{\mathbb{P}_n} + \sum_{n=1}^N \alpha_n \frac{1}{2\mathcal{B}} \|\nabla \Psi_n^*(x) - \nabla \Psi_n(x)\|_{\mathbb{P}_n}^2 = \\
& - \int_{\mathbb{R}^D} \sum_{n=1}^N \alpha_n \overline{\Psi_n}(y) d\bar{\mathbb{P}}(y) + \sum_{n=1}^N \alpha_n \langle x, \nabla \Psi_n^*(x) \rangle_{\mathbb{P}_n} - \\
& \sum_{n=1}^N \alpha_n \langle x, \nabla \Psi_n(x) \rangle_{\mathbb{P}_n} + \sum_{n=1}^N \alpha_n \frac{1}{2\mathcal{B}} \|\nabla \Psi_n^*(x) - \nabla \Psi_n(x)\|_{\mathbb{P}_n}^2. \tag{3.23}
\end{aligned}$$

Отметим, что

$$\begin{aligned}
& - \int_{\mathbb{R}^D} \sum_{n=1}^N \alpha_n \overline{\Psi_n}(y) d\bar{\mathbb{P}}(y) = \\
& \int_{\mathbb{R}^D} \left[ \frac{\|y\|^2}{2} - \sum_{n=1}^N \alpha_n \overline{\Psi_n}(y) \right] d\bar{\mathbb{P}}(y) - \int_{\mathbb{R}^D} \frac{\|y\|^2}{2} d\bar{\mathbb{P}}(y) \\
& \int_{\mathbb{R}^D} \left[ \frac{\|y\|^2}{2} - \sum_{n=1}^N \alpha_n \overline{\Psi_n}(y) \right] d\bar{\mathbb{P}}(y) - \frac{1}{2} \|\text{id}_{\mathbb{R}^D}\|_{\bar{\mathbb{P}}}^2. \tag{3.24}
\end{aligned}$$

Подставляя (3.23), (3.24), (3.19) и (3.20) в (3.18), мы получаем (3.9).  $\square$

*Доказательство теоремы 3.4.2.* Так как функция  $\hat{\varphi}_n$  является  $\beta$  сильно выпуклой, ее сопряженная  $\overline{\hat{\varphi}_n}$  является  $\frac{1}{\beta}$ -гладкой, т.е. ее градиент  $\nabla \overline{\hat{\varphi}_n}$   $\frac{1}{\beta}$ -липшицев [33]. Таким образом, при всех  $x, x' \in \mathbb{R}^D$ :

$$\|\nabla \overline{\hat{\varphi}_n}(x) - \nabla \overline{\hat{\varphi}_n}(x')\|^2 \leq (\frac{1}{\beta})^2 \cdot \|x - x'\|^2.$$

Подставляя  $x' = \nabla \hat{\varphi}_n(\nabla \hat{\psi}_n(y)) = (\nabla \overline{\hat{\varphi}_n})^{-1}(\nabla \hat{\psi}_n(y))$ , получаем

$$\|\nabla \overline{\hat{\varphi}_n}(x) - \nabla \hat{\psi}_n(x)\|^2 \leq (\frac{1}{\beta})^2 \|x - \nabla \hat{\varphi}_n(\nabla \hat{\psi}_n(x))\|^2. \tag{3.25}$$

Так как функция  $\hat{\varphi}_n$  является  $\mathcal{B}$ -гладкой, то при всех  $x \in \mathbb{R}^D$  имеем:

$$\begin{aligned}
& \hat{\varphi}_n(\nabla \hat{\psi}_n(x)) \leq \\
& \hat{\varphi}_n(\nabla \overline{\hat{\varphi}_n}(x)) + \underbrace{\langle \nabla \hat{\varphi}_n(\nabla \overline{\hat{\varphi}_n}(x)), \nabla \hat{\psi}_n(x) - \nabla \overline{\hat{\varphi}_n}(x) \rangle}_{=x} + \frac{\mathcal{B}}{2} \|\nabla \hat{\psi}_n(x) - \nabla \overline{\hat{\varphi}_n}(x)\|^2.
\end{aligned}$$

Это эквивалентно тому, что

$$\langle x, \nabla \hat{\psi}_n(x) \rangle - \hat{\varphi}_n(\nabla \hat{\psi}_n(x)) \geq$$

$$\underbrace{\langle x, \nabla \overline{\hat{\phi}_n}(x) \rangle - \hat{\phi}_n(\nabla \overline{\hat{\phi}_n}(x))}_{\overline{\hat{\phi}_n}(x)} - \frac{\mathcal{B}}{2} \|\nabla \hat{\psi}_n(x) - \nabla \overline{\hat{\phi}_n}(x)\|^2. \quad (3.26)$$

Теперь из (3.26) и (3.25) находим, что

$$\langle x, \nabla \hat{\psi}_n(x) \rangle - \hat{\phi}_n(\nabla \hat{\psi}_n(x)) \geq \overline{\hat{\phi}_n}(x) - \frac{\mathcal{B}}{2\beta^2} \cdot \|x - \nabla \hat{\phi}_n(\nabla \hat{\psi}_n(x))\|^2. \quad (3.27)$$

Для каждого  $n = 1, 2, \dots, N$ , интегрируя (3.27) по  $\mathbb{P}_n$  и суммируя соответствующий циклически согласованный регуляризационный член, находим:

$$\begin{aligned} \int_{\mathbb{R}^D} [\langle x, \nabla \hat{\psi}_n(x) \rangle - \hat{\phi}_n(\nabla \hat{\psi}_n(x))] d\mathbb{P}_n(x) + \lambda \cdot \|\nabla \hat{\phi}_n \circ \nabla \hat{\psi}_n - \text{id}_{\mathbb{R}^D}\|_{\mathbb{P}_n}^2 \geq \\ \int_{\mathbb{R}^D} \overline{\hat{\phi}_n}(x) d\mathbb{P}_n(x) + \left(\lambda - \frac{\mathcal{B}}{2\beta^2}\right) \cdot \underbrace{\|\nabla \hat{\phi}_n \circ \nabla \hat{\psi}_n - \text{id}_{\mathbb{R}^D}\|_{\mathbb{P}_n}^2}_{\mathcal{R}_2^{\mathbb{P}_n}(\hat{\psi}_n, \hat{\phi}_n)}. \end{aligned} \quad (3.28)$$

Суммируя (3.28) при  $n = 1, 2, \dots, N$  по весам  $\alpha_n$ , получаем

$$\begin{aligned} \sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} [\langle x, \nabla \hat{\psi}_n(x) \rangle - \hat{\phi}_n(\nabla \hat{\psi}_n(x))] d\mathbb{P}_n(x) + \lambda \sum_{n=1}^N \alpha_n \mathcal{R}_2^{\mathbb{P}_n}(\hat{\psi}_n, \hat{\phi}_n) \geq \\ \underbrace{\sum_{n=1}^N \alpha_n \int_{\mathbb{R}^D} \overline{\hat{\phi}_n}(x) d\mathbb{P}_n(x)}_{\text{MultiCorr}(\{\alpha_n, \mathbb{P}_n\} | \{\overline{\hat{\phi}_n}\})} + \sum_{n=1}^N \alpha_n \left(\lambda - \frac{\mathcal{B}}{2\beta^2}\right) \cdot \mathcal{R}_2^{\mathbb{P}_n}(\hat{\psi}_n, \hat{\phi}_n). \end{aligned}$$

Прибавляя  $\tau \cdot \mathcal{R}_1^{\widehat{\mathbb{P}}}(\{\hat{\phi}_n\})$  к обеим частям (3.29), имеем:

$$\begin{aligned} \text{MultiCorr}(\{\alpha_n, \mathbb{P}_n\} | \{\hat{\psi}_n\}, \{\hat{\phi}_n\}; \tau, \widehat{\mathbb{P}}, \lambda) \geq \text{MultiCorr}(\{\alpha_n, \mathbb{P}_n\} | \{\overline{\hat{\phi}_n}\}) + \\ \tau \cdot \mathcal{R}_1^{\widehat{\mathbb{P}}}(\{\hat{\phi}_n\}) + \sum_{n=1}^N \alpha_n \left(\lambda - \frac{\mathcal{B}}{2\beta^2}\right) \cdot \mathcal{R}_2^{\mathbb{P}_n}(\hat{\psi}_n, \hat{\phi}_n). \end{aligned} \quad (3.29)$$

Вычитая  $\text{MultiCorr}(\{\alpha_n, \mathbb{P}_n\} | \{\varphi_n^*\})$  из обеих частей и используя теорему 2.4.1, получаем, что

$$\begin{aligned} \Delta \geq \\ - \int_{\mathbb{R}^D} \sum_{n=1}^N \left[ \alpha_n \hat{\phi}_n(y) - \frac{\|y\|^2}{2} \right] d\overline{\mathbb{P}}(y) + \frac{\beta}{2} \sum_{n=1}^N \alpha_n \|\nabla \psi_n^*(x) - \nabla \overline{\hat{\phi}_n}(x)\|_{\mathbb{P}_n}^2 + \end{aligned} \quad (3.30)$$

$$\tau \cdot \mathcal{R}_1^{\widehat{\mathbb{P}}}(\{\hat{\phi}_n\}) + \sum_{n=1}^N \alpha_n \left(\lambda - \frac{\mathcal{B}}{2\beta^2}\right) \cdot \mathcal{R}_2^{\mathbb{P}_n}(\hat{\psi}_n, \hat{\phi}_n) \geq \quad (3.31)$$

$$\sum_{n=1}^N \alpha_n \left( \lambda - \frac{\beta}{2\beta^2} \right) \cdot \mathcal{R}_2^{\mathbb{P}_n}(\hat{\psi}_n, \hat{\phi}_n) + \frac{\beta}{2} \sum_{n=1}^N \alpha_n \|\nabla \psi_n^*(x) - \nabla \bar{\phi}_n(x)\|_{\mathbb{P}_n}^2. \quad (3.32)$$

При переходе от (3.31) к (3.32) мы использовали тот факт, что сумма первого члена в (3.30) с регуляризатором  $\tau \cdot \mathcal{R}_1^{\widehat{\mathbb{P}}}(\{\hat{\phi}_n\})$  неотрицательна. Так как  $\lambda > \frac{\beta}{2\beta^2}$ , то из (3.32) мы непосредственно заключаем, что  $\Delta \geq 0$ , что дает **оценку сверху для множественных корреляций** (3.15). С другой стороны, для каждого  $n = 1, 2, \dots, N$  имеем:

$$\|\nabla \psi_n^*(x) - \nabla \bar{\phi}_n(x)\|_{\mathbb{P}_n}^2 \leq \frac{2\Delta}{\alpha_n \beta} \text{ и } \|\nabla \hat{\phi}_n \circ \nabla \hat{\psi}_n - \text{id}_{\mathbb{R}^D}\|_{\mathbb{P}_n}^2 \leq \frac{2\Delta}{\alpha_n \cdot (\lambda - \frac{\beta}{2\beta^2})}. \quad (3.33)$$

Используя вторую часть (3.33), с учетом (3.25), проинтегрированного по  $\mathbb{P}_n$ , получим

$$\|\nabla \bar{\phi}_n - \nabla \hat{\psi}_n\|_{\mathbb{P}_n}^2 \leq \frac{2\Delta}{\alpha_n \cdot (\lambda \beta^2 - \frac{\beta}{2})}. \quad (3.34)$$

Окончательно, применяя неравенство треугольника для  $\|\cdot\|_{\mathbb{P}_n}$ , имеем:

$$\begin{aligned} \|\nabla \psi_n^* - \nabla \hat{\psi}_n\|_{\mathbb{P}_n} &\leq \|\nabla \bar{\phi}_n - \nabla \hat{\psi}_n\|_{\mathbb{P}_n} + \|\nabla \bar{\phi}_n - \nabla \psi_n^*\|_{\mathbb{P}_n} \leq \\ &\leq \sqrt{\frac{2\Delta}{\alpha_n}} \cdot \left( \sqrt{\frac{1}{\beta}} + \sqrt{\frac{1}{\lambda \beta^2 - \frac{\beta}{2}}} \right), \end{aligned} \quad (3.35)$$

т.е.

$$2\mathbb{W}_2^2(\nabla \hat{\psi}_n \# \mathbb{P}_n, \bar{\mathbb{P}}) \leq \|\nabla \psi_n^* - \nabla \hat{\psi}_n\|_{\mathbb{P}_n}^2 \leq \frac{2\Delta}{\alpha_n} \cdot \left( \sqrt{\frac{1}{\beta}} + \sqrt{\frac{1}{\lambda \beta^2 - \frac{\beta}{2}}} \right)^2 = O(\Delta),$$

где первое неравенство вытекает из леммы 2.5.1.  $\square$

### 3.6 Вычислительные эксперименты

Все эксперименты осуществлены с использованием одной видеокарты GTX 1080Ti. Программный код реализован на фреймворке **PyTorch** и публично доступен по адресу

<https://github.com/iamalexkorotin/Wasserstein2Barycenters>.

Мы сравниваем наш метод [ $\text{CW}_2\text{B}$ ] с основанным на потенциалах методе [ $\text{CRWB}$ ] из работы [8] (с расстоянием Васерштейна-2 и  $\mathcal{L}^2$ -регуляризацией), а также с основанным на мерах генеративным методом [ $\text{SCW}_2\text{B}$ ] из работы [52]. Все рассмотренные методы восстанавливают  $2N$  потенциалов  $\{\hat{\psi}_n, \hat{\phi}_n\} \approx \{\psi_n^*, \bar{\psi}_n^*\}$  и приближают барицентр мерами  $\{\nabla \hat{\psi}_n \sharp \mathbb{P}_n\}$ . Метод [ $\text{SCW}_2\text{B}$ ] дополнительно выдает сгенерированный барицентр  $g \sharp \mathbb{S} \approx \bar{\mathbb{P}}$ , где  $g$  – генеративная сеть и  $\mathbb{S}$  – распределение входного шума.

Для оценки качества вычисленного барицентра разумно рассмотреть **процент необъясненной дисперсии** (unexplained variance percentage), определяемый следующим образом:  $\text{UVP}(\tilde{\mathbb{P}}) = 100 \frac{\mathbb{W}_2^2(\tilde{\mathbb{P}}, \bar{\mathbb{P}})}{\text{Var}(\bar{\mathbb{P}})} \%$ . При  $\text{UVP} \approx 0\%$ ,  $\tilde{\mathbb{P}}$  является хорошей аппроксимацией  $\bar{\mathbb{P}}$ . Для значений  $\geq 100\%$  распределение  $\tilde{\mathbb{P}}$  является неподходящим: при тривиальном базовом выборе (baseline)  $\mathbb{P}^0 = \delta_{\mathbb{E}_{\mathbb{P}}[y]}$  мы получаем  $\text{UVP}(\mathbb{P}^0) = 100\%$ . Вычисление метрики UVP для больших размерностей технически неосуществимо: эмпирические оценки для  $\mathbb{W}_2^2$  являются недостоверными из-за высокой выборочной сложности [61]. Для преодоления этой сложности для барицентров, заданных  $\nabla \hat{\psi}_n \sharp \mathbb{P}_n$ , мы используем метрику  $\mathcal{L}^2$ -UVP, определяемую следующим образом:

$$\mathcal{L}^2\text{-UVP}(\nabla \hat{\psi}_n, \mathbb{P}_n) \stackrel{\text{def}}{=} 100 \frac{\|\nabla \hat{\psi}_n - \nabla \psi_n^*\|_{\mathbb{P}_n}^2}{\text{Var}(\bar{\mathbb{P}})} \% \quad \left[ \geq \text{UVP}(\nabla \hat{\psi}_n \sharp \mathbb{P}_n) \right], \quad (3.36)$$

где неравенство в скобках следует из леммы 2.5.1. В метрике  $\mathcal{L}^2$ -UVP происходит сравнение не только отображеного распределения  $\nabla \hat{\psi}_n \sharp \mathbb{P}_n$  с барицентром  $\bar{\mathbb{P}}$ , но также и полученного транспортного отображения с отображением оптимального транспорта  $\nabla \psi_n^*$ . В ней дается оценка для  $\text{UVP}(\nabla \hat{\psi}_n \sharp \mathbb{P}_n)$  сверху. Кроме того для метрики  $\mathcal{L}^2$ -UVP можно естественно получать *несмешанные оценки* по методу Монте-Карло, используя случайные выборки из  $\mathbb{P}_n$ .

Для барицентров, заданных в неявной форме  $g \sharp \mathbb{S}$ , мы вычисляем метрику **Буреса–Васерштейна** UVP, определяемую следующим образом:

$$\text{BW}_2^2\text{-UVP}(g \sharp \mathbb{S}) \stackrel{\text{def}}{=} 100 \frac{\text{BW}_2^2(g \sharp \mathbb{S}, \bar{\mathbb{P}})}{\frac{1}{2}\text{Var}(\bar{\mathbb{P}})} \% \quad \left[ \leq \text{UVP}(g \sharp \mathbb{S}) \right], \quad (3.37)$$

где  $\text{BW}_2^2(\mathbb{P}, \mathbb{Q}) = \mathbb{W}_2^2(\mathcal{N}(\mu_{\mathbb{P}}, \Sigma_{\mathbb{P}}), \mathcal{N}(\mu_{\mathbb{Q}}, \Sigma_{\mathbb{Q}}))$  – метрика Буреса–Васерштейна, здесь  $\mu_{\mathbb{P}}, \Sigma_{\mathbb{P}}$  обозначают среднее и ковариацию распределения  $\mathbb{P}$  [60]. Для мер  $\mathbb{P}, \mathbb{Q}$  с невырожденными матрицами ковариации она задается формулой

$$\text{BW}_2^2(\mathbb{P}, \mathbb{Q}) = \frac{1}{2} \|\mu_{\mathbb{P}} - \mu_{\mathbb{Q}}\|^2 + \left[ \frac{1}{2} \text{Tr} \Sigma_{\mathbb{P}} + \frac{1}{2} \text{Tr} \Sigma_{\mathbb{Q}} - \text{Tr}(\Sigma_{\mathbb{P}}^{\frac{1}{2}} \Sigma_{\mathbb{Q}} \Sigma_{\mathbb{P}}^{\frac{1}{2}})^{\frac{1}{2}} \right].$$

В метрике Буреса–Васерштейна сравнение  $\mathbb{P}$  и  $\mathbb{Q}$  осуществляется только по их первым и вторым моментам. Известно, что  $B\mathbb{W}_2^2(\mathbb{P}, \mathbb{Q})$  представляет собой оценку снизу для  $\mathbb{W}_2^2(\mathbb{P}, \mathbb{Q})$  (см. [62]). Таким образом, имеем  $B\mathbb{W}_2^2\text{-UVP}(g\sharp\mathbb{S}) \leq \text{UVP}(g\sharp\mathbb{S})$ . На практике при вычислении  $B\mathbb{W}_2^2\text{-UVP}(g\sharp\mathbb{S})$  используются оценки средних и матрицы ковариаций распределений с использованием  $10^5$  случайных выборок.

### 3.6.1 Линейно-разбросанные семейства в большой размерности

В этом параграфе в качестве весов мы рассматриваем  $N = 4$  с  $(\alpha_1, \dots, \alpha_4) = (0.1, 0.2, 0.3, 0.4)$ . Изучим **линейно-разбросанное семейство** распределений [36, §4], для которого можно вычислить истинный барицентр. Пусть  $\mathbb{P}_0 \in \mathcal{P}_{2,\text{ac}}$ . Рассмотрим следующее линейно-разбросанное семейство распределений  $\mathcal{F}(\mathbb{P}_0) = \{f_{S,u}\sharp\mathbb{P}_0 \mid S \in \mathcal{M}_{D \times D}^+, u \in \mathbb{R}^D\}$ , где  $f_{S,u} : \mathbb{R}^D \rightarrow \mathbb{R}^D$  – линейное отображение  $f_{S,u}(x) = Sx + u$  с положительно определенной матрицей  $S \in \mathcal{M}_{D \times D}^+$ . При  $\{\mathbb{P}_n\} \subset \mathcal{F}(\mathbb{P}_0)$  барицентр  $\bar{\mathbb{P}}$  также лежит в  $\mathcal{F}(\mathbb{P}_0)$  и может быть вычислен через итерационный процесс приближений к неподвижной точке [36].

На рис. 3.1а мы показываем 2-мерное линейно-разбросанное семейство, сгенерированное с использованием распределения типа “швейцарский рулет”  $\mathbb{P}_0$ . Истинный барицентр показан на рис. 3.1б. Сгенерированный барицентр  $g\sharp\mathbb{S}$  из [SCW<sub>2</sub>B] представлен на рис. 3.1в. Отображенные меры  $\nabla\hat{\psi}_n\sharp\mathbb{P}_n$  для каждого из методов показаны на рис. 3.1г, 3.1д и 3.1е соответственно. В этом примере все отображенные меры  $\nabla\hat{\psi}_n\sharp\mathbb{P}_n$  дают хорошую аппроксимацию для  $\bar{\mathbb{P}}$ , в то время как для сгенерированного барицентра  $g\sharp\mathbb{S}$  для [SCW<sub>2</sub>B] (см. рис. 3.1в) визуально видна недостаточность обучения.

Для количественного сравнения мы рассматриваем два выбора  $\mathbb{P}_0$ :  $D$ -мерное стандартное гауссовское распределение и равномерное распределение на  $[-\sqrt{3}, +\sqrt{3}]^D$ . Каждое  $\mathbb{P}_n$  строится как  $f_{S_n^T \Lambda S_n, 0}\sharp\mathbb{P}_0 \in \mathcal{F}(\mathbb{P}_0)$ , где  $S_n$  – матрица случайных вращений и  $\Lambda$  – диагональная матрица с элементами  $[\frac{1}{2}b^0, \frac{1}{2}b^1, \dots, 2]$ , где  $b = \sqrt[D-1]{4}$ . Мы рассматриваем только центрированные распределения (т.е. с нулевым средним), поскольку барицентр нецентрированного  $\{\mathbb{P}_n\} \in \mathcal{P}_{2,\text{ac}}(\mathbb{R}^D)$  есть барицентр  $\{\mathbb{P}'_n\}$ , сдвинутый на  $\sum_{n=1}^N \alpha_n \mu_{\mathbb{P}_n}$ , где  $\{\mathbb{P}'_n\}$  – центрированные копии  $\{\mathbb{P}_n\}$  [36]. Полученные результаты указаны в табл. 4 и 5.

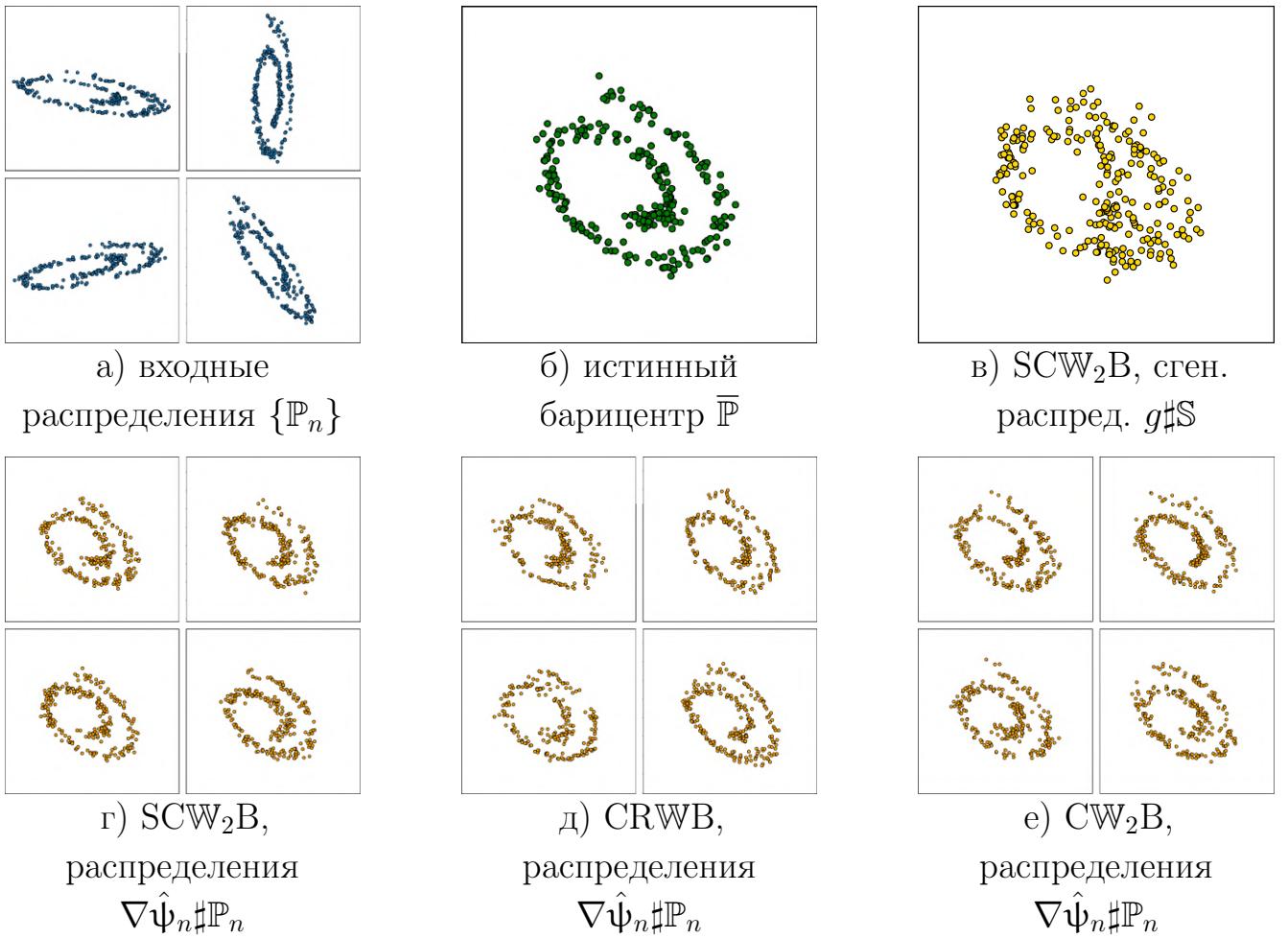


Рисунок 3.1 — Барицентр линейно-разбросанной популяции “швейцарский рулет”, вычисленный по трем методам.

В этих экспериментах наш метод работает лучше, чем [CRWB] и [SCW<sub>2</sub>B]. Для [CRWB] размерность  $\sim 16$  является точкой перелома: метод плохо масштабируется на большие размерности. Метод [SCW<sub>2</sub>B] лучше масштабируется с увеличением размерности, но его ошибки  $\mathcal{L}^2$ -UVP и  $\mathbb{W}_2^2$ -UVP в два раза превосходят ошибки нашего метода. По-видимому, это можно объяснить генеративной аппроксимацией и сложной min-max-min оптимизацией в [SCW<sub>2</sub>B]. Для полноты мы также сравниваем наш алгоритм с предложенным в [55] алгоритмом, в котором барицентр аппроксимируется дискретным распределением на фиксированном числе точек. В наших экспериментах, аналогично [8], в качестве носителя мы берем 5000. Как и ожидалось, ошибка  $\mathbb{W}_2^2$ -UVP метода сильно возрастает при увеличении размерности, так что этот метод проигрывает нашему подходу.

Для демонстрации способности нашего метода к масштабированию (при числе  $N$  входных распределений) мы проводим аналогичный эксперимент

Метрика	Метод	D=2	4	8	16	32	64	128	256
BW <sub>2</sub> <sup>2</sup> -UVP, %	[FCWB], [55]	0.7	0.68	1.41	3.87	8.85	14.08	18.11	21.33
	[SCW <sub>2</sub> B], [52]	0.07	0.09	0.16	0.28	0.43	0.59	1.28	2.85
$\mathcal{L}_2$ -UVP, % (потенциалы)	[CRWB], [8]	0.08	0.10	0.17	0.29	0.47	0.63	1.14	1.50
	[CW <sub>2</sub> B], <u>наш</u>	0.99	2.52	8.62	22.23	67.01	>100		
		0.06	0.05	0.07	0.11	0.19	0.24	0.42	0.83

Таблица 4 — Сравнение UVP в случае  $\{\mathbb{P}_n\} \subset \mathcal{F}(\mathbb{P}_0)$ ,  $\mathbb{P}_0 = \mathcal{N}(0, I_D)$ ,  $N = 4$ .

Метрика	Метод	D=2	4	8	16	32	64	128	256
BW <sub>2</sub> <sup>2</sup> -UVP, %	[FCWB], [55]	0.64	0.77	1.22	3.75	8.92	14.3	18.46	21.64
	[SCW <sub>2</sub> B], [52]	0.12	0.10	0.19	0.29	0.46	0.6	1.38	2.9
$\mathcal{L}_2$ -UVP, % (потенциалы)	[CRWB], [8]	0.17	0.12	0.2	0.31	0.47	0.62	1.21	1.52
	[CW <sub>2</sub> B], <u>наш</u>	0.58	1.83	8.09	21.23	55.17	> 100		
		0.17	0.08	0.06	0.1	0.2	0.25	0.42	0.82

Таблица 5 — Сравнение UVP в случае  $\{\mathbb{P}_n\} \subset \mathcal{F}(\mathbb{P}_0)$ ,  $\mathbb{P}_0 = \text{Uniform}([-\sqrt{3}, +\sqrt{3}]^D)$ ,  $N = 4$ .

Метрика	Метод	D=32	64	128
BW <sub>2</sub> <sup>2</sup> -UVP, %	[FCWB], [55]	14.09	26.21	38.43
	[SCW <sub>2</sub> B], [52]	0.62	0.93	1.83
$\mathcal{L}_2$ -UVP, % (потенциалы)	[CRWB], [8]	0.60	0.86	1.52
	[CW <sub>2</sub> B], <u>наш</u>	0.31	0.58	1.45

Таблица 6 — Сравнение UVP в случае  $\{\mathbb{P}_n\} \subset \mathcal{F}(\mathbb{P}_0)$ ,  $\mathbb{P}_0 = \text{Uniform}([-\sqrt{3}, +\sqrt{3}]^D)$ ,  $N = 20$ .

с линейно-разбросанным семейством большой размерности при  $N = 20$ . Положим  $\alpha_n = \frac{2n}{N(N+1)}$  для  $n = 1, 2, \dots, 20$  и выберем равномерное распределение на  $[-\sqrt{3}, +\sqrt{3}]^D$  как  $\mathbb{P}_0$  и, как и выше, построим распределения  $\mathbb{P}_n \in \mathcal{F}(\mathbb{P}_0)$ . Результаты для размерностей 32, 64 и 128 представлены в табл. 6. Как и в таблицах 4 и 5, мы получаем, что наш метод работает лучше, чем альтернативны.

### 3.6.2 Агрегация апостериорных распределений на подмножествах

Мы применим наш метод для агрегации апостериорных распределений по подмножествам данных. Барицентр апостериорных распределений подмножеств сходится к истинному апостериорному распределению [51]. Таким образом, вычисление барицентра апостериорных распределений подмножеств является эффективной альтернативой для получения полного апостериорного распределения для задач больших данных [8; 50; 57].

Аналогично работе [8], мы рассматриваем пуассоновские и отрицательные биномиальные регрессии для задачи предсказания почасового числа аренд велосипедов, используя такие признаки, как день недели и погодные условия<sup>2</sup>. Мы рассматриваем апостериорное распределение 8-мерных коэффициентов регрессии для пуассоновских и отрицательных биномиальных регрессий. Данные случайно разбиваются на  $N = 5$  подмножеств одинакового размера, берутся  $10^5$  выборок из каждого апостериорных распределений подмножеств с использованием библиотеки “Stan” [63]. В результате получаем дискретное равномерное распределения  $\{\mathbb{P}_n\}$  с носителем на выборках. В качестве эталонного барицентра  $\bar{\mathbb{P}}$  мы рассматриваем апостериорное распределение по всей выборке, также состоящее из  $10^5$  точек.

Мы используем  $\text{BW}_2^2\text{-UVP}(\tilde{\mathbb{P}}, \bar{\mathbb{P}})$  для сравнения полученного барицентра  $\bar{\mathbb{P}}$  (образ меры  $\nabla\hat{\Psi}_n\#\mathbb{P}_n$  или сгенерированная мера  $g\#\mathbb{S}$ ) с истинным барицентром. Результаты приведены в табл. 7. Все рассмотренные методы работают хорошо ( $\text{UVP} < 2\%$ ), однако наш метод дает лучшие результаты.

	Регрессия	SCW <sub>2</sub> B, [52]	[CRWB], [8]	CW <sub>2</sub> B, <u>наш</u>
		$\tilde{\mathbb{P}} = g\#\mathbb{S}$	$\tilde{\mathbb{P}} = \nabla\hat{\Psi}_n\#\mathbb{P}_n$	
$\text{BW}_2^2\text{-UVP, \%}$	пуассоновск.	0.67	0.41	1.53
	отр. бином.	0.15	0.15	0.11

Таблица 7 — Сравнение UVP для восстановленных барицентров в нашей задаче агрегации апостериорных распределений подмножеств.

### 3.6.3 Усреднение цветовой палитры

Для понимания качественной картины мы применяем наш метод для усреднения цветовой палитры изображений. Для RGB-изображения  $\mathcal{I}$  его цветовая палитра определяется дискретным равномерным распределением  $\mathbb{P}(\mathcal{I})$  по всем его пикселям  $\in [0,1]^3$ . Для 3 изображений  $\{\mathcal{I}_n\}$  мы вычисляем барицентр  $\bar{\mathbb{P}}$  для каждой из цветовых палитр  $\mathbb{P}_n = \mathbb{P}(\mathcal{I}_n)$  относительно равномерных весов  $\alpha_n = \frac{1}{3}$ . Каждый вычисленный потенциал  $\nabla\hat{\Psi}_n$  применяется попиксельно к  $\mathcal{I}_n$  для получения “отображенного” изображения  $\nabla\hat{\Psi}_n\#\mathcal{I}_n$ . Эти “отображенные” изображения должны быть близки к барицентру  $\bar{\mathbb{P}}$  распределения  $\{\mathbb{P}_n\}$ .

<sup>2</sup><http://archive.ics.uci.edu/ml/datasets/Bike+Sharing+Dataset>

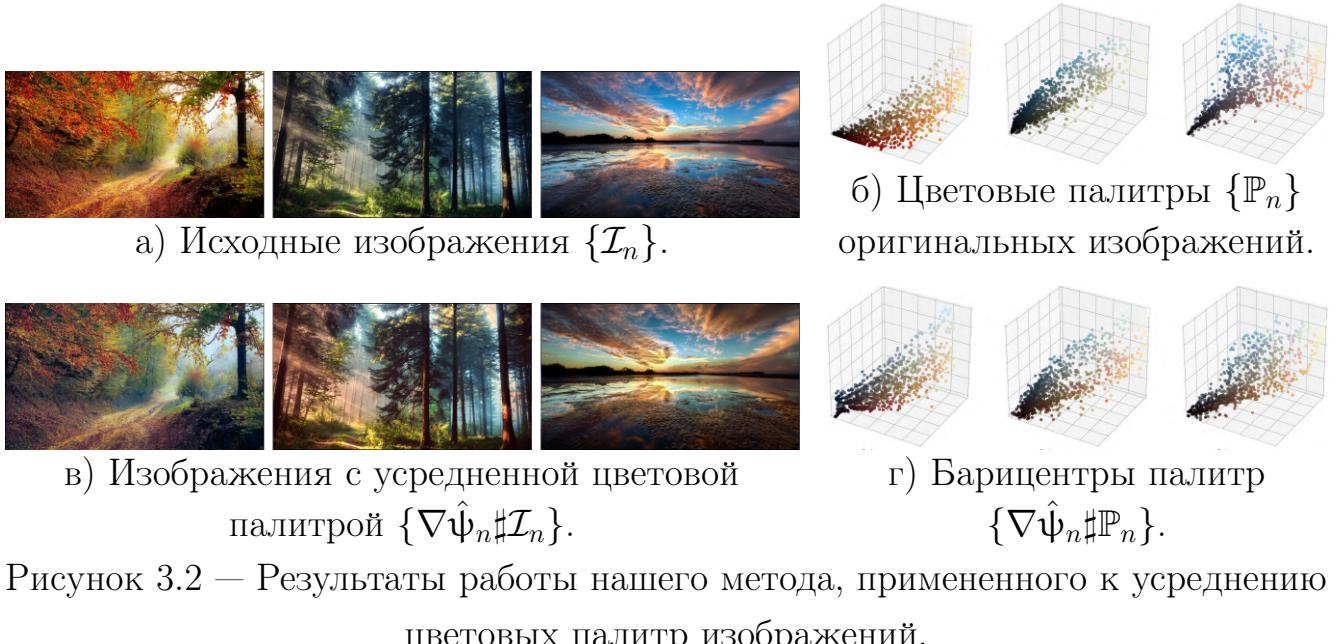


Рисунок 3.2 — Результаты работы нашего метода, примененного к усреднению цветовых палитр изображений.

Полученные результаты приведены на рис. 3.2. Отметим, что изображение  $N\hat{\psi}_1 \# I_1$  наследует некоторые атрибуты изображений  $I_2$  и  $I_3$ : небо становится более синим и деревья становятся более зелеными. С другой стороны, солнечный свет в изображениях  $N\hat{\psi}_2 \# I_2, N\hat{\psi}_3 \# I_3$  становится более оранжевым благодаря преобладанию этого цвета в  $I_1$ .

### 3.6.4 Технические детали экспериментов

**Метод CW<sub>2</sub>B (наш метод).** Мы используем  $\tau = 5$  и  $\hat{P} = \mathcal{N}(0, I_D)$  в нашем условии регуляризатора на конгруэнтность  $\tau \cdot \mathcal{R}_1^{\hat{P}}$ . Мы также используем  $\lambda = 10$  для циклической регуляризации  $\lambda \cdot \mathcal{R}_2^{\hat{P}_n}$  при всех  $n = 1, 2, \dots, N$ . Для аппроксимации потенциалов  $\{\hat{\psi}_n, \hat{\phi}_n\}$  в размерности  $D$  мы используем сеть

$$\text{DenseICNN}[2; \max(64, 2D), \max(64, 2D), \max(32, D)]$$

с функцией активации CELU. Сеть DenseICNN обладает выпуклой по входу полной архитектурой с дополнительными выпуклыми квадратичными обходными связями. Здесь 2 – ранг гессиана обходных связей. Каждое последующее число  $\max(\cdot, \cdot)$  представляет собой размер скрытого полно связного слоя в последующей части сети. По поводу подробного обсуждения архитектуры см. п. 2.7. Обучение выполнялось в соответствии с алгоритмом 2. Пусть  $K = 1024$  – размер случайных выборок и  $\gamma = 0.2$  – балансировочный коэффициент. Мы

используем оптимизатор Adam из [34] с фиксированной скоростью обучения  $10^{-3}$ . Общее число итераций 50000.

**Алгоритм SCW<sub>2</sub>B.** Для распределения входного шума в генеративной модели мы используем  $\mathbb{S} = \mathcal{N}(0, I_D)$ . Для генеративной сети  $g : \mathbb{R}^D \rightarrow \mathbb{R}^D$  мы используем полносвязную последовательную сеть ReLU с числом скрытых слоев

$$[\max(100, 2D), \max(100, 2D), \max(100, 2D)].$$

До начала основной оптимизации мы предварительно обучаем сеть для выполнения приближенного равенства  $g(z) \approx z$  при всех  $z \in \mathbb{R}^D$ . Это условие было эмпирически проверено и превосходит случайную инициализацию весов сети. Для потенциалов были использованы в точности те же сети, что и в нашем методе. Процесс обучения был выполнен согласно min-max-min-процедуре, описанной в алгоритме 1 работы [52]. Размер случайной выборки был выбран равным 1024. Мы использовали оптимизатор Adam из [34] с фиксированной скоростью обучения  $10^{-3}$  для потенциалов и  $10^{-4}$  для генеративной сети  $g$ . Число итераций во внешнем (min-max-min) цикле полагается равным 15000. Следуя [52], мы использовали 10 итераций для внутреннего (min-max-min) цикла и 6 итераций для внутреннего (min-max-min) цикла.

**Алгоритм CRWB.** Метод [CRWB] использует регуляризацию для сохранения сопряженности потенциалов. Используется энтропия или  $\mathcal{L}^2$ -регуляризация по некоторой предложенной мере  $\hat{\mathbb{P}}$  (по поводу дальнейших деталей см. п. 3 в [8]). Используя предоставленный авторами код, мы применили  $\mathcal{L}_2$ -регуляризацию (эксперименты показывают, что ее применение дает лучшую устойчивость, чем энтропийная регуляризация). В качестве меры регуляризации  $\hat{\mathbb{P}}$  бралась равномерная мера на пространстве, содержащем носитель всех исходных распределений; оценка производилась по выборке. Параметр регуляризации  $\Delta$  был выбран равным  $10^{-4}$ . Для аппроксимации потенциалов  $\{\hat{\psi}_n, \hat{\phi}_n\}$  в размерности  $D$  мы используем полносвязные нейронные сети ReLU с размерами слоев

$$[\max(128, 4D), \max(128, 4D), \max(128, 4D)].$$

Мы также попытались применить архитектуру DenseICNN, но не получили никакого улучшения работы. Обучение выполнялось по алгоритму 1 из [8] при размере случайных выборок 1024. Был использован оптимизатор Adam из [34] с фиксированной скоростью обучения  $10^{-3}$ . Общее число итераций выбиралось равным 50000.

## Глава 4. Метод для моделирования Васерштейн-2 градиентных потоков функционалов

Градиентные потоки Васерштейна представляют собой мощный инструмент для понимания и решения многих диффузионных уравнений. В частности, решение уравнения Фоккера–Планка, которое моделирует диффузию вероятностных мер, можно рассматривать как градиентный поток относительно функционалов энтропии в пространстве Васерштейна. Эта эквивалентность, которая была введена Джорданом, Киндерлерером и Отто, послужила отправной точкой для создания так называемой JKO-схемы для аппроксимации этих диффузионных процессов посредством неявной дискретизации градиентного потока в пространстве Васерштейна. Однако решение оптимизационной задачи, связанной с каждым шагом схемы JKO, представляет собой серьезную вычислительную проблему. Мы вводим масштабируемый метод, аппроксимирующий градиентные потоки Васерштейна и нацеленный на приложения в задачах машинного обучения. Наш подход опирается на использование выпуклых по входу нейронных сетей (ICNN) для дискретизации JKO-шагов, после чего оптимизация достигается при использовании стохастического метода градиентного спуска. В отличие от предыдущих работ наш метод не требует дискретизации домена или моделирования по методу частиц. В результате на каждом временном этапе диффузии может быть осуществлена выборка из меры и вычислена ее плотность вероятности. Для оценки производительности нашего алгоритма мы рассматриваем задачу вычисления диффузии для уравнения Фоккера–Планка и применяем его к задаче моделирования апостериорного распределения.

### 4.1 Введение

Стохастические дифференциальные уравнения (SDE), использующиеся для моделирования эволюции во времени случайных диффузионных процессов, имеют приложения в физике [64], финансовой математике [65; 66] и динамике популяций [67]. В машинном обучении диффузионные процессы также возни-

кают в задачах фильтрации [68; 69] и задачах ненормированной выборки из апостериорных распределений путем дискретизации диффузии Ланжевена [70].

Поведение нестационарной по времени плотности  $\rho_t$  таких диффузионных процессов описывается уравнением Фоккера–Планка. В работе [5] Джордан, Киндерлерер и Отто показали, что уравнение Фоккера–Планка эквивалентно градиентному потоку функционала энтропии в пространстве Васерштейна, т.е. в пространстве вероятностных мер с конечными вторыми моментами, снабженном расстоянием Васерштейна. Это наблюдение привело к созданию простой минимизационной схемы (называемой JКО-схемой), в которой используется неявная эйлерова дискретизация градиентного потока Васерштейна. Однако каждый шаг JКО-схемы является весьма затратным с вычислительной точки зрения, поскольку в нем требуется решение минимизационной задачи, связанной с расстоянием Васерштейна.

Одним из способов вычисления диффузии является использование фиксированной дискретизации домена и применение стандартных методов численного интегрирования [71—75] для нахождения  $\rho_t$ . Например, в [76] предлагается метод аппроксимации диффузии, основанный на шагах схемы JКО и регуляризации по энтропии в задаче оптимального транспорта. Однако применение этих методов ограничивается лишь небольшими размерностями, поскольку требуемая дискретизация пространства растет экспоненциально.

Альтернативой дискретизации домена является моделирование методом частиц. В этом методе случайные выборки (частицы) получаются из начального распределения, а их эволюция описывается такими стандартными методами, как схема Эйлера–Маруямы (см. [77, §9.2]). После сходимости оказывается, что распределение частиц приблизительно соответствует стационарному распределению, но при этом метод не дает оценку плотности.

Другим способом отказа от дискретизации является параметризация плотности  $\rho_t$ . Большинство существующих методов приближают только первые и вторые моменты плотности  $\rho_t$  (т.е. используют гауссовскую аппроксимацию). После этого требуемую динамику можно получить, используя подходы, опирающиеся на технику фильтров Кальмана (см. [68; 78—80]). Также исследовались более сложные аппроксимации смеси гауссовых распределений [81; 82] и более общие параметрические семейства [83]. В работе [84] вариационные методы используются для минимизации расхождения между прогнозируемой и истинной плотностями.

В работе [11] был введен параметрический метод для вычисления JKO-шагов при помощи регуляризованного энтропией оптимального транспорта. В этой работе регуляризация расстояния Васерштейна осуществлялась на каждом JKO-шаге и обеспечивала строгую выпуклость задачи, а плотность представлялась как линейная комбинация базисных функций и оценивалась с помощью стохастических методов. Этот метод дает *ненормированную* плотность вероятности без прямого доступа к возможности получения выборки.

Недавно был предложен ряд масштабируемых методов задачи оптимального транспорта между непрерывными вероятностными распределениями, в которых решения параметризуются при помощи воспроизводящих ядер (reproducing kernels) [85], полносвязных нейронных сетей (fully-connected neural networks) [2], или выпуклых по входу нейронных сетей (Input Convex Neural Networks, ICNN) [7]. В частности, использование сетей ICNN привлекло внимание к задаче транспорта относительно расстояния Васерштейна-2, поскольку в этой постановке градиенты  $\nabla \psi_\theta : \mathbb{R}^D \rightarrow \mathbb{R}^D$  могут представлять отображения задачи оптимального транспорта для квадратичной функции стоимости. Такие методы обладают лучшей способностью к масштабированию на случай высоких размерностей без дискретизации входных мер, но в общем их использование очень вычислительно затратно, что не позволяет обеспечить их прямое применение на JKO-шагах.

**Основные результаты, полученные в главе.** Мы предлагаем масштабируемый параметрический метод для аппроксимации градиентных потоков Васерштейна с использованием JKO-шагов и выпуклых по входу нейронных сетей (ICNN) [20]. В частности, мы используем теорему Бренье, чтобы избежать дорогостоящих вычислений расстояния Васерштейна. При этом оптимальное транспортное отображение параметризуется как градиент сети ICNN. При наличии выборочного доступа к начальной мере  $\rho_0$  мы используем метод стохастического градиентного спуска (SGD) для последовательного обучения дискретизованной по времени динамики JKO для  $\rho_t$ . Обученная модель обладает способностью получать выборки из непрерывной аппроксимации  $\rho_t$  и вычислять ее плотность  $\frac{d\rho_t}{dx}(x)$ . Мы вычисляем градиентные потоки для функционала свободной энергии Фоккера–Планка  $\mathcal{F}_{\text{FP}}$ , заданного (4.5), однако наш метод обобщается и на другие случаи. Производительность нашего подхода оценивается путем вычисления диффузии для уравнения Фоккера–Планка и ее применения к задаче получения выборок из ненормированной плотности.

**Обозначения.**  $\mathcal{P}_2(\mathbb{R}^D)$  – множество борелевских вероятностных мер на  $\mathbb{R}^D$  с конечными вторыми моментами.  $\mathcal{P}_{2,ac}(\mathbb{R}^D)$  – подмножество  $\mathcal{P}_2(\mathbb{R}^D)$ , состоящее из вероятностных мер, абсолютно непрерывных относительно лебеговой меры. Для  $\rho \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$  через  $\frac{d\rho}{dx}(x)$  мы обозначаем ее плотность относительно меры Лебега. Через  $\Pi(\mu, \nu)$  мы обозначаем множество вероятностных мер на  $\mathbb{R}^D \times \mathbb{R}^D$  с маргиналами  $\mu$  и  $\nu$ . Для измеримого отображения  $T : \mathbb{R}^D \rightarrow \mathbb{R}^D$  через  $T^\sharp$  мы обозначаем соответствующий оператор отображения между мерами.

## 4.2 Основные сведения о градиентных потоках Васерштейна

Мы рассматриваем градиентные потоки в пространстве Васерштейна ( $\mathcal{P}_2(\mathbb{R}^D), \mathcal{W}_2$ ), т.е. в пространстве вероятностных мер с конечными вторыми моментами  $\mathbb{R}^D$ , наделенном метрикой Васерштейна-2  $\mathcal{W}_2$ .

**Расстояние Васерштейна-2.** Для мер  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^D)$  (квадрат) метрики Васерштейна-2  $\mathcal{W}_2$  между ними определяется следующим образом:

$$\mathcal{W}_2^2(\mu, \nu) \stackrel{\text{def}}{=} \min_{\pi \in \Pi(\mu, \nu)} \int_{\mathbb{R}^D \times \mathbb{R}^D} \|x - y\|_2^2 d\pi(x, y), \quad (4.1)$$

где минимум берется по мерам  $\pi$  на  $\mathbb{R}^D \times \mathbb{R}^D$  с маргиналами  $\mu$  и  $\nu$  соответственно [25]. Отметим, что в данной главе квадрат расстояния Васерштейна-2 мы определяем без множителя  $\frac{1}{2}$ , то есть  $\mathcal{W}_2^2 = 2\mathbb{W}_2^2$ , где  $\mathbb{W}_2^2$  определено в (1.1).

Для  $\mu \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$  существует  $\mu$ -единственное отображение  $\nabla\psi^* : \mathbb{R}^D \rightarrow \mathbb{R}^D$ , являющееся градиентом выпуклой функции  $\psi^* : \mathbb{R}^D \rightarrow \mathbb{R} \sqcup \{\infty\}$ , удовлетворяющей  $\nabla\psi^*\sharp\mu = \nu$  (см. [17]). Из теоремы Бренье [26] следует, что  $\pi^* = [\text{id}_{\mathbb{R}^D}, \nabla\psi^*]\sharp\mu$  – единственный минимизатор для (4.1), т.е.

$$\mathcal{W}_2^2(\mu, \nu) = \int_{\mathbb{R}^D} \|x - \nabla\psi^*(x)\|_2^2 d\mu(x).$$

**Градиентные потоки Васерштейна.** В евклидовом случае градиентные потоки для функции  $f : \mathbb{R} \rightarrow \mathbb{R}$  направлены в направлении ее наибольшего убывания и определяются дифференциальным уравнением  $\frac{dx_t}{dt} = -\nabla f(x_t)$ . Дискретизация этого потока приводит к алгоритму минимизации по методу градиентного спуска. Если функционалы действуют на пространстве мер, снабженном метрикой Васерштейна-2, то аналогичный поток называется градиентным потоком Васерштейна. Здесь идея аналогична: поток направлен

в направлении наибольшего убывания, но теперь градиент определяется более сложно. По поводу некоторых сведений о градиентных потоках в метрических пространствах мы отсылаем читателя к [86]; доступное введение в эту теорию дается в главе 8 работы [9].

Функция  $f_\rho : \mathbb{R}^D \rightarrow \mathbb{R}$  называется *первой вариацией* [87] функционала  $\mathcal{F} : \mathcal{P}_2(\mathbb{R}^D) \rightarrow \mathbb{R} \cup \{\infty\}$  в точке  $\rho$ , если выполняется  $\frac{d}{d\varepsilon}\mathcal{F}(\rho + \varepsilon\chi)|_{\varepsilon=0} = \int_{\mathbb{R}^D} f_\rho(x)d\chi(x)$  для любой меры  $\chi$  со знаком на  $\mathbb{R}^D$  такой, что  $\int 1d\chi(x) = 0$  и существует  $\varepsilon_0 > 0$ , при котором для любого  $\varepsilon \in [0, \varepsilon_0]$  выполнено  $\rho + \varepsilon\chi \in \mathcal{P}_2(\mathbb{R}^D)$ , то есть мера  $\rho + \varepsilon\chi$  является вероятностной. Первая вариация не всегда существует. При этом, если первая вариация в точке  $\rho$  существует, то она определена с точностью до константы и обозначается через  $\mathcal{F}'(\rho) = f_\rho$ .

**Васерштейн градиентным потоком** функционала  $\mathcal{F}$  называется кривая, состоящая из мер  $\{\rho_t\}_{t \in \mathbb{R}_+}$ , являющаяся решением уравнения [87]:

$$\frac{\partial \rho_t}{\partial t} = \operatorname{div}(\rho_t \nabla_x \mathcal{F}'(\rho_t)), \quad \text{s.t. } \rho_0 = \rho^0. \quad (4.2)$$

Член в правой части можно понимать, как градиент  $\mathcal{F}$  в пространстве Васерштейна, т.е., векторное поле,двигающее массу  $\rho_t$  с целью получения наиболее крутого возможного локального изменения  $\mathcal{F}$ .

Градиентные потоки Васерштейна используются в различных прикладных задачах, например, в задачах обучения [88—90] или улучшения [91] неявных генеративных моделей. В обучении с подкреплением градиентные потоки упрощают оптимизационную политику [92; 93]. Другие применения включают в себя моделирование движения толпы [76; 94; 95], оптимизацию наборов данных [96] и промежуточную анимацию (in-between animation) [97].

Во многих приложениях используется связь между градиентными потоками Васерштейна и стохастическими дифференциальными уравнениями. Рассмотрим  $\mathbb{R}^D$ -значный стохастический процесс  $\{X_t\}_{t \in \mathbb{R}_+}$ , задаваемый следующим стохастическим дифференциальным уравнением Ито:

$$dX_t = -\nabla \Phi(X_t)dt + \sqrt{2\beta^{-1}}dW_t, \quad \text{s.t. } X_0 \sim \rho^0 \quad (4.3)$$

где  $\Phi : \mathbb{R}^D \rightarrow \mathbb{R}$  — потенциальная функция,  $W_t$  — стандартный винеровский процесс,  $\beta > 0$  — его величина. Решение задачи (4.3) называется диффузионным процессом *адвекции-диффузии*. В каждый момент времени маргинальная мера  $\rho_t$  описывающая распределение  $X_t$  удовлетворяет *уравнению Фоккера–Планка*

с фиксированным коэффициентом диффузии:

$$\frac{\partial \rho_t}{\partial t} = \operatorname{div}(\nabla \Phi(x) \rho_t) + \beta^{-1} \Delta \rho_t, \quad \text{s.t. } \rho_0 = \rho^0. \quad (4.4)$$

Уравнение (4.4) есть градиентный поток Васерштейна (4.2) для  $\mathcal{F}$ , заданный функционалом свободной энергии Фоккера–Планка [5]

$$\mathcal{F}_{\text{FP}}(\rho) = \mathcal{U}(\rho) - \beta^{-1} \mathcal{E}(\rho), \quad (4.5)$$

где  $\mathcal{U}(\rho) = \int_{\mathbb{R}^D} \Phi(x) d\rho(x)$  называется *потенциальной энергией*, а  $\mathcal{E}(\rho) = - \int_{\mathbb{R}^D} \log \frac{d\rho}{dx}(x) d\rho(x)$  – *энтропия*. В результате для решения стохастического дифференциального уравнения (4.3) можно вычислить градиентный поток Васерштейна функционала  $\mathcal{F}_{\text{FP}}$  (4.5).

**JKO-схема.** В общем случае задача вычисления градиентных потоков Васерштейна является сложной. Обычно решение в явной форме получить не удается и поэтому приходится применять методы численной аппроксимации. Джордан, Киндерлерер и Отто предложили метод (позже получивший сокращенное название “JKO-интегрирование”) для аппроксимации динамики  $\rho_t$  в (4.2) (см. [5]). В этом методе осуществляется дискретное по времени уточнение непрерывного потока, заданного следующим образом:

$$\rho^{(k)} \leftarrow \arg \min_{\rho \in \mathcal{P}_2(\mathbb{R}^n)} \left[ \mathcal{F}(\rho) + \frac{1}{2h} \mathcal{W}_2^2(\rho^{(k-1)}, \rho) \right], \quad (4.6)$$

где  $\rho^{(0)} = \rho^0$  – начальное условие и  $h > 0$  – величина шага дискретизации по времени. Дискретный временной градиентный поток сходится к непрерывному потоку при  $h \rightarrow 0$ , т.е.  $\rho^{(k)} \approx \rho_{kh}$ . Этот метод был далее развит в [86; 87]. Тем не менее, стоит отметить, что выполнение JKO-итераций является сложной задачей из-за наличия минимизации по  $\mathcal{W}_2$ . Типичным подход к выполнению JKO-шага основан на дискретизации пространственного домена. При размере носителя  $\lesssim 10^6$  уравнение (4.6) может быть решено стандартными алгоритмами задачи оптимального транспорта [6]. В размерностях  $D \geq 3$  дискретные носители в общем случае не точно приближают непрерывные распределения и, как следствие, динамику градиентных потоков. Чтобы обойти эту трудность в [11] предложен стохастический параметрический метод для аппроксимации плотности  $\rho_t$ . В методе работы [11] используется регуляризованный по энтропии оптимальный транспорт (ОТ). Однако его использование приводит к смешению.

### 4.3 Вычисление градиентных потоков Васерштейна с использованием сетей ICNN

Опишем наш подход к вычислению градиентных потоков Васерштейна черезJKO-схему с использованием сетей ICNN.

#### 4.3.1 Переформулировка JKO-схемы с использованием оптимальных операторов отображения

Наша ключевая идея состоит в том, чтобы заменить оптимизацию (4.6) по вероятностным мерам на оптимизацию по выпуклым функциям – данная идея навеяна работой [98]. Согласно теореме Бренье для любого  $\rho \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$  существует единственный  $\rho^{(k-1)}$ -измеримый градиент  $\nabla\psi : \mathbb{R}^D \rightarrow \mathbb{R}^D$  выпуклой функции  $\psi$ , удовлетворяющий  $\rho = \nabla\psi\#\rho^{(k-1)}$ . Положим  $\rho = \nabla\psi\#\rho^{(k-1)}$  и запишем (4.6) как оптимизационную задачу с выпуклой  $\psi$ :

$$\psi^{(k)} \leftarrow \arg \min_{\text{Convex } \psi} \left[ \mathcal{F}(\nabla\psi\#\rho^{(k-1)}) + \frac{1}{2h} \mathcal{W}_2^2(\rho^{(k-1)}, \nabla\psi\#\rho^{(k-1)}) \right]. \quad (4.7)$$

Для перехода на следующий шаг JKO-схемы определим  $\rho^{(k)} \stackrel{\text{def}}{=} \nabla\psi^{(k)}\#\rho^{(k-1)}$ .

Так как  $\rho$  является образом под действием отображения  $\rho^{(k-1)}$  градиента выпуклой функции  $\nabla\psi$ , то член  $\mathcal{W}_2^2$  в (4.7) может быть вычислен явно, что приводит к упрощению члена для расстояния Васерштейна-2 в (4.7):

$$\psi^{(k)} \leftarrow \arg \min_{\text{Convex } \psi} \left[ \mathcal{F}(\nabla\psi\#\rho^{(k-1)}) + \frac{1}{2h} \int_{\mathbb{R}^D} \|x - \nabla\psi(x)\|_2^2 d\rho^{(k-1)}(x) \right]. \quad (4.8)$$

При такой переформулировке мы избегаем сложностей, связанных с вычислением расстояния Васерштейна-2. Дополнительным преимуществом является возможность получения *выборки* из  $\rho^{(k)}$ . Так как  $\rho^{(k)} = [\nabla\psi^{(k)} \circ \dots \circ \nabla\psi^{(1)}]^\# \rho^0$ , то можно выбрать  $x_0 \sim \rho^{(0)}$ , и тогда  $\nabla\psi^{(k)} \circ \dots \circ \nabla\psi^{(1)}(x_0)$  дает выборку из  $\rho^{(k)}$ . Более того, если функции  $\psi^{(\cdot)}$  сильно выпуклы, то градиенты  $\nabla\psi^{(\cdot)}$  обратимы. В этом случае *плотность*  $\frac{d\rho^{(k)}}{dx}$  для  $\rho^{(k)} = \nabla\psi^{(k)} \circ \dots \circ \nabla\psi^{(1)} \#\rho^0$  может быть вычислена по формуле замены переменных (в предположении, что  $\psi^{(\cdot)}$

дважды дифференцируема)

$$\frac{d\rho^{(k)}}{dx}(x_k) = [\det \nabla^2 \Psi^{(k)}(x_{k-1})]^{-1} \cdots [\det \nabla^2 \Psi^{(1)}(x_0)]^{-1} \cdot \frac{d\rho^{(0)}}{dx}(x_0), \quad (4.9)$$

где  $x_i = \nabla \Psi^{(i)}(x_{i-1})$  для  $i = 1, \dots, k$  и  $\frac{d\rho^{(0)}}{dx}$  – плотность  $\rho^{(0)}$ .

### 4.3.2 Стохастическая оптимизация для JKO-схемы с использованием сетей ICNN

В общем, решение  $\Psi^{(k)}$  задачи (4.8) получить сложно, поскольку в нем участвует оптимизация по всем выпуклым функциям. Чтобы обойти эту сложность, в работе [98] используется дискретизация в пространства выпуклой функции. При этом подход также требует дискретизации мер  $\rho^{(k)}$ , что заведомо ограничивает использование этого метода только на случай малых размерностей.

Мы предлагаем параметризацию пространства поиска с использованием выпуклых по входу нейронных сетей (ICNN) [20], удовлетворяющих универсальному свойству аппроксимации среди выпуклых функций [31]. Напомним, что сети ICNN являются параметрическими моделями вида  $\Psi_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ , где функция  $\Psi_\theta$  выпукла по входу. Сеть ICNN строится из обычных слоев нейронных сетей с добавлением ограничений на веса и функции активации для сохранения выпуклости по входу (см. [20, §3.1]). Параметры оптимизируются с использованием оптимизационных методов глубокого обучения, таких, как метод стохастического градиентного спуска (SGD).

В этой постановке шаг схемы JKO состоит в нахождении оптимальных параметров  $\theta^*$  для  $\Psi_\theta$ :

$$\theta^* \leftarrow \arg \min_{\theta} \left[ \mathcal{F}(\nabla \Psi_\theta \# \rho^{(k-1)}) + \frac{1}{2h} \int_{\mathbb{R}^D} \|x - \nabla \Psi_\theta(x)\|_2^2 d\rho^{(k-1)}(x) \right]. \quad (4.10)$$

Если функционал  $\mathcal{F}$  может быть стохастически оценен с использованием случайных пакетов из  $\rho^{(k-1)}$ , то метод стохастического градиентного спуска может быть использован для оптимизации  $\theta$ . Примером такого функционала является функционал  $\mathcal{F}_{\text{FP}}$ , заданный в (4.5).

**Теорема 4.3.1** (оценка для  $\mathcal{F}_{\text{FP}}$ ). *Пусть  $\rho \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$  и  $T : \mathbb{R}^D \rightarrow \mathbb{R}^D$  – диффеоморфизм и пусть  $x_1, \dots, x_N \sim \rho$  – случайная выборка. Тогда выражение*

$$[\widehat{\mathcal{U}}_T(x_1, \dots, x_N) - \beta^{-1} \widehat{\Delta\mathcal{E}}_T(x_1, \dots, x_N)], \text{ где}$$

$$\begin{aligned}\widehat{\mathcal{U}}_T(x_1, \dots, x_N) &\stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \Phi(T(x_n)) \text{ и} \\ \widehat{\Delta\mathcal{E}}_T(x_1, \dots, x_N) &\stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \log |\det \nabla T(x_n)|,\end{aligned}$$

является оценкой для  $\mathcal{F}_{FP}(T\#\rho)$  с точностью до постоянного (относительно  $T$ ) сдвига, задаваемого  $\beta^{-1}\mathcal{E}(\rho)$ .

*Доказательство.*  $\widehat{\mathcal{U}}_T$  – несмещенная оценка для  $\mathcal{U}(T\#\rho)$ . Пусть  $p$  и  $p_T$  – плотности  $\rho$  и  $T\#\rho$  соответственно. Так как  $T$  – диффеоморфизм, имеем  $p_T(y) = p(x) \cdot |\det \nabla T(x)|^{-1}$ , где  $x = T^{-1}(y)$ . Используя формулу замены переменных, получаем

$$\begin{aligned}\mathcal{E}(T\#\rho) &= - \int_{\mathbb{R}^D} p_T(y) \log p_T(y) dy = \\ &= - \int_{\mathbb{R}^D} p(x) \cdot |\det \nabla T(x)|^{-1} \log \left[ p(x) \cdot |\det \nabla T(x)|^{-1} \right] \cdot |\det \nabla T(x)| dx = \\ &\quad - \int_{\mathbb{R}^D} p(x) \log p(x) dx + \int_{\mathbb{R}^D} p(x) \log |\det \nabla T(x)| dx = \\ &\quad \mathcal{E}(\rho) + \int_{\mathbb{R}^D} p(x) \log |\det \nabla T(x)| dx, \\ \implies \Delta\mathcal{E}_T(\rho) &\stackrel{\text{def}}{=} \mathcal{E}(T\#\rho) - \mathcal{E}(\rho) = \int_{\mathbb{R}^D} \log |\det \nabla T(x)| d\rho(x)\end{aligned}$$

что показывает, что  $\widehat{\Delta\mathcal{E}}_T$  – несмещенная оценка для  $\Delta\mathcal{E}_T(\rho)$ . В результате получаем, что  $\widehat{\mathcal{U}}_T - \beta^{-1} \widehat{\Delta\mathcal{E}}_T$  – оценка для  $\mathcal{F}_{FP}(T\#\rho) = \mathcal{U}(T\#\rho) - \beta^{-1}\mathcal{E}(T\#\rho)$  с точностью до сдвига  $\beta^{-1}\mathcal{E}(\rho)$ .  $\square$

Для того, чтобы применить теорему 4.3.1 к нашему случаю, получим  $T \leftarrow \nabla \Psi_\theta$  и  $\rho \leftarrow \rho^{(k-1)}$ , что дает стохастическую оценку для  $\mathcal{F}_{FP}(\nabla \Psi_\theta \# \rho^{(k-1)})$  в (4.10). Здесь  $\beta^{-1}\mathcal{E}(\rho^{(k-1)})$  является  $\theta$ -независимым и постоянным, поскольку  $\rho^{(k-1)}$  фиксировано, и, значит, смещение оценки не играет роли в оптимизации по  $\theta$ .

Наш стохастический JKо-метод для  $\mathcal{F}_{FP}$  реализован в алгоритме 3. Обучение проводится исключительно на основе случайных выборок из исходной меры  $\rho^0$  (её плотность не требуется).

В данном алгоритме предполагается, что  $\mathcal{F}$  – функционал Фоккера–Планка (4.5). Стоит отметить, что наш метод допускает прямое обобщение на случай

---

**Алгоритм 3:** Осуществление схемы JKO для уравнения Фоккера–Планка с помощью сетей ICNN

---

**Вход :** Начальная мера  $\rho^0$  с выборочным доступом;

шаг дискретизации схемы JKO  $h > 0$ ;

число шагов схемы JKO  $K > 0$ ;

целевой потенциал  $\Phi(x)$ ;

температура  $\beta^{-1}$  диффузационного процесса;

размер случайной выборки  $N$ ;

**Выход:** обученные ICNN модели  $\{\psi^{(k)}\}_{k=1}^K$ , представляющие шаги схемы JKO

**for**  $k = 1, 2, \dots, K$  **do**

$\psi_\theta \leftarrow$  основная модель ICNN;

**for**  $i = 1, 2, \dots$  **do**

Получение случайной выборки  $Z \sim \rho^0$  размера  $N$ ;

$X \leftarrow \nabla \psi^{(k-1)} \circ \dots \circ \nabla \psi^{(1)}(Z)$ ;

$\widehat{\mathcal{W}}_2^2 \leftarrow \frac{1}{N} \sum_{x \in X} \|\nabla \psi_\theta(x) - x\|_2^2$ ;

$\widehat{\mathcal{U}} \leftarrow \frac{1}{N} \sum_{x \in X} \Phi(\nabla \psi_\theta(x))$ ;

$\widehat{\Delta \mathcal{E}} \leftarrow \frac{1}{N} \sum_{x \in X} \log \det \Delta \psi_\theta(x)$ ;

$\widehat{\mathcal{L}} \leftarrow \frac{1}{2h} \widehat{\mathcal{W}}_2^2 + \widehat{\mathcal{U}} - \beta^{-1} \widehat{\Delta \mathcal{E}}$ ;

Выполнить градиентный шаг по  $\theta$ , используя  $\frac{\partial \widehat{\mathcal{L}}}{\partial \theta}$ ;

**end**

$\psi^{(k)} \leftarrow \psi_\theta$

**end**

---

любого функционала  $\mathcal{F}$ , допускающего стохастическое оценивание – изучение таких функций составляет многообещающую тему для будущих исследований.

#### 4.3.3 Вычисление плотности диффузационного процесса

Наш алгоритм позволяет *вычислить плотность* для  $\rho^{(k)}$ . Как отмечалось в п. 4.3.1, оказывается возможным получать выборку из  $\rho^{(k)}$  и одновременно

вычислять её плотность. Однако этот подход не дает прямого способа для оценки  $\frac{d\rho^{(k)}}{dx}(x_k)$  при любом  $x_k \in \mathbb{R}^D$ . Ниже мы решаем эту проблему.

Если выпуклая функция также является сильно выпуклой, тогда ее градиент биективен на  $\mathbb{R}^D$ . Используя формулу замены переменных для  $x_k \in \mathbb{R}^D$ , имеем  $\frac{d\rho^{(k)}}{dx}(x_k) = \frac{d\rho^{(k-1)}}{dx}(x_{k-1}) \cdot [\det \nabla^2 \Psi^{(k)}(x_{k-1})]^{-1}$  где  $x_k = \nabla \Psi^{(k)}(x_{k-1})$ . Для вычисления  $x_{k-1}$  требуется решить выпуклую оптимизационную задачу:

$$x_k = \nabla \Psi^{(k)}(x_{k-1}) \iff x_{k-1} = \arg \max_{x \in \mathbb{R}^D} [\langle x, x_k \rangle - \Psi^{(k)}(x)]. \quad (4.11)$$

Если мы знаем плотность  $\rho^0$ , то для вычисления плотности  $\rho^{(k)}$  в  $x_k$  нам следует решить  $k$  выпуклых задач

$$x_{k-1} = \arg \max_{x \in \mathbb{R}^D} [\langle x, x_k \rangle - \Psi^{(k)}(x)] \quad \dots \quad x_0 = \arg \max_{x \in \mathbb{R}^D} [\langle x, x_1 \rangle - \Psi^{(1)}(x)]$$

для получения  $x_{k-1}, \dots, x_0$  и дальнейшей оценки плотности вида

$$\frac{d\rho_k}{dx}(x_k) = \frac{d\rho^0}{dx}(x_0) \cdot \left[ \prod_{i=1}^k \det \nabla^2 \Psi^{(i)}(x_{i-1}) \right]^{-1}.$$

Следует отметить, что приведенные выше шаги обеспечивают общий метод отслеживания положения частицы вдоль потока, а вычисление плотности является просто побочным продуктом.

#### 4.4 Вычислительные эксперименты

В этом параграфе мы оцениваем наш метод на синтетических и практических приложениях. Наш программный код написан в среде PyTorch:

<https://github.com/iamalexkorotin/LargeScaleWassersteinGradientFlows>

Эксперименты выполнялись на видеокарте GTX 1080Ti. В большинстве случаев мы выполняли несколько случайных перезапусков, чтобы получить среднее значение и вариацию рассматриваемой метрики. В результате для экспериментов потребовалось около 100–150 часов вычисления. Детали приведены в п. 4.5.

**Архитектура нейронных сетей.** Во всех экспериментах мы используем архитектуру DenseICNN (п. 2.7) для  $\Psi_\theta$  в алгоритме 3 с *SoftPlus* активациями.

Сеть  $\psi_\theta$  дважды дифференцируема по входу  $x$  и имеет биективный градиент  $\nabla\psi_\theta : \mathbb{R}^D \rightarrow \mathbb{R}^D$  с неотрицательно определенным гессианом  $\nabla^2\psi_\theta(x) \succeq 0$  при каждом  $x$ . Для вычисления  $\nabla\psi_\theta$  и  $\nabla^2\psi_\theta$  мы используем автоматическое дифференцирование.

**Метрика.** Чтобы качественно сравнить меры, мы используем симметричную дивергенцию Кульбака–Лейблера

$$\text{SymKL}(\rho_1, \rho_2) \stackrel{\text{def}}{=} \text{KL}(\rho_1 \| \rho_2) + \text{KL}(\rho_2 \| \rho_1), \quad (4.12)$$

где  $\text{KL}(\rho_1 \| \rho_2) \stackrel{\text{def}}{=} \int_{\mathbb{R}^D} \log \frac{d\rho_1}{d\rho_2}(x) d\rho_1(x)$  – дивергенция Кульбака–Лейблера. Для методов, основанных на симуляции частиц, мы получаем аппроксимацию распределения путем оценки ядерной плотности.

#### 4.4.1 Сходимость к стационарному решению

Начиная с произвольной начальной меры  $\rho^0$ , процесс адвекции-диффузии (4.4) сходится к единственному стационарному решению  $\rho^*$  [99] с плотностью

$$\frac{d\rho^*}{dx}(x) = Z^{-1} \exp(-\beta\Phi(x)), \quad (4.13)$$

где  $Z = \int_{\mathbb{R}^D} \exp(-\beta\Phi(x)) dx$  – нормировочная константа. Это свойство открывает возможность для вычисления симметричной дивергенции Кульбака–Лейблера между распределением, к которому наш метод сходится, и эталонным распределением (при условии, что мы знаем  $Z$ ).

Мы используем  $\mathcal{N}(0, 16I_D)$  в качестве начальной меры  $\rho^0$  и случайные смеси гауссовских распределений в качестве стационарной меры  $\rho^*$ . В нашем методе мы выполняем  $K = 40$  шагов схемы JKO с величиной шага  $h = 0.1$ . Выполняется сравнение с методом, основанным на симуляции частиц (с  $10^3, 10^4, 10^5$  частицами), в котором используется аппроксимация Эйлера–Маруямы [EM] (см. [77, §9.2]). Эксперимент повторялся пять раз; усредненные результаты показаны на рис. 4.1.

На рис. 4.2 представлены качественные результаты нашего метода, сходящегося к эталону в размерностях  $D = 13, 32$ .

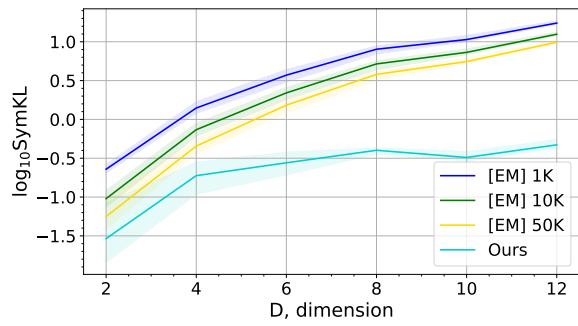


Рисунок 4.1 — Симметричная дивергенция Кульбака–Лейблера SymKL между вычисленной и стационарной мерой  $D = 2, 4, \dots, 12$

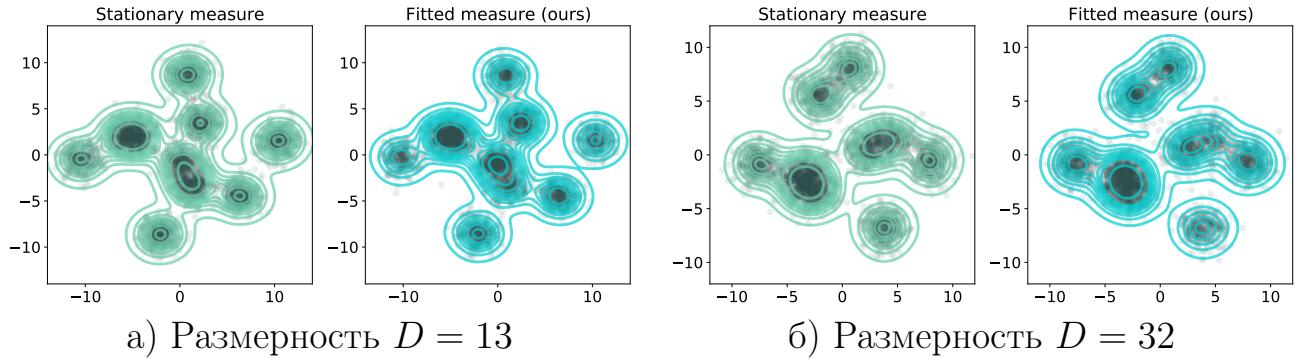


Рисунок 4.2 — Проекции на две первые РСА-компоненты (анализ главных компонент) истинной стационарной меры и меры, полученной нашим методом для размерностей  $D = 13$  (слева) и  $D = 32$  (справа).

#### 4.4.2 Моделирование процесса Орнштейна–Уленбека

Процессы Орнштейна–Уленбека являются диффузионными процессами адвекции–диффузии (4.4) с  $\Phi(x) = \frac{1}{2}(x - b)^T A(x - b)$  для симметричной положительно определенной матрицы  $A \in \mathbb{R}^{D \times D}$  и  $b \in \mathbb{R}^D$ . Они представляют собой один из немногих примеров, в которых аналитически известны  $\rho_t$  для любого  $t \in \mathbb{R}^+$ , когда начальная мера  $\rho^0$  является гауссовой (см. [100]). Это позволяет нам количественно оценить вычисленную динамику процесса, а не только стационарную меру.

Пусть  $A, b$  выбраны случайно и пусть  $\rho^0$  – стандартная гауссовская мера  $\mathcal{N}(0, I_D)$ . Мы аппроксимируем динамику процесса при помощи нашего метода с шагом JKO-схемы  $h = 0.05$  и вычисляем дивергенцию compute SymKL между истинной мерой  $\rho_t$  и приближенной мерой в момент времени  $t = 0.5$  и  $t = 0.9$ . Мы повторяем эксперимент 15 раз для размерностей  $D = 1, 2, \dots, 12$ ;

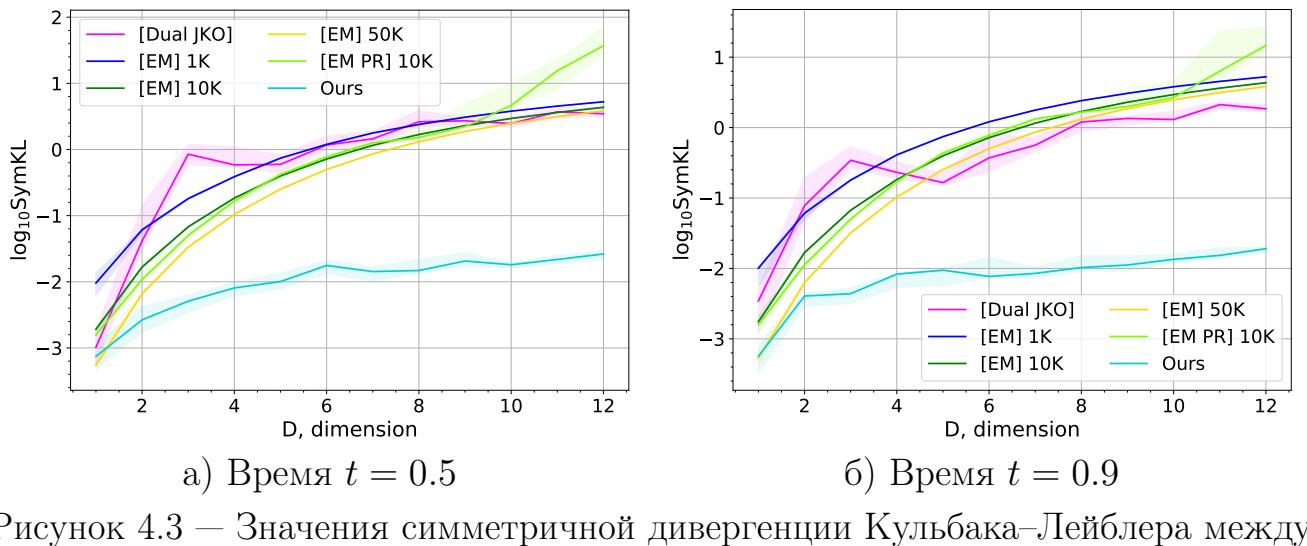


Рисунок 4.3 – Значения симметричной дивергенции Кульбака–Лейблера между вычисленной мерой и истинной мерой  $\rho_t$  при  $t = 0.5$  (слева) и  $t = 0.9$  (справа) для размерностей  $D = 1, 2, \dots, 12$ .

степень эффективности его функционирования показана на рис. 4.3. Основные используемые алгоритмы: [EM] с  $10^3, 10^4, 5 \times 10^4$  частицами, алгоритм [EM PR] для моделирования частиц с оператором проксимальной рекурсии с  $10^4$  частицами [101], а также параметрический метод двойственного вывода [Dual JKO] (см. [11]) для шагов схемы JKO.

#### 4.4.3 Выборка из ненормированных апостериорных распределений в байесовской логистической регрессии

Важной задачей в байесовском машинном обучении, к которому можно применять наш алгоритм, является получение выборок из ненормированного апостериорного распределения. При заданных модельных параметрах  $x \in \mathbb{R}^D$  с априорным распределением  $p_0(x)$  и условной плотностью  $p(\mathcal{S}|x) = \prod_{m=1}^M p(s_m|x)$  данных  $\mathcal{S} = \{s_1, \dots, s_M\}$  апостериорное распределение имеет вид

$$p(x|\mathcal{S}) = \frac{p(\mathcal{S}|x)p_0(x)}{p(\mathcal{S})} \propto p(\mathcal{S}|x)p_0(x) = p_0(x) \cdot \prod_{m=1}^M p(s_m|x).$$

Вычисление нормировочной постоянной  $p(\mathcal{S})$  обычно является трудноразрешимой задачей, что подчеркивает необходимость в методах оценки, в которых необходимо получать выборки из  $p(\mathcal{S}|x)$ , известной с точностью до нормировочной постоянной.

В нашей постановке выборка из  $p(x|\mathcal{S})$  может быть осуществлена, как в задаче из п. 4.4.1. Из (4.13) следует, что процесс адвекции-диффузии с температурой  $\beta > 0$  и  $\Phi(x) = -\frac{1}{\beta} \log [p_0(x) \cdot p(\mathcal{S}|x)]$  в качестве стационарного распределения имеет  $\frac{d\rho^*}{dx}(x) = p(x|\mathcal{S})$ . Таким образом, мы можем использовать наш метод для аппроксимации диффузионного процесса, получая в результате выборку для  $p(x|\mathcal{S})$ .

Потенциальную энергию  $\mathcal{U}(\rho) = \int_{\mathbb{R}^D} \Phi(x)d\rho(x)$  можно эффективно оценить, используя прием, аналогичный приему из стохастической градиентной динамики Ланжевена (stochastic gradient Langevin dynamics) [70]. В этом приеме повторные выборки из  $\mathcal{S}$  осуществляются равномерно. Для оценки мы рассматриваем байесовскую линейную регрессию из [102]. Используется 8 наборов данных из [103]. Число признаков варьируется от 2 до 60, размер набора данных – от 700 до 7400 точек. Мы также используем набор данных Covertype<sup>1</sup> с 500К точками данных и 54 признаками. Априорное распределение для коэффициента регрессии  $w$  имеет вид  $p_0(w|\alpha) = \mathcal{N}(w|0, \alpha^{-1})$  с  $p_0(\alpha) = \text{Gamma}(\alpha|1, 0.01)$ , поэтому априорное распределение с параметрами  $x = [w, \alpha]$  модели задается формулой  $p_0(x) = p_0(w, \alpha) = p_0(w|\alpha) \cdot p_0(\alpha)$ . Мы случайно разбиваем каждый набор данных на обучаемую часть  $\mathcal{S}_{\text{train}}$  и тестируемую часть  $\mathcal{S}_{\text{test}}$  с отношением 4:1, после чего применяется вывод для апостериорного распределения  $p(x|\mathcal{S}_{\text{train}})$ . В табл. 8 показаны точность и логарифмическое правдоподобие прогнозного распределения на  $\mathcal{S}_{\text{test}}$ . В качестве основного метода для сравнения использовался метод Штейна вариационного градиентного спуска с симуляцией на частицах (particle-based Stein Variational Gradient Descent; см. [102]). Использована авторская реализация с гиперпараметрами, заданными по умолчанию.

## 4.5 Технические детали экспериментов

**Общие детали.** Мы используем архитектуру DenseICNN (см. п. 2.7.2) для  $\Psi_\theta$  с двумя скрытыми слоями и изменяем ширину модели в зависимости от задачи. Также используется оптимизатор Adam, у которого скорость обучения уменьшается с увеличением JKO-шагов. Инициализация модели ICNN происходит или через предварительное обучение для удовлетворения условия

---

<sup>1</sup><https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/binary.html>

Набор данных	Точность		Лог. правдоподоб.	
	Наш	[SVGD]	Наш	[SVGD]
covtype	0.75	0.75	-0.515	-0.515
german	0.67	0.65	-0.6	-0.6
diabetis	0.775	0.78	-0.45	-0.46
twonorm	0.98	0.98	-0.059	-0.062
ringnorm	0.74	0.74	-0.5	-0.5
banana	0.55	0.54	-0.69	-0.69
splice	0.845	0.85	-0.36	-0.355
waveform	0.78	0.765	-0.485	-0.465
image	0.82	0.815	-0.43	-0.44

Таблица 8 — Сравнение нашего метода с [SVGD] [102] для байесовской логистической регрессии.

$\nabla\psi_\theta(x) \approx x$ , или путем использования параметров  $\theta$ , полученных из предыдущего JKO-шага.

Для [Dual JKO] мы использовали реализацию, предложенную авторами, при этом гиперпараметры выбирались по умолчанию. Для [EM PR] оператор проксимальной рекурсии (Proximal Recursion operator) был реализован согласно псевдокоду из [101]. Для целей сравнения с простым алгоритмом [EM] мы использовали этот оператор с гиперпараметрами, заданными по умолчанию, но с увеличенным числом частиц. Отметим, что из-за высокой вычислительной сложности метода число частиц было ограничено сверху константой  $N = 10^4$ . Для [SVGD] мы использовали официальную реализацию, доступную по адресу

<https://github.com/dilinwang820/Stein-Variational-Gradient-Descent>

В моделировании по методу частиц для [EM], [BBF] и [EM PR] мы использовали временной шаг распространения частиц  $dt = 10^{-3}$ .

Величина SymKL (см. (4.12)) оценивалась по методу Монте-Карло (MC) на  $10^4$  выборках. В нашем методе оценка по методу Монте-Карло является прямой, поскольку метод позволяет осуществлять как выборку, так и подсчет плотности. В методах на основе симуляции частиц мы используем ядерную оценку плотности для аппроксимации плотности с использованием реализации `scipy` для `gaussian_kde` с шириной полосы (`bandwidth`), выбираемой по правилу Скотта. В [Dual JKO] используется процедура выборки с оценкой важности (importance sampling procedure) и оценкой нормировочной постоянной, как в [11]. Мы используем  $\beta = 1$  во всех наших экспериментах.

$D$	$M$	$l$	$w$
2	5	10	256
4	6	10	384
6	7	10	512
8	8	10	512
10	9	10	512
12	10	10	1024
13	10	10	512
32	10	6	1024

Таблица 9 — Гиперпараметры, используемые в эксперименте.

#### 4.5.1 Сходимость к стационарному распределению

В качестве стационарной меры  $\rho^*$  мы рассматриваем случайные смеси гауссовых распределений  $\frac{1}{N_p} \sum_{m=1}^M \mathcal{N}(\mu_m, I_D)$ , где  $\mu_1, \dots, \mu_M \sim \text{Uniform}\left([- \frac{l}{2}, \frac{l}{2}]^D\right)$ . Предполагается, что ширина  $w$  используемой сети ICNN  $\psi_\theta$  зависит от размерности  $D$ . Используемые параметры приведены в табл. 9.

На каждом JKO-шаге используется 1000 итераций градиентного спуска в алгоритме 3. Для размерностей  $D = 2, 4, \dots, 12$  первые 20 JKO-переходов оптимизируются с  $lr = 5 \cdot 10^{-3}$ , а на оставшихся шагах используется  $lr = 2 \cdot 10^{-3}$ . Для качественных экспериментов при  $D = 13, 32$  мы выполняем 50 и 70 JKO-шагов с длиной шага  $h = 0.1$ . В этих случаях параметры для определения скорости обучения аналогичны параметрам для проведения количественных экспериментов, но в первом случае выполняется дополнительный цикл с  $lr = 5 \cdot 10^{-4}$  на заключительных шагах JKO. Размер выборки:  $N = 512$ .

#### 4.5.2 Моделирование процессов процесса Орнштейна–Уленбека

Матрицы  $A \in \mathbb{R}^{D \times D}$  случайно генерируются с использованием `sklearn.datasets.make_spd_matrix`. Векторы  $b \in \mathbb{R}^D$  выбираются из стандартной гауссовой меры. Для всех сетей ICNN  $\psi_\theta$  мы полагаем  $w = 64$ ; каждая из них обучается за 500 итераций за один шаг JKO с  $lr = 5 \cdot 10^{-3}$  и размером выборки  $N = 1024$ .

Набор данных	$w$	$lr$	$iter$	batch	$K$
covtype	512	$2 \cdot 10^{-5}$	$10^4$	1024	6
german	512	$2 \cdot 10^{-4}$	5000	512	5
diabetis	128	$5 \cdot 10^{-5}$	6000	1024	16
twonorm	512	$5 \cdot 10^{-5}$	5000	1024	7
ringnorm	512	$5 \cdot 10^{-5}$	5000	1024	2
banana	128	$2 \cdot 10^{-4}$	5000	1024	5
splice	512	$2 \cdot 10^{-3}$	2000	512	5
waveform	512	$5 \cdot 10^{-5}$	5000	512	2
image	512	$5 \cdot 10^{-5}$	5000	512	5

Таблица 10 — Гиперпараметры, используемые в экспериментах по байесовской логистической регрессии.

#### 4.5.3 Получение выборки из апостериорного распределения

Для избавления от условия положительности  $\alpha$  мы рассматриваем  $[w, \log(\alpha)]$  в качестве параметров регрессионной модели (вместо  $[w, \alpha]$ ). Для обучения апостериорного распределения  $p(x|S_{\text{train}})$  мы используем размер JКО-шага  $h = 0.1$ . Пусть  $iter$  – число градиентных шагов по  $\theta$  на каждый JКО-шаг. Использованные гиперпараметры для каждого набора данных указаны в табл. 10.

Для оценки логарифмического правдоподобия и точности прогнозируемого распределения на  $S_{\text{test}}$  по  $p(x|S_{\text{train}})$  мы используем прямую оценку по методу Монте-Карло на  $2^{12}$  случайных выборках параметров.

### 4.6 Обсуждение

**Вычислительная сложность обучения и получения выборки.** Пусть  $T$  – число операций, требуемых для вычисления  $\Psi_\theta(x)$  сетью ICNN. Предположим, что временные затраты на вычисление  $\Phi(x)$  в потенциальной энергии  $\mathcal{U}$  составляют по порядку  $O(1)$ .

Напомним, что вычислительные затраты на вычисление градиента пре-восходят лишь на постоянный множитель затраты на вычисление самой функции (см. [104]). Таким образом, вычисление градиента  $\nabla \Psi_\theta(x) : \mathbb{R}^D \rightarrow \mathbb{R}^D$

Операция	Временная сложность
Вычисл. $\Psi_\theta, \nabla \Psi_\theta, \nabla^2 \Psi_\theta$	$T, O(T), O(DT)$
Вычисл. $\log \det \nabla^2 \Psi_\theta$	$O(DT + D^3)$
Выборк. $x \sim \rho^{(k)}$	$O((k-1)T)$
Вычисл. $\widehat{\mathcal{L}}$ на $x \sim \rho^{(k)}$	$O(DT + D^3)$
Вычисл. $\frac{\partial \widehat{\mathcal{L}}}{\partial \theta}$ на $x \sim \rho^{(k)}$	$O(DT + D^3)$
Sample $x \sim \rho^{(k)}$ и Eval. $\frac{d\rho^{(k)}}{dx}(x)$	$O((k-1)(TD + D^3))$

Таблица 11 — Вычислительная сложность операций в нашем метод для вычисления шагов схемы JKO с помощью сетей ICNN.

требует  $O(T)$  операций, а временные затраты на вычисление гессиана  $\nabla^2 \Psi_\theta(x) : \mathbb{R}^D \rightarrow \mathbb{R}^{D \times D}$  равны  $O(DT)$ . Для вычисления  $\log \det \nabla^2 \Psi_\theta(x)$  требуется  $O(D^3)$  дополнительных операций. Получение выборки из  $\rho^{(k-1)} = \nabla \Psi^{(k-1)} \circ \dots \circ \nabla \Psi^{(1)} \# \rho_0$  включает в себя прямое отображение  $x_0 \sim \rho^0$  при помощи последовательности сетей  $\Psi^{(\cdot)}$  длины  $k-1$ , что требует  $O((k-1)T)$  операций. Прямое прохождение для вычисления целевой функции  $\widehat{\mathcal{L}}$  в шаге схемы JKO в алгоритме 3 требует  $O(DT + D^3)$  операций, равно как и операция обратного прохода для вычисления градиента  $\frac{\partial \widehat{\mathcal{L}}}{\partial \theta}$  по  $\theta$ .

Более сложной оказывается оценка *затрат памяти*, поскольку они зависят от реализации автодифференцирования. Затраты, связанные с памятью, не превосходят временную сложность алгоритма и являются линейными по числу  $k$  шагов схемы JKO.

**Физическое время.** Все основанные на моделировании частиц методы, которые были рассмотрены в п.4.4 и [Dual JKO], требуют от нескольких секунд до нескольких минут процессорного времени. Наш метод требует от нескольких минут до нескольких часов процессорного времени на графическом процессоре (это время объясняется необходимостью обучения новой сети на каждом шаге).

**Преимущества.** Благодаря использованию непрерывной аппроксимации наш метод хорошо масштабируется на случай больших размерностей, что будет показано в п. 4.4.1 и п. 4.4.2. После обучения можно организовать бесконечное множество выборок  $x_k \sim \rho^{(k)}$  вместе с их траекториями  $x_{k-1}, x_{k-2}, \dots, x_0$  вдоль потока градиента. Более того, плотности выборок в потоке flow  $\frac{d\rho^{(k)}}{dx}(x_k), \frac{d\rho^{(k-1)}}{dx}(x_{k-1}), \dots, \frac{d\rho^{(0)}}{dx}(x_0)$  могут быть вычислены непосредственно.

Напротив, методы, основанные на моделировании частиц, и методы, основанные на дискретизации доменов, обладают плохой способностью к масштабированию с изменением размерности (рис. 4.3), и по ним нельзя получить плотность. Несмотря на применяемую в [Dual JKO] параметрическую аппроксимацию, функциональные характеристики метода [Dual JKO] сравнимы с характеристиками методов, основанных на моделировании методом частиц, но хуже, чем у нашего метода (см. также рис. 3 в работе [11]).

**Ограничения.** Для обучения  $k$  JKO-шагов временные затраты нашего метода пропорциональны  $k^2$  из-за повышенной сложности, связанной с выборкой  $x \sim \rho^{(k)}$ . Это может быть невыгодным при обучении длинных диффузий. Также для очень больших размерностей  $D$  точное вычисление  $\log \det \nabla^2 \Psi_\theta(x)$  является весьма затратным по времени.

**Возможные направления дальнейших исследований.** Для снижения вычислительной сложности процесса получения выборок из  $\rho^{(k)}$  на шаге  $k$  можно обучить обратимую сеть  $H : \mathbb{R}^D \rightarrow \mathbb{R}^D$  [105; 106], для которой выполнено  $H(x_0) \approx \nabla \Psi^{(k)} \circ \dots \circ \nabla \Psi^{(1)}(x_0)$ , и использовать  $H \sharp \rho_0 \rightarrow \rho^{(k)}$  для упрощения процесса получения выборок. Другим вариантом является использование вариационных выводов (см. [107—109] для аппроксимации  $\rho^{(k)}$ ). Чтобы уменьшить сложность вычисления  $\log \det \nabla \Psi_\theta(x)$ , можно использовать быструю аппроксимацию [110; 111]. Отметим, что разработка сетей ICNN с легко вычисляемыми точными гессианами является чрезвычайно важна для дальнейших исследований, поскольку сети ICNN продолжают привлекать внимание для задач, связанных с машинным обучением [7; 52; 112; 113].

**Возможные применения.** Процессы диффузии имеют важное значение в решении многочисленных научных и промышленных задач, включая машинное обучение, финансы, физику и динамику популяций. Наш метод улучшает модели в этих областях, обеспечивая лучшую масштабируемость. Производительность, однако, может зависеть от сетей ICNN, что подчеркивает исключительную важность теоретического анализа сходимости в будущих исследованиях для увеличения доверия к нашей модели.

Подводя итог всему вышесказанному, мы разработали эффективный метод для моделирования диффузионных процессов, возникающих во многих практических задачах. Мы применяем наш метод к общей байесовской задаче семплирования из ненормированного апостериорного распределения (п. 4.4.3). Также отметим ряд потенциальных приложений:

- **Динамика популяций.** В этой задаче требуется восстановить  $\Phi(x)$  в функционале свободной энергии Фоккера–Планка  $\mathcal{F}_{\text{FP}}$ , основываясь на выборках из диффузии, полученной с временными шагами  $t_1, \dots, t_n$  (см. [114]). Такого рода задачи возникают в вычислительной биологии (см. п. 6 в [114]). В недавней статье [115] для моделирования динамики популяций используются JКО, основанные на сетях ICNN.
- **Обучение с подкреплением (Reinforcement learning).** Градиентные потоки Васерштейна дают теоретически обоснованный способ для оптимизации политики агента в обучении с подкреплением (см. [92; 93]). Идея этого метода состоит в максимизации ожидаемого общего вознаграждения (см. (10) в [93]) с использованием градиента потока, связанного с функционалом Фоккера–Планка (см. (12) в [93]). Авторы оригинальной статьи предложили дискретный метод аппроксимации решения соответствующей JКО-схемы с помощью симуляции частиц. Можно ожидать улучшения полученных результатов при комбинации их подхода с нашей JKO-схемой, основанной на сетях ICNN-based.
- **Уточнение генеративных состязательных сетей.** В задачах, связанных с генеративными состязательными сетями (GAN), можно использовать дискриминатор  $D$  для улучшения выборки из  $G$ . Для этого, можно рассмотреть градиентный поток относительно регуляризованной энтропией  $f$ -дивергенции между реальными и сгенерированными распределениями данных (см. [91], и, в частности, формулу (4)). Использование дивергенции Кульбака–Лейблера позволит применить наш метод для нахождения потока градиента: функционал  $\mathcal{F}$ , определяющий поток, имеет только энтропийные члены и члены, входящие в потенциальную энергию. Использование нашего метода вместо метода симуляции частиц может улучшить модель генератора  $G$ .
- **Молекулярный анализ.** В работе [116], параллельно нашим исследованиям, также была предложена JKO-ICNN схема. В этой работе в качестве приложения авторы рассматривают задачи молекулярного анализа, и, в частности, задачу повышения лекарственных свойств для заданного распределения  $\rho$  молекул при условии близости к оригинальному распределению  $\rho_0$ . Эта задача сводится к задаче оптимизации функционала  $\mathcal{F}(\rho) = \mathbb{E}_{x \sim \rho} \Phi(x) + \mathcal{D}(\rho, \rho_0)$  при некотором потенциале  $\Phi$  ( $V$  в обозначениях работы [116]) и несоответствия (discrepancy)  $\mathcal{D}$ . Метод

JKO-ICNN был применен для минимизации  $\mathcal{F}$  на наборе данных MOSES [117] (молекулы). Полученные результаты являются весьма перспективными.

## Глава 5. Методология построения Васерштейн-2 эталонных пар вероятностных распределений

В последнее время большой популярностью пользуются основанные на нейронных сетях численные методы (*решатели*) для задачи 1 оптимального транспорта (ОТ). Тем не менее в настоящее время до сих пор не существует стандартный способ для оценки их функциональных характеристик. В настоящей главе мы рассматриваем этот вопрос для задачи оптимального транспорта с квадратичной стоимостью, более конкретно, для расстояния Васерштейн-2, широко используемого в задачах оптимального транспорта в машинном обучении. Чтобы преодолеть проблему вычисления эталонных транспортных отображений между непрерывными мерами, необходимыми для оценки этих решателей, мы используем выпуклые по входу нейронные сети (ICNN), чтобы построить пары мер, для которых эталонные транспортные отображения можно получить аналитически (задача 4). Эта стратегия дает пары непрерывных эталонных мер в таких многомерных пространствах, как пространства изображений. Мы аккуратно оцениваем существующие оптимальные решатели для задачи оптимального транспорта с использованием этих эталонных мер. Несмотря на то, что эти решатели хорошо работают в ряде текущих задач, многие из них плохо восстанавливают оптимальные транспортные отображения. Чтобы исследовать причину такого несоответствия, мы дополнительное проверяем решатели в задачах генерации изображений. Наши исследования позволяют выявить основные ограничения, присущие существующим решателям, и показывают, что повышенная точность в задаче оптимального транспорта не обязательно коррелирует с лучшими результатами по решению практических задач.

### 5.1 Введение

В настоящий момент в машинном обучении широко используются *параметрические* (или т.н. *непрерывные*) методы для решения задач оптимального транспорта (ОТ). Эти методы включают в себя методы для крупномасштабных задач оптимального транспорта [2; 58] и методы, связанные с популярной

генеративной состязательной сетью Васерштейна (W-GAN) [1; 35]. В отличие от методов, использующих дискретизацию задачи (см. [6]), в параметрических алгоритмах оптимального транспорта применяются нейронные сети, а также ядерные разложения (kernel expansions) для оценки транспортных отображений или двойственных решений. Такой подход позволяет масштабировать задачи оптимального транспорта на крупномасштабные и многомерные задачи, не решаемые дискретными методами. Существенный успех применения параметрических методов решения задач оптимального транспорта достигнут в генеративном моделировании [43; 118—120] и задачах адаптации доменов [59; 121; 122].

В таких приложениях стоимость оптимального транспорта обычно используется как функция потерь для генеративных нейронных сетей. Например, в сетях W-GAN стоимость оптимального транспорта используется как функция потерь для генератора; данная модель включает в себя основанный на нейронных сетях решатель задачи оптимального транспорта, используемый для оценки потерь. Отметим, что недавно введенные сети W-GAN обеспечивают генеративную производительность на современном уровне. Однако остается неясным, в какой степени этот успех связан с задачей оптимального транспорта. Например, в работах [123—125] показывается, что популярные решатели для расстояния Васерштейна-1 ( $\mathbb{W}_1$ ) в сетях GAN не могут корректно оценить расстояние  $\mathbb{W}_1$ . Несмотря на то, что сети W-GANs были изначально введены для расстояния  $\mathbb{W}_1$  в [1], в настоящее время современные решатели используют оба расстояния  $\mathbb{W}_1$  и  $\mathbb{W}_2$  ( $\mathbb{W}_2$  – расстояние *Васерштейна-2*, т.е. расстояние с квадратичной функцией стоимости в задаче оптимального транспорта). Эксперименты показывают, что для сетей GAN такие расстояния показывают сходные функциональные характеристики, но  $\mathbb{W}_2$ -решатели имеют тенденцию сходиться быстрее (см. табл. 4 в работе [43]) и с лучшими теоретическими оценками [7; 43].

**Краткое описание полученных результатов главы.** В настоящей главе разработан общий подход к оценке параметрических решателей задачи оптимального транспорта с квадратичной стоимостью ( $\mathbb{W}_2$ ). Основные результаты состоят в следующем:

- Выпуклые по входу нейронные сети (ICNN [20]) используются для построения эталонных пар непрерывных мер с аналитически известным (по построению) ОТ отображением между ними (п. 5.4, п. 5.5.1).
- Этalonные меры используются для оценки популярных параметрических решателей оптимального транспорта с квадратичной стоимостью в

пространствах большой размерности (п. 5.5.3), включая пространство изображений лиц размера  $64 \times 64$  из набора данных CelebA (п. 5.5.4).

- Решатели задачи оптимального транспорта 1 тестируются в качестве оценщиков функции потерь при генеративном моделировании изображений (п. 5.5.5).

Наши эксперименты показывают, что некоторые решатели задачи оптимального транспорта выдают большую ошибку даже в малых размерностях (п. 5.5.3) и показывают результаты, сравнимые с тривиальными подходами (п. 5.5.2). В наиболее успешных решателях используется параметризация сетей ICNN. Однако достаточно удивительным оказывается то, что решатели, правильно восстанавливающие  $\mathbb{W}_2$ , не достигают наилучших показателей в задачах генеративного моделирования.

Наши эталонные меры могут быть использованы для оценки перспективных  $\mathbb{W}_2$ -решателей в пространствах большой размерности, что является ключевым шагом для улучшения прозрачности и воспроизводимости параметрических численных методов оптимального транспорта. Отметим, что методология [126] не удовлетворяет этой цели, поскольку она предназначена для тестирования дискретных методов оптимального транспорта и использует дискретные меры малой размерности с ограниченным носителем.

**Обозначения.** В главе  $\mathcal{P}_2(\mathbb{R}^D)$  обозначает множество борелевских вероятностных мер на  $\mathbb{R}^D$  с конечными вторыми моментами,  $\mathcal{P}_{2,ac}(\mathbb{R}^D)$  — подмножество всех абсолютно непрерывных вероятностных мер,  $\Pi(\mathbb{P}, \mathbb{Q})$  — множество всех вероятностных мер на  $\mathbb{R}^D \times \mathbb{R}^D$  с маргиналами  $\mathbb{P}$  и  $\mathbb{Q}$ . Для заданного измеримого отображения  $T : \mathbb{R}^D \rightarrow \mathbb{R}^D$  через  $T^\sharp$  обозначим соответствующий оператор отображения. Для функции  $\varphi : \mathbb{R}^D \rightarrow \mathbb{R}$  посредством  $\bar{\varphi}$  обозначим ее преобразование Лежандра–Фенхеля [29], заданное формулой  $\bar{\varphi}(y) = \max_{x \in \mathbb{R}^D} [\langle x, y \rangle - \varphi(x)]$ . Напомним, что  $\bar{\varphi}$  есть выпуклая функция (даже при невыпуклой  $\varphi$ ).

## 5.2 Некоторые сведения из теории оптимального транспорта

Для начала напомним определение и некоторые свойства задачи 1 оптимального транспорта с квадратичной функцией стоимости. По поводу формальных постановок и доказательств см. гл. 1 работы [9].

**Постановка задачи.** Для  $\mathbb{P}, \mathbb{Q} \in \mathcal{P}_2(\mathbb{R}^D)$  определение Монжа квадрата расстояния *Васерштейна-2*, т.е. задачи оптимального транспорта с квадратичной функцией стоимости, имеет следующий вид:

$$\mathbb{W}_2^2(\mathbb{P}, \mathbb{Q}) \stackrel{\text{def}}{=} \min_{T \sharp \mathbb{P} = \mathbb{Q}} \int_{\mathbb{R}^D} \frac{\|x - T(x)\|^2}{2} d\mathbb{P}(x), \quad (5.1)$$

где минимум берется по измеримым функциям (транспортным отображениям)  $T : \mathbb{R}^D \rightarrow \mathbb{R}^D$ , отображающим  $\mathbb{P}$  в  $\mathbb{Q}$ . Оптимальное значение  $T^*$  в этой задаче называется *отображением оптимального транспорта* (отображением ОТ). Отметим, что функционал (5.1) не является симметричным и что в этой формулировке не возможно расщепление массы, т.е. при некоторых  $\mathbb{P}, \mathbb{Q} \in \mathcal{P}_2(\mathbb{R}^D)$  может не существовать отображения  $T$  со свойством  $T \sharp \mathbb{P} = \mathbb{Q}$ . В этой связи Канторович [127] предложил следующую постановку задачи:

$$\mathbb{W}_2^2(\mathbb{P}, \mathbb{Q}) \stackrel{\text{def}}{=} \min_{\pi \in \Pi(\mathbb{P}, \mathbb{Q})} \int_{\mathbb{R}^D \times \mathbb{R}^D} \frac{\|x - y\|^2}{2} d\pi(x, y), \quad (5.2)$$

где минимум берется по всем транспортным планам  $\pi$ , т.е. мерам на  $\mathbb{R}^D \times \mathbb{R}^D$ , в которых  $\mathbb{P}$  и  $\mathbb{Q}$  – маргины. Оптимальное  $\pi^* \in \Pi(\mathbb{P}, \mathbb{Q})$  называется *оптимальным транспортным планом* (ОТ-планом). Если  $\pi^*$  имеет вид  $[\text{id}_{\mathbb{R}^D}, T^*] \sharp \mathbb{P} \in \Pi(\mathbb{P}, \mathbb{Q})$  при некотором  $T^*$ , то  $T^*$  – минимизатор в задаче (5.1).

**Двойственная формулировка.** Для  $\mathbb{P}, \mathbb{Q} \in \mathcal{P}_2(\mathbb{R}^D)$  *двойственная формулировка* задачи  $\mathbb{W}_2^2$  имеет следующий вид (см. [27]):

$$\mathbb{W}_2^2(\mathbb{P}, \mathbb{Q}) = \max_{f \oplus g \leqslant \frac{1}{2}\|\cdot\|^2} \left[ \int_{\mathbb{R}^D} f(x) d\mathbb{P}(x) + \int_{\mathbb{R}^D} g(y) d\mathbb{Q}(y) \right], \quad (5.3)$$

где максимум берется по всем  $f \in \mathcal{L}^1(\mathbb{P}, \mathbb{R}^D \rightarrow \mathbb{R})$  и  $g \in \mathcal{L}^1(\mathbb{Q}, \mathbb{R}^D \rightarrow \mathbb{R})$ , удовлетворяющим  $f(x) + g(y) \leqslant \frac{1}{2}\|x - y\|^2$  при всех  $x, y \in \mathbb{R}^D$ . По оптимальному двойственному потенциалу  $f^*$  мы можем восстановить оптимальное отображение  $T^*(x) = x - \nabla f^*(x)$  (см. пояснения в теореме 1.17 из [9]).

Существуют оптимальные  $f^*$  и  $g^*$ , которые удовлетворяют условиям  $(f^*)^c = g^*$  и  $(g^*)^c = f^*$ , где  $u^c : \mathbb{R}^D \rightarrow \mathbb{R}$  –  $c$ -преобразование  $u$ , определенное формулой:  $u^c(y) = \min_{x \in \mathbb{R}^D} [\frac{1}{2}\|x - y\|^2 - u(x)]$ . Задачу (5.3) можно переписать в следующем виде:

$$\mathbb{W}_2^2(\mathbb{P}, \mathbb{Q}) = \max_f \left[ \int_{\mathbb{R}^D} f(x) d\mathbb{P}(x) + \int_{\mathbb{R}^D} f^c(y) d\mathbb{Q}(y) \right], \quad (5.4)$$

где максимум берется по всем  $f \in \mathcal{L}^1(\mathbb{P}, \mathbb{R}^D \rightarrow \mathbb{R})$ . Так как  $f^*$  и  $g^*$  получаются друг из друга  $c$ -преобразованием, каждая из этих функций является  $c$ -вогнутой (см. [9, §1.6]), что эквивалентно тому, что функции  $\Psi^* : x \mapsto \frac{1}{2}\|x\|^2 - f^*(x)$  и  $\varphi^* : x \mapsto \frac{1}{2}\|x\|^2 - g^*(x)$  выпуклы, см. предложение 1.21 в [9]. В частности,  $\overline{\Psi^*} = \varphi^*$  и  $\overline{\varphi^*} = \Psi^*$ . Так как

$$T^*(x) = x - \nabla f^*(x) = \nabla \left( \frac{\|x\|^2}{2} - f^*(x) \right) = \nabla \Psi^*, \quad (5.5)$$

мы получаем, что отображения оптимального транспорта являются градиентами выпуклых функций. Этот результат известен как теорема Бренье [26].

**Существующие подходы к “решению” задач оптимального транспорта.** В приложениях для заданных  $\mathbb{P}, \mathbb{Q} \in \mathcal{P}_2(\mathbb{R}^D)$  оптимальная транспортная задача  $\mathbb{W}_2$  обычно рассматривается в следующих трех аналогичных, но не эквивалентных постановках:

- **Вычисление  $\mathbb{W}_2^2(\mathbb{P}, \mathbb{Q})$ .** Расстояние Васерштейна-2, представляет собой геометрически осмысленный способ вычисления вероятностных мер и задает метрику на  $\mathcal{P}_2(\mathbb{R}^D)$ .
- **Вычисление оптимального отображения  $T^*$  или плана  $\pi^*$ .** Отображение  $T^*$  дает интуитивный способ интерполяции между мерами. Оно часто используется в качестве генеративного отображения между мерами в таких задачах, как, например, адаптация доменов [2; 59] и перенос стилей изображений (см. главу 2 или [59]).
- **Использование градиента  $\partial \mathbb{W}_2^2(\mathbb{P}_\alpha, \mathbb{Q}) / \partial \alpha$  для обновления генеративных моделей.** Производные  $\mathbb{W}_2^2$  неявно используются в генеративном моделировании, включающем в себя  $\mathbb{W}_2$ -потери [43; 128]; в этом случае  $\mathbb{P} = \mathbb{P}_\alpha$  – параметрическая мера и  $\mathbb{Q}$  – мера данных. В типичной ситуации  $\mathbb{P}_\alpha = G_\alpha \sharp \mathbb{S}$  – мера, сгенерированная из латентной меры  $\mathbb{S}$  с помощью параметризованной функции  $G_\alpha$ , т.е. нейронной сетью. Целью является нахождение параметров  $\alpha$ , минимизирующих  $\mathbb{W}_2^2(\mathbb{P}_\alpha, \mathbb{Q})$  с использованием метода градиентного спуска.

В постановке задачи с использованием генеративных моделей по определению отображения  $\mathbb{P}_\alpha = G_\alpha \sharp \mathbb{S}$  имеем

$$\mathbb{W}_2^2(\mathbb{P}_\alpha, \mathbb{Q}) = \int_z f^*(G_\alpha(z)) d\mathbb{S}(z) + \int_{\mathbb{R}^D} g^*(y) d\mathbb{Q}(y),$$

где  $f^*$  и  $g^*$  – оптимальные двойственные потенциалы. На каждом шаге обучения генератора  $f^*$  и  $g^*$  фиксируются, поэтому при взятии градиента по  $\alpha$  по правилу

дифференцирования сложной функции имеем

$$\frac{\partial \mathbb{W}_2^2(\mathbb{P}_\alpha, \mathbb{Q})}{\partial \alpha} = \int_z \mathbf{J}_\alpha G_\alpha(z)^T \nabla f^*(G_\alpha(z)) d\mathbb{S}(z), \quad (5.6)$$

где  $\mathbf{J}_\alpha G_\alpha(z)^T$  – транспонированная матрица Якоби для  $G_\alpha(z)$  с параметром  $\alpha$ . Используя теорему об огибающей (the envelope theorem) [129], можно доказать, что этот результат также имеет место и без предположения, что потенциалы фиксированы. Для нахождения градиента нам нужна хорошая оценка  $\nabla f^* = \text{id}_{\mathbb{R}^D} - T^*$ , для получения которой можно использовать (5.5). Эта задача немного отличается от задачи вычисления отображения оптимального транспорта  $T^*$ : поскольку оценка  $\nabla f^*$  используется только при обновлении градиента для генератора, она может отличаться, но всё равно давать хорошую генеративную модель.

Под фразой *решатель задачи оптимального транспорта* мы будем понимать метод для решения любой из упомянутых выше задач.

### **Численная оценка решателей задачи оптимального транспорта.**

Для дискретных методов задачи оптимального транспорта существует эталонный набор данных (см. [126]), но при этом механизм для генерации набора данных не переносится на случай задач оптимального транспорта между непрерывными мерами. Существующие параметрические решатели обычно оцениваются на множестве простых примеров или тестируются в генеративных моделях без оценки ее фактической производительности в задаче оптимального транспорта. Часто используются следующие две метрики:

В явных метриках вычисленное транспортное отображение  $\hat{T}$  сравнивается с истинным  $T^*$ , например, с использованием  $\mathcal{L}^2$ -метрики необъяснимого процента расхождения (Unexplained Variance Percentage;  $\mathcal{L}^2$ -UVP), см. главу 2. В настоящее время имеется сравнительно немного явных метрик, поскольку число примеров  $\mathbb{P}, \mathbb{Q}$  с известным эталоном  $T^*$  достаточно мало. Известно, что  $T^*$  может быть аналитически получено или явно вычислено в дискретном случае [6, §3], в 1-мерном случае [6, §2.6], а также в гауссовском или линейно-разбросанном случаях [36].

*Няявные метрики* применяют решатель ОТ как компоненту в многоэтапной схеме обучения. При этом качество конечного решения задачи используется как косвенная оценка качества самого решателя. Например, в генеративном моделировании, при котором оптимальный транспорт используется как функция потерь для генератора [43; 130], качество генератора можно оценить через метрики для сетей GAN, например, через метрику Fréchet Inception distance (FID) [42].

Неявные метрики не дают ясного понимания о качестве самого решателя, но зависят от компонент модели, не связанных с задачей оптимального транспорта.

### 5.3 Параметрические двойственные решатели для задач оптимального транспорта с квадратичной функцией стоимости

Несмотря на то, что наш разрабатываемый **бенчмарк** (*эталонный набор пар мер*) может быть использован для тестирования любого параметрического решателя, вычисляющего отображение  $T^*$  или градиент  $\nabla f^*$ , в настоящей главе вычисления выполняются только для *параметрических решателей*, использующих двойственные форму ОТ (5.3) или (5.4). Для таких решателей имеются простые и понятные оптимизационные процедуры, которые могут быть адаптированы к различным наборам данных без затратного поиска гиперпараметров. Напротив, основанные на форме (5.1) решатели (см., например, [59; 131—133]) обычно параметризуют  $T^*$ , используя сложные методы теории генеративного моделирования и требуют тщательного поиска гиперпараметров [134].

В табл. 12 мы приводим основные сведения о существующих параметрических решателях, использующих двойственную форму. Решатели обучаются параметрическую функцию  $f_\theta$  (или  $\Psi_\theta$ ) для аппроксимации  $f^*$  (или  $\Psi^* = \text{id}_{\mathbb{R}^D} - f^*$ ). Получающаяся в итоге функция  $f_\theta$  дает приближенное отображение задачи оптимального транспорта  $\text{id}_{\mathbb{R}^D} - \nabla f_\theta = \nabla \Psi_\theta \approx T^*$  и производную  $\nabla f_\theta = \text{id}_{\mathbb{R}^D} - \nabla \Psi_\theta$ , которая нужна для обучения генеративных моделей (5.6).

Насколько нам известно, ни один из представленных выше решателей не был оценен в количественном отношении в негауссовской постановке. Для [ММ], [ММ-В] и [QC] качество восстановленных производных  $\nabla f^*$  для  $\partial W_2^2(\mathbb{P}_\alpha, \mathbb{Q}) / \partial \alpha$  оценивалось только неявно через GAN-метрики. Более того, для этих трех решателей не производилась их количественная оценка при решении задач оптимального транспорта. Рассмотрим более подробно каждый из решателей из табл. 12.

Решатель [LS] оптимизирует регуляризованную двойственную форму для (5.3) без граничных условий [2]:

$$\max_{f,g} \left[ \int_{\mathbb{R}^D} f(x) d\mathbb{P}(x) + \int_{\mathbb{R}^D} g(y) d\mathbb{Q}(y) \right] - \mathcal{R}(f,g). \quad (5.7)$$

Решатель	Ссылки	Параметризация потенциалов или отображений	Количественно тестируется как метод ОТ.	Тестирование в GAN
Regularized [LS]	[2; 58; 128]	$f_\theta, g_\omega : \mathbb{R}^D \rightarrow \mathbb{R}$ - NNs	Гаусс. случай (глава 2)	Энтроп. регуляриз. WGAN [128]
Maximin [MM]	[135]	$f_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ - NN $H_\omega : \mathbb{R}^D \rightarrow \mathbb{R}^D$ - NN	-	WGAN [135] с тремя игроками
Maximin (Batch-wise) [MM-B]	[130; 136]	$f_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ - NN	-	(q,p)-WGAN [130]
Квадратичн. стоим. [QC]	[43]	$f_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ - NN	-	WGAN-QC [43]
Maximin + ICNN [MMv1]	[21]	$\psi_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ - ICNN	Гаусс. случай (глава 2)	-
Maximin + 2 ICNN [MMv2]	[7; 52]	$\psi_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ - ICNN $H_\omega : \mathbb{R}^D \rightarrow \mathbb{R}^D$ - VICNN	Гаусс. случай (глава 2)	-
Non-Maximin [W2]	Глава 2	$\psi_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ - ICNN $H_\omega : \mathbb{R}^D \rightarrow \mathbb{R}^D$ - VICNN	Гаусс. случай (глава 2)	-

Таблица 12 — Сравнение существующих параметрических двойственных решателей задач оптимального транспорта 1.

Энтропийный или квадратичный регуляризатор  $\mathcal{R}$  штрафует потенциалы  $f, g$  за нарушение ограничений  $f \oplus g \leq \frac{1}{2}\|\cdot\|^2$  [2, §3]. На практике  $f = f_\theta$  и  $g = g_\omega$  являются линейными комбинациями ядерных функций [58] или нейронных сетей [2]. Параметры  $\theta, \omega$  получаются применением метода стохастического градиентного подъема (SGA) по случайным выборкам из  $\mathbb{P}, \mathbb{Q}$ .

Большинство других решателей основано на разложении (5.4):

$$\mathbb{W}_2^2(\mathbb{P}, \mathbb{Q}) = \max_f \int_{\mathbb{R}^D} f(x) d\mathbb{P}(x) + \int_{\mathbb{R}^D} \underbrace{\min_{x \in \mathbb{R}^D} \left[ \frac{1}{2}\|x - y\|^2 - f(x) \right]}_{=f^c(y)} d\mathbb{Q}(y). \quad (5.8)$$

Сложностью (5.8) является внутренняя минимизация по  $x \in \mathbb{R}^D$ , т.е. вычисление  $f^c(y)$ . Основное отличие от существующих решателей заключается в процедуре решения внутренней задачи.

Решатель [MM-B] использует нейронную сеть  $f_\theta$  как потенциал, обученный SGA [130] на случайных выборках. Для решения внутренней задачи минимизация по  $x$  ограничивается текущей выборкой из  $\mathbb{P}$  (вместо  $\mathbb{R}^D$ ). Этот подход обеспечивает быстроту вычислений, но приводит к переоценке решения внутренней задачи, поскольку минимум берется по подмножеству.

В решателе [MM-v1] используется следующее свойство:  $f^* = \frac{1}{2}\|\cdot\|^2 - \Psi^*$ , где  $\Psi^*$  — выпуклая функция [21]. В [21] используется параметризация  $f_\theta = \frac{1}{2}\|\cdot\|^2 - \Psi_\theta$ ,

где  $\Psi_\theta$  – выпуклая по входу нейронная сеть (ICNN) [20]. Следовательно, для каждого  $y \in \mathbb{R}^D$  внутренняя подзадача в (5.8) становится выпуклой по  $x$ . Эта задача может быть решена с использованием SGA с высокой точностью, но этот подход очень затратен с вычислительной точки зрения (см. главу 2).

В решателе [MM] используется задача, эквивалентная (5.8) [135]:

$$\mathbb{W}_2^2(\mathbb{P}, \mathbb{Q}) = \max_f \int_{\mathbb{R}^D} f(x) d\mathbb{P}(x) + \int_{\mathbb{R}^D} \min_H \left[ \frac{1}{2} \|H(y) - y\|^2 - f(H(y)) \right] d\mathbb{Q}(y), \quad (5.9)$$

где минимизация производится по функциям  $H : \mathbb{R}^D \rightarrow \mathbb{R}^D$ . Используются нейронные сети  $f_\theta$  и  $H_\omega$  для параметризации потенциала и оптимизатора внутренней задачи. Для обучения  $\theta, \omega$  применяется метод стохастического градиентного подъема/спуска с использованием случайных выборок из  $\mathbb{P}, \mathbb{Q}$ . Метод [MM] является достаточно общим и может быть модифицирован для вычисления произвольных транспортных стоимостей и производных (т.е. он может применяться не только для расстояния  $\mathbb{W}_2^2$ ), хотя в указанной работе рассматривалось только расстояние Васерштейн-1 ( $\mathbb{W}_1$ ).

Аналогично [MMv1], метод [MMv2] параметризует  $f_\theta = \frac{1}{2} \|\cdot\|^2 - \Psi_\theta$ , где  $\Psi_\theta$  – сеть ICNN [7]. При фиксированном  $f_\theta$  оптимальное решение  $H$  имеет вид  $H = (\nabla \Psi_\theta)^{-1}$  (обратный градиент выпуклой функции) и, следовательно, также является градиентом выпуклой функции. С учетом этого в указанной работе параметризуется  $H_\omega = \nabla \varphi_\omega$ , где  $\varphi_\omega$  – сеть ICNN, а [MM] используется для обучения  $\theta, \omega$ .

В главе 2 данной работы мы предложили метод [W2]. В нем используется такая же параметризация ICNN, как и в [7], но, в отличие от этого подхода, вводится *циклически согласованная* регуляризация во избежание появления максиминной задачи.

В заключение этого краткого обзора упомянем также решатель [QC] из [43]. Как и в [MM-B], в нем нейронная сеть  $f_\theta$  используется, как потенциал. После получения случайных выборок  $\{x_n\}, \{y_n\}$  из  $\mathbb{P}, \mathbb{Q}$  в [43] решается *дискретная* задача оптимального транспорта для получения двойственных переменных  $\{f_n^*\}, \{g_n^*\}$ , которые, в свою очередь, используются для регрессии  $f_\theta(x_n)$  на  $f_n^*$ .

**Отклонение градиентов.** Указанные выше решатели применяются для оптимизации потенциалов типа  $f_\theta$  (или  $\Psi_\theta$ ), но именно градиент  $f_\theta$  (ор  $\Psi_\theta$ ) используется для восстановления отображения оптимального транспорта посредством  $T = x - \nabla f_\theta$ . Даже если норма  $\|f - f^*\|_{\mathcal{L}^2(\mathbb{P})}^2$  мала, отклонение  $\|\nabla f_\theta - \nabla f^*\|_{\mathcal{L}^2(\mathbb{P})}^2$  может быть сколь угодно велико, поскольку градиент  $\nabla f_\theta$

участвует в оптимизационном процессе неявно. Такую возникающую проблему мы будем называть проблемой *отклонения градиентов*. Для решателей [MMv1], [MMv2], [W2], основанных на сетях ICNN, это не является проблемой (см. теорему 3.6 в работе [7] и главу 2 диссертации).

**Реверсированные решатели.** В решателях [MM], [MMv2], [W2] восстанавливаются не только прямое отображение задачи оптимального транспорта  $\nabla\psi_\theta \approx \nabla\psi^* = T^*$ , но также и обратное отображение вида  $H_\omega \approx (T^*)^{-1} = (\nabla\psi^*)^{-1} = \nabla\bar{\Psi}^*$  (см. [7, §3] или главу 2). Данные решатели не являются симметричными по  $\mathbb{P}, \mathbb{Q}$ , поэтому возможным подходом здесь является перемена мест  $\mathbb{P}$  и  $\mathbb{Q}$  при обучении. Такие *реверсированные* решатели мы будем обозначать следующим образом: [MM:R], [MMv2:R], [W2:R]. В п. 5.5 мы покажем (и это неожиданно!), что решатель [MM:R] работает лучше в генеративном моделировании, чем решатель [MM].

## 5.4 Тестирование решателей для задачи оптимального транспорта

В этом параграфе мы предлагаем общий метод для создания *эталонных пар*, т.е. таких мер  $(\mathbb{P}, \mathbb{Q})$ , что  $\mathbb{Q} = T\sharp\mathbb{P}$  с выборочным доступом и аналитически известным решением задачи оптимального транспорта  $T^*$  между ними.

**Ключевая идея.** Наш метод основывается на том факте, что для дифференцируемой выпуклой функции  $\psi : \mathbb{R}^D \rightarrow \mathbb{R}$  ее градиент  $\nabla\psi$  является *оптимальным* транспортным отображением между произвольным  $\mathbb{P} \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$  и его образом  $\nabla\psi\sharp\mathbb{P}$  под действием  $\nabla\psi : \mathbb{R}^D \rightarrow \mathbb{R}^D$ . Это следует из теоремы Бренье [26] (теорема 2.12 в [25]). Таким образом, для непрерывной меры  $\mathbb{P}$  с выборочным доступом и *известным* выпуклым  $\psi$  в качестве эталонной пары можно использовать  $(\mathbb{P}, \nabla\psi\sharp\mathbb{P})$ . Выборка из  $\nabla\psi\sharp\mathbb{P}$  осуществляется путем взятия выборки из  $\mathbb{P}$  и применения  $\nabla\psi$ .

**Произвольные пары  $(\mathbb{P}, \mathbb{Q})$ .** Для произвольной пары  $(\mathbb{P}, \mathbb{Q})$  вычисление точного непрерывного решения задачи оптимального транспорта является трудной задачей. В качестве компромисса мы вычисляем *приближение* транспортного отображения как градиент сети ICNN с использованием [W2]. На этом пути мы находим функцию  $\psi_\theta$  с параметризацией, как у ICNN, и такую, что  $\nabla\psi_\theta\sharp\mathbb{P} \approx \mathbb{Q}$ . Тогда модифицированная пара  $(\mathbb{P}, \nabla\psi_\theta\sharp\mathbb{P})$  может быть использова-

на для тестирования решателя задачи оптимального транспорта. Здесь выбор [W2] обеспечивает хорошее функционирование в больших размерностях; при этом остальные решатели также могут быть использованы при условии, что  $\psi_\theta$  выпукла. При таком выборе [W2] последующие вычисления могут оказаться лучше для метода, основанного на сети ICNN.

**Расширения.** Выпуклые функции можно модифицировать для генерации большего числа эталонных пар. Если  $\psi_1, \dots, \psi_N$  выпуклы, то  $\sigma(\psi_1, \dots, \psi_N)$  также выпукла при условии, что  $\sigma : \mathbb{R}^N \rightarrow \mathbb{R}$  выпукла и монотонна. Например,  $c \cdot \psi_1$  ( $c \geq 0$ ),  $\sum_n \psi_n$ ,  $\max_n \psi_n$  выпуклы, поэтому по их градиентам можно получить новые эталонные пары.

**Обращение.** Если функция  $\nabla\psi_\theta$  биективна, то обратное транспортное отображение для  $(\mathbb{P}, \nabla\psi_\theta \sharp \mathbb{P})$  существует и имеет вид  $(\nabla\psi_\theta)^{-1}$ . Для каждого  $y \in \mathbb{R}^D$  значение  $(\nabla\psi_\theta)^{-1}(y)$  можно получить, решая *выпуклую* задачу [21, §6]. Все используемые нами сети ICNN  $\psi_\theta$  имеют биективные градиенты  $\nabla\psi_\theta$  (см. п. 5.7.1).

## 5.5 Детали сравнительных тестов и полученные результаты

Тестирование проводилось на фреймворке PyTorch. Программный код доступен публично по адресу

<https://github.com/iamalexkorotin/Wasserstein2Benchmark>

Наши эксперименты были выполнены на четырех видеокартах GTX 1080ti и заняли около 100 ч. (на один графический процессор). Подробности реализации даются в п. 5.7.

### 5.5.1 Наборы данных

**Меры большой размерности.** Были разработаны эталонные пары для проверки того, могут ли решатели задачи оптимального транспорта правильно транспортировать массы между мерами. С этой целью мы использовали смеси

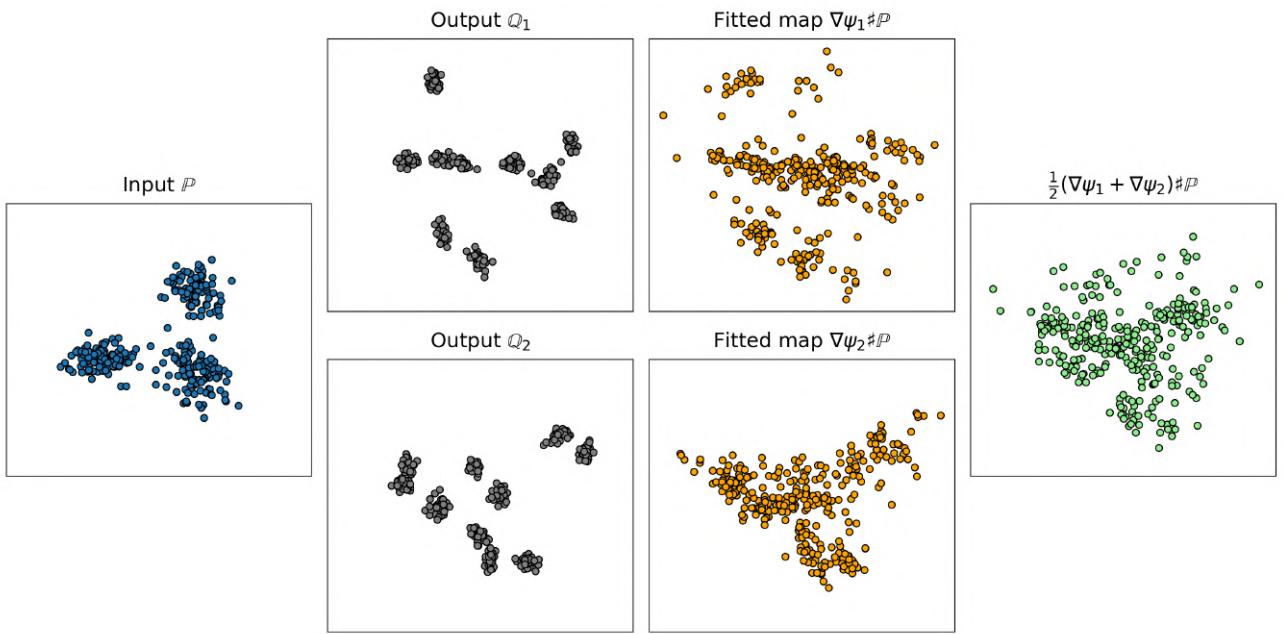


Рисунок 5.1 — Пример создания эталонной пары для размерности  $D = 16$ . Сначала инициализируются 3 случайные смеси гауссовых распределений  $\mathbb{P}$  и  $\mathbb{Q}_1, \mathbb{Q}_2$  и обучаются два приближенные отображения задачи оптимального транспорта  $\nabla\psi_i \# \mathbb{P} \approx \mathbb{Q}_i$ ,  $i = 1, 2$ . Усреднение потенциалов используется для определения выходной меры:  $\frac{1}{2}(\nabla\psi_1 + \nabla\psi_2) \# \mathbb{P}$ . Каждый рисунок содержит 512 случайных точек, спроектированных на две принципиальные компоненты меры

$$\frac{1}{2}(\nabla\psi_1 + \nabla\psi_2) \# \mathbb{P}.$$

гауссовых распределений для размерностей  $D = 2^1, 2^2, \dots, 2^8$ . В каждой размерности  $D$  мы рассмотрели случайную смесь  $\mathbb{P}$  из 3 гауссовых распределений и двух случайных смесей  $\mathbb{Q}_1, \mathbb{Q}_2$  из 10 гауссовых распределений. Приближенное транспортное отображение  $\nabla\psi_i \# \mathbb{P} \approx \mathbb{Q}_i$  ( $i = 1, 2$ ) было обучено с использованием решателя [W2]. Каждый потенциал являлся сетью ICNN с архитектурой DenseICNN, см. п. 2.7. Тестовая пара определялась через полусумму вычисленных потенциалов  $(\mathbb{P}, \frac{1}{2}(\nabla\psi_1 + \nabla\psi_2) \# \mathbb{P})$ . Первая мера  $\mathbb{P}$  является смесью 3 гауссовых распределений, а вторая получена путем усреднения потенциалов, которые преобразуют ее в приближенные смеси из 10 гауссовых распределений (по поводу дальнейших деталей см. п. 5.6.1 и рис. 5.1).

**Изображения.** Мы используем выровненные изображения из набора данных CelebA64 (лица)<sup>1</sup> [41] для генерации дополнительных эталонных пар. Для начала мы обучаем 3 генеративные модели (WGAN-QC [43]) на наборе данных и выбираем промежуточные контрольные точки для получения непре-

<sup>1</sup><http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html>

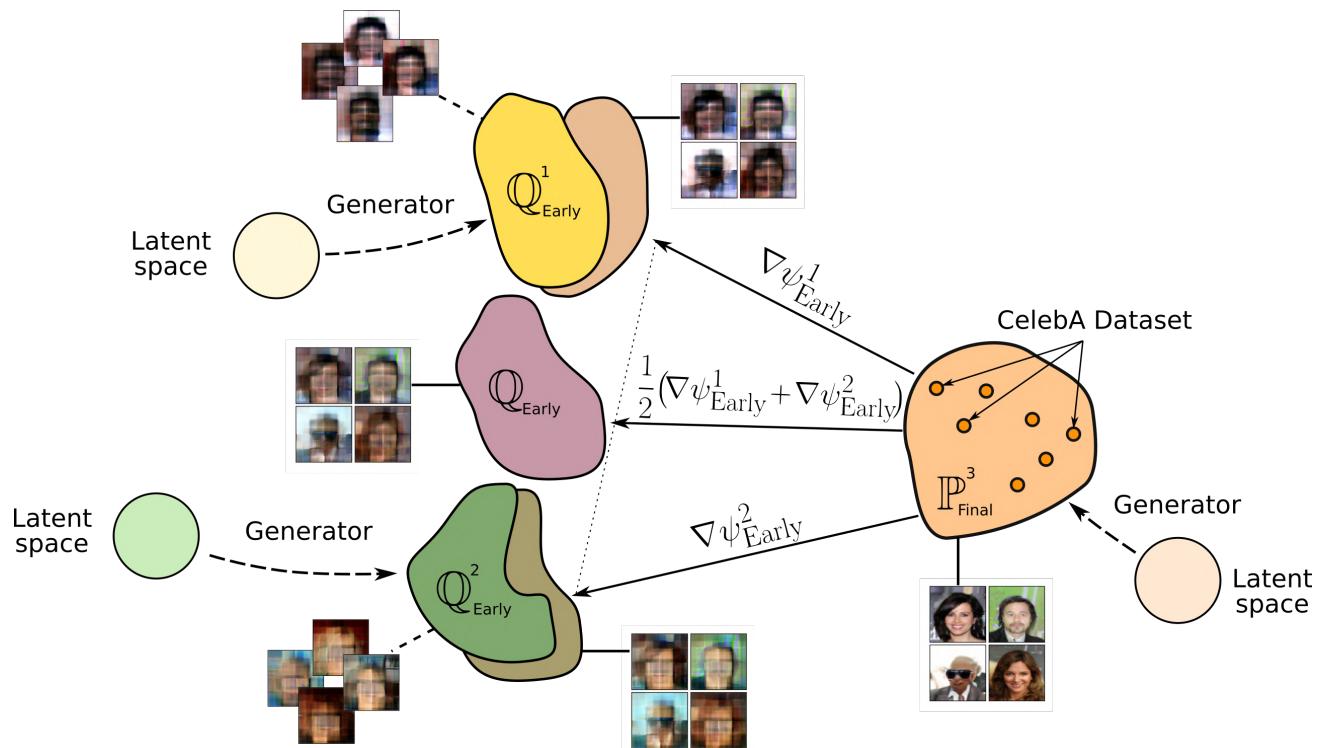


Рисунок 5.2 — Методология создания эталонных пар на пространстве изображений. Мы используем 3 генеративных модели:  $\mathbb{P}_{\text{Final}}^3$  (хорошо обученная) и  $\mathbb{Q}_{\text{Cpkt}}^1$ ,  $\mathbb{Q}_{\text{Cpkt}}^2$  (недообученные). Для  $k = 1, 2$  мы обучаем приближенное ОТ отображение  $\mathbb{P}_{\text{Final}}^3 \rightarrow \mathbb{Q}_{\text{Cpkt}}^k$  как  $\nabla \psi_{\text{Cpkt}}^k$ , т.е. градиент сети ICNN. Определим эталонную пару через  $(\mathbb{P}_{\text{CelebA}}, \mathbb{Q}_{\text{Cpkt}}) \stackrel{\text{def}}{=} (\mathbb{P}_{\text{Final}}^3, \frac{1}{2}(\nabla \psi_{\text{Cpkt}}^1 + \nabla \psi_{\text{Cpkt}}^2) \# \mathbb{P}_{\text{Final}}^3)$ . На визуализации представлена эталонная пара Early.

рывных мер  $\mathbb{Q}_{\text{Early}}^k, \mathbb{Q}_{\text{Mid}}^k, \mathbb{Q}_{\text{Late}}^k$  для первых двух моделей ( $k = 1, 2$ ) и финальную контрольную точку третьей модели ( $k = 3$ ) для получения меры  $\mathbb{P}_{\text{Final}}^3$ . Для обеспечения абсолютной непрерывности мер мы добавляем небольшой гауссовский шум на выход генератора. Каждая контрольная точка (Early, Mid, Late, Final) представляет изображения лиц определенного качества. Далее, для  $k \in \{1, 2\}$  и точки  $\text{Cpkt} \in \{\text{Early}, \text{Mid}, \text{Late}\}$  мы используем решатель [W2] для обучения приближенного транспортного отображения  $\nabla \psi_{\text{Cpkt}}^k$  для пары  $(\mathbb{P}_{\text{Final}}^3, \mathbb{Q}_{\text{Cpkt}}^k)$ , т.е.  $\nabla \psi_{\text{Cpkt}}^k \# \mathbb{P}_{\text{Final}}^3 \approx \mathbb{Q}_{\text{Cpkt}}^k$ . Потенциал  $\psi_{\text{Cpkt}}^k$  есть сверточная сеть ICNN с архитектурой ConvICNN64 (п. 5.7.1). Для каждой точки Cpkt мы получаем эталонную пару  $(\mathbb{P}_{\text{CelebA}}, \mathbb{Q}_{\text{Cpkt}}) \stackrel{\text{def}}{=} (\mathbb{P}_{\text{Final}}^3, [(\nabla \psi_{\text{Cpkt}}^1 + \nabla \psi_{\text{Cpkt}}^2)/2] \# \mathbb{P}_{\text{Final}}^3)$  (по поводу деталей см. п. 5.6.2 и рис. 5.2).

### 5.5.2 Метрики и базовые решатели

**Базовые решатели (бейзлайны).** Мы предлагаем следующие бейзлайны: тождественный (identity)  $[ID]$ , постоянный (constant)  $[C]$  и линейный (linear)  $[L]$ . *Тождественный* решатель в качестве транспортного отображения выдает  $T^{\text{id}} = \text{id}_{\mathbb{R}^D}$ . *Постоянный* решатель выдает среднее значение  $\mathbb{Q}$ , т.е.  $T^0 \equiv \mathbb{E}_{\mathbb{Q}}[y] \equiv \mu_{\mathbb{Q}}$ . *Линейный* решатель выдает  $T^1(x) = \Sigma_{\mathbb{P}}^{-\frac{1}{2}} (\Sigma_{\mathbb{P}}^{\frac{1}{2}} \Sigma_{\mathbb{Q}} \Sigma_{\mathbb{P}}^{\frac{1}{2}})^{\frac{1}{2}} \Sigma_{\mathbb{P}}^{-\frac{1}{2}} (x - \mu_{\mathbb{P}}) + \mu_{\mathbb{Q}}$ , т.е. отображение задачи оптимального транспорта между мерами, укрупненными до гауссовых распределений, см. теорему 2.3 [36].

**Метрики.** Для оценки качества восстановленного транспортного отображения  $\hat{T} : \mathbb{R}^D \rightarrow \mathbb{R}^D$  из  $\mathbb{P}$  в  $\mathbb{Q}$  мы используем *процент необъясненной дисперсии* (unexplained variance percentage; UVP), см. главу 2:  $\mathcal{L}^2\text{-UVP}(\hat{T}) \stackrel{\text{def}}{=} 100 \cdot \|\hat{T} - T^*\|_{\mathcal{L}^2(\mathbb{P})}^2 / \text{Var}(\mathbb{Q})\%$ . Здесь  $T^*$  – эталонное отображение задачи оптимального транспорта. Для значений  $\approx 0\%$  отображение  $\hat{T}$  хорошо аппроксимирует  $T^*$ . Для значений  $\geq 100\%$  отображение  $\hat{T}$  далеко от оптимального. Постоянный решатель дает  $\mathcal{L}^2\text{-UVP}(T^0) = 100\%$ .

Для оценки качества аппроксимации производной потенциала  $[\text{id}_{\mathbb{R}^D} - \hat{T}] \approx \nabla f^*$ , применяемого для обучения генеративных моделей (5.6), мы используем косинусную близость ( $\cos$ ):

$$\cos(\text{id} - \hat{T}, \text{id} - T^*) \stackrel{\text{def}}{=} \frac{\langle \hat{T} - \text{id}, \nabla \psi^* - \text{id} \rangle_{\mathcal{L}^2(\mathbb{P})}}{\|T^* - \text{id}\|_{\mathcal{L}^2(\mathbb{P})} \cdot \|\hat{T} - \text{id}\|_{\mathcal{L}^2(\mathbb{P})}} \in [-1, 1].$$

Для оценки  $\mathcal{L}^2\text{-UVP}$  и метрик  $\cos$  используются выборки из  $\mathbb{P}$  размера  $2^{14}$ .

### 5.5.3 Оценка решателей на многомерных эталонных парах

Мы обучаем решатели на разработанных эталонных парах и получаем вычисленные значения метрик для обученного транспортного отображения. Для объективного сравнения в каждом из методов потенциал  $f$  и отображение  $H$  (в применимых случаях) параметризуются как  $f_{\theta} = \frac{1}{2} \|\cdot\|^2 - \psi_{\theta}$  и  $H_{\omega} = \nabla \varphi_{\omega}$ , соответственно, где  $\psi_{\theta}, \varphi_{\omega}$  использует архитектуру DenseICNN (см. п. 2.7). В решателях  $[QC]$ ,  $[LS]$ ,  $[MM-B]$ ,  $[MM]$  мы не налагаем никаких ограничений на

веса  $\theta, \omega$ , т.е.  $\psi_{\theta, \omega}$  являются обычными полносвязными сетями с дополнительными обходными связями. Вычисленные значения для метрик приведены в табл. 13, графически (для  $D = 64$ ) результаты представлены на рис. 5.3.

В размерности  $D = 2$  все решатели показали хорошие результаты ( $\mathcal{L}^2\text{-UVP} \approx 0$ ,  $\cos \approx 1$ ). Однако для больших размерностей приемлемые результаты дали только  $\lfloor \text{MMv1} \rfloor$ ,  $\lfloor \text{MM} \rfloor$ ,  $\lfloor \text{MMv2} \rfloor$ ,  $\lfloor \text{W2} \rfloor$  и их реверсированные варианты. При этом решатель  $\lfloor \text{MMv1} \rfloor$  работает медленно, так как на каждом оптимизационном шаге решается сложная подзадача вычисления  $f^c$ . Максиминные решатели  $\lfloor \text{MM} \rfloor, \lfloor \text{MMv2} \rfloor, \lfloor \text{MM:R} \rfloor$  также трудно оптимизируются, поскольку они или расходятся с самого начала ( $\rightarrow$ ), или расходятся после схождения к почти оптимальной седловой точке ( $\leftarrow$ ). По-видимому, такое поведение, типичное для задач оптимизации на максимум, можно улучшить при более тщательном выборе гиперпараметров.

Для решателей  $\lfloor \text{QC} \rfloor$ ,  $\lfloor \text{LS} \rfloor$ ,  $\lfloor \text{MM-B} \rfloor$  при возрастании размерности величина  $\mathcal{L}^2\text{-UVP}$  существенно возрастает. В сравнении с тривиальным решателем  $\lfloor \text{L} \rfloor$  значительно лучшие результаты показывает только  $\lfloor \text{MM-B} \rfloor$ . Ошибки при работе решателя  $\lfloor \text{MM-B} \rfloor$  могут быть объяснены переоценкой решения внутренней задачи в (5.8), что приводит к смещенным оптимальным потенциалам. Ошибки в решателе  $\lfloor \text{LS} \rfloor$  могут быть объяснены смещением из-за регуляризации [2]. В решателе  $\lfloor \text{QC} \rfloor$  ошибки возникают вследствие того, что дискретная задача оптимального транспорта (эта задача обычно приводит к смещению, см. теорему 1 в [137]), решаемая на случайных выборках, используется для обновления  $f_\theta$ . Стоит отметить, что хотя решатели  $\lfloor \text{QC} \rfloor$ ,  $\lfloor \text{LS} \rfloor$  не являются точными в терминах показателя  $\mathcal{L}^2\text{-UVP}$ , оказывается, что они дают высокие значения для метрики  $\cos$ .

Из-за сложностей, связанных с оптимизацией и различием в функционировании, физическое время сходимости не является репрезентативным. Все решатели, за исключением  $\lfloor \text{MMv1} \rfloor$ , сходятся за несколько часов. Для решателей, существенно превосходящих линейный (т.е. для решателей  $\lfloor \text{MM} \rfloor$ ,  $\lfloor \text{MMv1} \rfloor$ ,  $\lfloor \text{MMv2} \rfloor$ ,  $\lfloor \text{W2} \rfloor$ ,  $\lfloor \text{MM-B} \rfloor$ ) наибольшую скорость сходимости показал  $\lfloor \text{MM-B} \rfloor$ , но для него результаты будут смещены. Решатели  $\lfloor \text{MM} \rfloor$ ,  $\lfloor \text{MMv2} \rfloor$ ,  $\lfloor \text{W2} \rfloor$  требуют большего времени.

Dim	<b>2</b>	<b>4</b>	<b>8</b>	<b>16</b>	<b>32</b>	<b>64</b>	<b>128</b>	<b>256</b>		Dim	<b>2</b>	<b>4</b>	<b>8</b>	<b>16</b>	<b>32</b>	<b>64</b>	<b>128</b>	<b>256</b>
[MMv1]	0.2	1.0	1.8	1.4	6.9	8.1	2.2	2.6		[MMv1]	0.99	0.99	0.99	0.99	0.98	0.97	0.99	0.99
[MM]	0.1	0.3	0.9	2.2	4.2	3.2	3.1↑	4.1↑		[MM]	0.99	0.99	0.99	0.99	0.99	0.99	0.999↑	0.999↑
[MM:R]	0.1	0.3	0.7	1.9	2.8	4.5	→	→		[MM:R]	0.99	1.00	1.00	0.99	1.00	0.98	→	→
[MMv2]	0.1	0.68	2.2	3.1	5.3	10.1↑	3.2↑	7.7↑		[MMv2]	0.99	0.99	0.99	0.99	0.99	0.96↑	0.999↑	0.999↑
[MMv2:R]	0.1	0.7	4.4	7.7	5.8	6.8	2.1	2.8		[MMv2:R]	0.99	1.00	0.97	0.96	0.99	0.97	0.99	1.00
[W2]	0.1	0.7	2.6	3.3	6.0	7.2	2.0	2.7		[W2]	0.99	0.99	0.99	0.99	0.99	0.97	1.00	1.00
[W2:R]	0.2	0.9	4.0	5.3	5.2	7.0	2.0	2.7		[W2:R]	0.99	1.00	0.98	0.98	0.99	0.97	1.00	1.00
[MM-B]	0.1	0.7	3.1	6.4	12.0	13.9	19.0	22.5		[MM-B]	0.99	1.00	0.98	0.96	0.96	0.94	0.93	0.93
[LS]	5.0	11.6	21.5	31.7	42.1	40.1	46.8	54.7		[LS]	0.94	0.86	0.80	0.80	0.81	0.83	0.82	0.81
[L]	14.1	14.9	27.3	41.6	55.3	63.9	63.6	67.4		[L]	0.75	0.80	0.73	0.73	0.76	0.75	0.77	0.77
[QC]	1.5	14.5	28.6	47.2	64.0	75.2	80.5	88.2		[QC]	0.99	0.84	0.78	0.70	0.70	0.70	0.69	0.66
[C]	100	100	100	100	100	100	100	100		[C]	0.29	0.32	0.38	0.46	0.55	0.58	0.60	0.62
[ID]	32.7	42.0	58.6	87	121	137	145	153		[ID]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Таблица 13 — Значения метрик  $\mathcal{L}^2$ -UVP (%) (слева) и  $\cos \in [-1,1]$  (справа) для транспортных отображений, обученных при помощи решателей задачи оптимального транспорта многомерных на эталонных парах в размерностях  $D = 2, 2^2, \dots, 2^8$ . Оранжевый цвет —  $\mathcal{L}^2$ -UVP  $> 10\%$  и  $\cos < 0.95$ . Красный цвет — функционирование хуже, чем у бейзлайна [L].

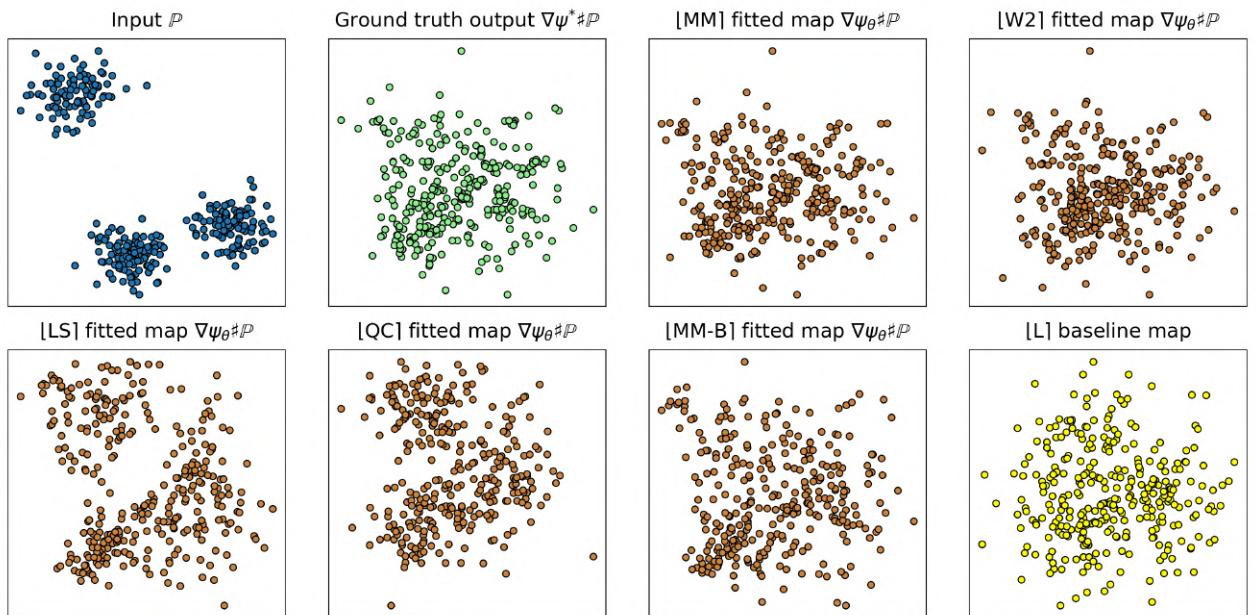
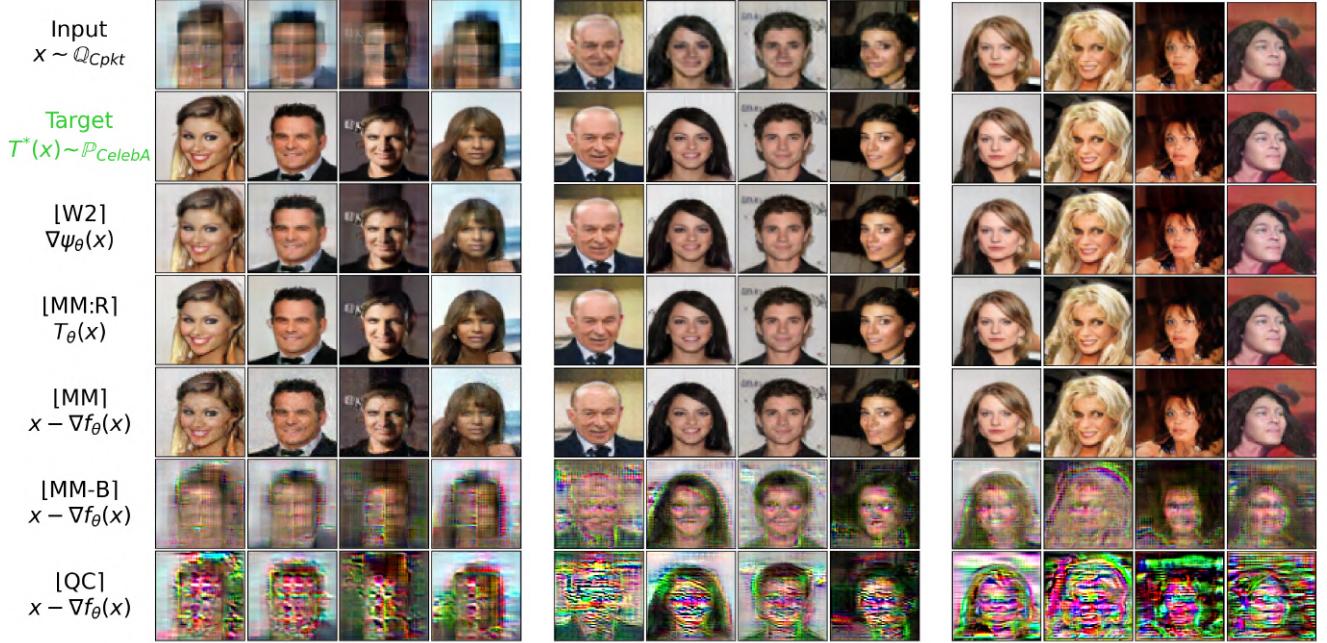


Рисунок 5.3 — Визуализация 64-мерной эталонной пары и ОТ отображений обученных решателями. Диаграммы рассеяния содержат 512 случайных выборок, спроектированных на две главные компоненты меры  $\nabla\psi^*\sharp\mathbb{P}$ .

#### 5.5.4 Оценка решателей на эталонных парах изображений размера $64 \times 64$ из набора CelebA

При проведении сравнительных оценок на наборе CelebA мы не рассматриваем решатели [LS] и [MMv1], поскольку первый является неустойчивым в больших размерностях [128], а второй обладает слишком медленной скоростью сходимости. Основанные на сетях ICNN решатели [MMv2], [W2] и их ревер-



а) Обученные

отображения

$$\mathbb{Q}_{\text{Early}} \rightarrow \mathbb{P}_{\text{CelebA}}.$$

б) Обученные

отображения

$$\mathbb{Q}_{\text{Mid}} \rightarrow \mathbb{P}_{\text{CelebA}}.$$

в) Обученные

отображения

$$\mathbb{Q}_{\text{Late}} \rightarrow \mathbb{P}_{\text{CelebA}}.$$

Рисунок 5.4 — Отображения, обученные решателями на эталонных парах  $(\mathbb{Q}_{\text{Cpkt}}, \mathbb{P}_{\text{CelebA}})$ . В первой строке даны случайные  $x \sim \mathbb{Q}_{\text{Cpkt}}$ . Вторая строка содержит выборки из  $\mathbb{P}_{\text{CelebA}}$ , полученные прямым отображением  $x$  посредством ОТ отображения  $T^* = \nabla\psi^*$ . Далее показаны образы  $x$  под действием отображений, обученных нашими решателями задачи оптимального транспорта.

Cpkt	Early	Mid	Late	Cpkt	Early	Mid	Late
[W2]	1.7	0.5	0.25	[W2]	<b>0.99</b>	0.95	0.93
[MM]	2.2	0.9	0.53	[MM]	0.98	0.90	0.87
[MM:R]	<b>1.4</b>	<b>0.4</b>	<b>0.22</b>	[MM:R]	<b>0.99</b>	<b>0.96</b>	<b>0.94</b>
[ID]	31.2	4.26	2.06	[ID]	0.00	0.00	0.00
[MM-B]	<b>45.9</b>	<b>46.1</b>	<b>47.74</b>	[MM-B]	<b>0.28</b>	<b>-0.08</b>	<b>-0.14</b>
[C]	100	100	100	[C]	0.03	-0.14	-0.20
[QC]	<b>94.7</b>	<b>&gt;&gt;100</b>	<b>&gt;&gt;100</b>	[QC]	<b>0.17</b>	<b>-0.01</b>	<b>0.05</b>

Таблица 14 — Значения метрик  $\mathcal{L}^2$ -UVP (%) (слева) и  $\cos \in [-1,1]$  (справа) для транспортных отображений  $\mathbb{Q}_{\text{Cpkt}} \rightarrow \mathbb{P}_{\text{CelebA}}$ , обученных решателями оптимального транспорта на 3 разработанных CelebA64  $\mathbb{W}_2$ -эталонных наборах.

сированные версии показали практически равные результаты. Для простоты мы их будем просто называть решателем [W2].

В решателе [W2] мы параметризуем  $f_\theta = \frac{1}{2}\|\cdot\|^2 - \Psi_\theta$  и  $H_\omega = \nabla\varphi_\omega$ , где  $\Psi_\theta, \varphi_\omega$  — выпуклые по входу нейронные сети с выпуклой архитектурой ConvexICNN64 (см. п. 5.7.1). Все другие решатели основаны на генеративном

моделировании с применением сверточных архитектур для изображений. Таким образом, в [MM], [QC], [MM-B] мы параметризуем сети  $f_\theta$  как ResNet и  $H_\omega$  как U-Net (in [MM]). В свою очередь, в [MM:R] сеть  $T_\theta$  – UNet, а  $g_\omega$  – ResNet.

Транспортное отображение  $\mathbb{Q}_{\text{Cpkt}} \rightarrow \mathbb{P}_{\text{CelebA}}$  вычисляется для каждого решателя на трех эталонных наборах изображений. Полученные результаты показаны на рис. 5.4 и в табл. 14; они приблизительно аналогичны результатам, полученным на задачах большой размерности (см. п. 5.5.3). Недостатком решателей [QC], [MM-B] является наличие сильной смещенности из-за размерности изображений, при этом вычисленные этими решателями производные  $\mathbb{W}_2^2$  почти ортогональны истинной производной ( $\cos \approx 0$ ). Это означает, что *такие решатели не точно вычисляют  $\mathbb{W}_2^2$ .* При этом решатели [MM], [MM:R], [W2] хорошо восстанавливают транспортное отображение. Отображение решателя [MM] немного более зашумлено, чем отображение решателя [MM:R] (это представляет собой незначительный пример отклонения градиентов).

### 5.5.5 Тестирование решателей в генеративном моделировании изображений размера $64 \times 64$

Предыдущие вычисления показывают, что много существующих решателей задачи оптимального транспорта очень неточны. Поэтому возникает вопрос: *до какой степени качество решателя имеет значение в практических приложениях?*

Для ответа на этот вопрос мы оценим наиболее перспективные решатели на задаче генеративного моделирования изображений лиц размера  $64 \times 64$  из набора данных CelebA. Для сравнения мы также рассмотрим решатель [QC], обладающий хорошими генеративными характеристиками [43]. Для каждого решателя мы обучаем генеративную сеть  $G_\alpha$  с архитектурой ResNet из [43] для отображения 128-мерного нормального распределения  $\mathbb{S}$  в распределение данных  $\mathbb{Q}$ . В качестве функции потерь для генератора мы используем  $\mathbb{W}_2^2(\mathbb{P}_\alpha, \mathbb{Q}) = \mathbb{W}_2^2(G_\alpha \# \mathbb{S}, \mathbb{Q})$ , которое оценивается решателями. Обучение производилось по типу GAN, где уточнения генератора перемежаются с градиентными шагами решателя задачи оптимального транспорта (дискrimинатора). Выбранные сгенерированные изображения показаны в первой строке каждого

подрисунка на рис. 5.5; выдаются полученные показатели FID (см. [42]). В последней строке показаны образы под действием отображения задачи оптимального транспорта из  $\mathbb{P}_\alpha = G_\alpha \sharp \mathbb{S}$  в  $\mathbb{Q}$ , извлеченные из решателя задачи оптимального транспорта. Так как обучение сошлось ( $\mathbb{P}_\alpha \approx \mathbb{Q}$ ), отображение должно почти совпадать с тождественным.

Решатель [W2] не дает достаточного качества в этой задаче (рис. 5.5а). Это может быть объяснено использованием сети ConvICNN (в других решателях используется обычная сверточная архитектура, и они работают лучше). В общем, для применения сетей ICNN в задачах, связанных с изображениями (см. главу 2), требуется разработка новых выпуклых сверточных архитектур, что является перспективным направлением будущих исследований.

Генеративные свойства решателя [QC] хороши (рис. 5.5б). Однако, как и в п. 5.5.3-5.5.4, восстановленное отображение далеко от оригинала. В этой связи мы предполагаем, что решатель имеет хорошие генеративные свойства, поскольку он аппроксимирует некоторую другую (не- $\mathbb{W}_2^2$ ) метрику на практике.

Решатель [MM] дает генеративную модель с размытыми изображениями (рис. 5.5в). Вычисленное транспортное отображение  $\text{id}_{\mathbb{R}^D} - \nabla f_\theta$  существенно отличается от оригинала из-за отклонений градиента. Это приводит к неточностям в вычислении градиента, используемого для обновления генератора, а также объясняет, почему генератор не улучшается. Стоит отметить, что результаты применения этого решателя в п. 5.5.4 не выявили ощутимых недостатков в смысле отклонения градиента. Возможно, это связано с тем, что меры абсолютно непрерывны и имеют носитель на всем  $\mathbb{R}^D$ . Однако эта ситуация отлична от рассматриваемой нами в нашем генеративном моделировании, когда сгенерированные меры и меры данных имеют носители на маломерных многообразиях в  $\mathbb{R}^D$ .

Реверсированный решатель [MM:R] не имеет недостатков, присущих из-за отклонения градиента решателю [MM], но всё равно приводит к появлению размытых изображений (рис. 5.5г). Достаточно интересно, что обученное транспортное отображение  $T_\theta$  значительно улучшает качество; при этом изображения  $T_\theta(G_\alpha(z))$  сравнимы с изображениями, полученными с помощью [QC] (рис. 5.5б).

Следует отметить, что работа решателей [MM], [MM:R] основана на решении максиминной задачи, поэтому их использование в сетях GAN требует решения сложной *min-max-min* оптимизационной задачи. Мы используем

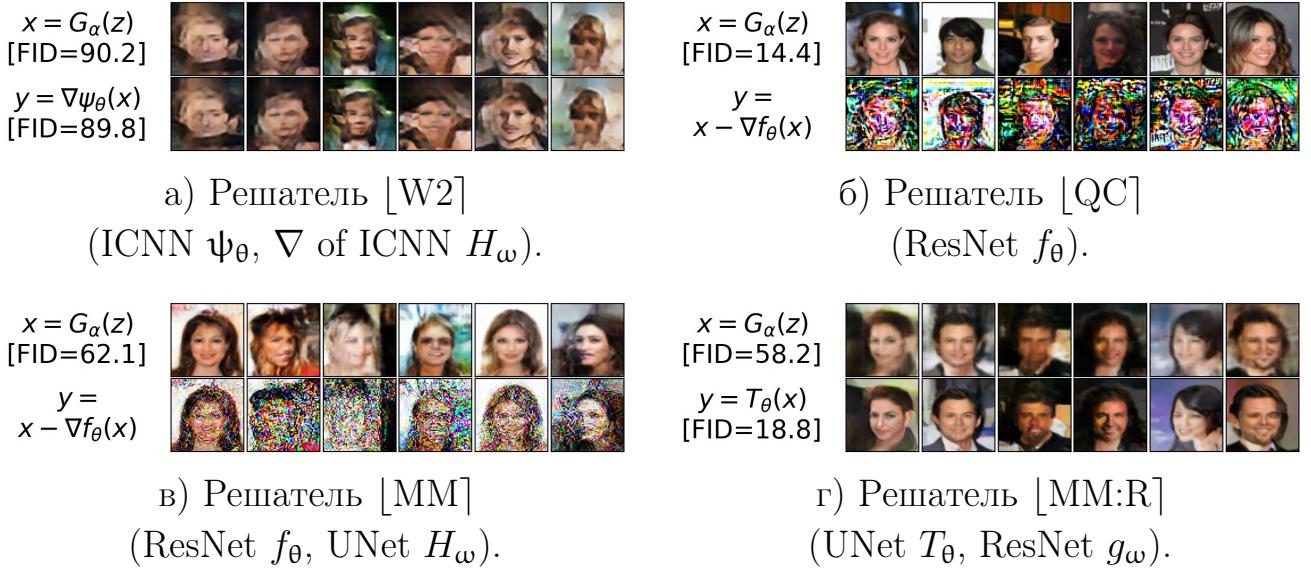
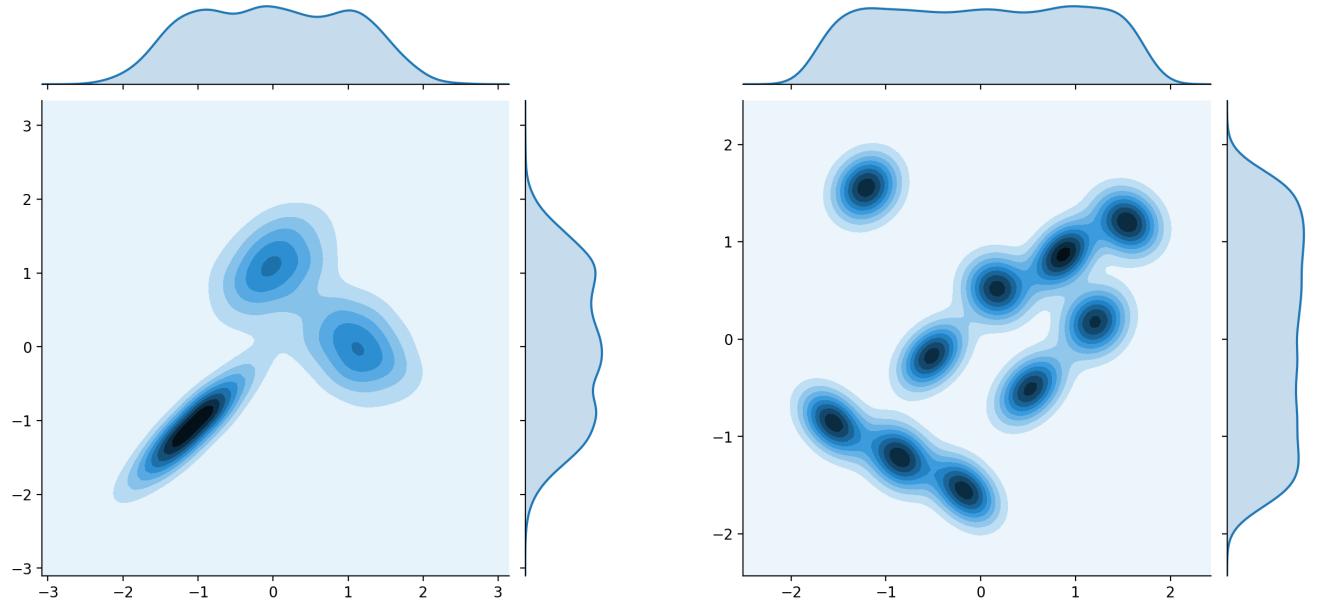


Рисунок 5.5 — Случайные изображения, полученные в результате обучения генеративной модели с помощью решателя. В первой строке показаны случайные сгенерированные изображения  $x = G_\alpha(z) \sim \mathbb{P}_\alpha$ ,  $z \sim \mathbb{S}$ . Во второй строке показано вычислённое транспортное отображение из сгенерированного  $x = G_\alpha(z) \sim \mathbb{P}_\alpha$  в  $\mathbb{Q}$ .

три вложенных цикла обучения и стохастический метод градиентного спуска-подъема-спуска. В наших экспериментах обучение было нестабильным и часто расходилось: в приведенных результатах использованы наилучшие найденные нами гиперпараметры (хотя, возможно, существуют и лучшие). Соответственно, имеющиеся сложности в выборе гиперпараметров и нестабильность процесса обучения составляют проблему для таких решателей, с которой следует разобраться до начала их практического использования.

## 5.6 Технические детали построения эталонных пар

В п. 5.6.1 обсуждаются детали эталонных пар большой размерности. В п. 5.6.2 изучаются изображения размера  $64 \times 64$  эталонных пар из набора данных Celeba.



a) Случайная смесь из 3 гауссовских распределений.

б) Случайная смесь из 10 гауссовских распределений.

Рисунок 5.6 — Случайно сгенерированные смеси гауссовских распределений. Проекция на первые две размерности.

### 5.6.1 Эталонные пары большой размерности

Пример эталонной пары дается на рис. 5.1. В каждой размерности мы фиксируем случайные смеси гауссовских распределений  $\mathbb{P}, \mathbb{Q}_1, \mathbb{Q}_2$  (в программном коде мы определяем конкретные смеси и используем их для построения эталонных пар).

Для генерации случайной смеси из  $M$  гауссовских мер в размерности  $D$  мы используем следующую процедуру. Пусть  $\delta, \sigma > 0$  (мы используем  $\delta = 1$ ,  $\sigma = \frac{2}{5}$ ). Рассмотрим  $M$ -мерную сетку

$$G = \left\{ -\frac{\delta \cdot M}{2} + i \cdot \delta \text{ для } i = 1, 2, \dots, M \right\}^D \subset \mathbb{R}^D.$$

Мы выбираем  $M$  случайных точек  $\mu'_1, \dots, \mu'_M \in G$  так, что никакая пара точек не имеет одинаковых координат. Далее инициализируются случайные матрицы  $A'_1, \dots, A'_M \in \mathbb{R}^{D \times D}$ , где каждая строка любой матрицы случайно выбирается из  $D - 1$ -мерной сферы в  $\mathbb{R}^D$ . Пусть  $\Sigma'_m = \sigma^2 \cdot (A'_m) \cdot (A'_m)^\top$  для  $m = 1, 2, \dots, M$ . Отметим, что  $[\Sigma'_m]_{dd} = \sigma^2$  при  $d = 1, 2, \dots, D$ . Далее, мы рассматриваем смеси гауссовских распределений  $\frac{1}{M} \sum_{m=1}^M \mathcal{N}(\mu'_m, \Sigma'_m)$ . Окончательно, мы нормализуем смесь так, чтобы покоординатная дисперсия была равна 1, т.е.

мы рассматриваем окончательную смесь  $\frac{1}{M} \sum_{m=1}^M \mathcal{N}(\mu_m, \Sigma_m)$ , где  $\mu_m = a\mu'_m$  и  $\Sigma_m = a^2\Sigma_m$ . Значение  $a \in \mathbb{R}_+$  находится из равенства

$$a^{-1} = \sqrt{\frac{\sum_{m=1}^M \|\mu'_m\|^2}{M \cdot D} + \sigma^2}.$$

Смеси гауссовых распределений, получаемые по этой процедуре, имеют  $D$  равных маргиналов (см. рис. 5.6).

### 5.6.2 Эталонные пары изображений

Мы обучаем три генеративные модели на наборе данных CelebA64 (выровненные лица) с 128-мерной латентной гауссовой мерой. Выборка из распределения осуществляется с помощью WGAN-QC [43] с сетью генератора ResNet. В испытаниях  $k = 1, 2$  мы сохраняем контрольные точки генератора после 1000, 5000, 10000 итераций для получения мер  $\mathbb{Q}_{\text{Early}}^k$ ,  $\mathbb{Q}_{\text{Mid}}^k$ ,  $\mathbb{Q}_{\text{Late}}^k$  соответственно. В последнем испытании  $k = 3$  мы сохраняем только контрольные точки финальной сети генератора после 50000 итераций, в результате чего получаем меру  $\mathbb{P}_{\text{Final}}^3$ . Для обеспечения абсолютной непрерывности мер мы добавляем нормальный белый шум (коэффициент по осям  $\sigma = 0.01$ ) к выходам генераторов.

Мы используем сгенерированные меры для построения изображения эталонных пар согласно подходу из п. 5.5.1. Графическая схема представлена на рис. 5.2.

## 5.7 Технические детали экспериментов

В п. 5.7.1 обсуждаются архитектуры нейронных сетей, использованные в экспериментах. Гиперпараметры приведены в п. 5.7.2.

Решатель	Многомерный бенчмарк	Бенчмарк (CelebA)	Генерация изобр. CelebA
[LS]	$\psi_\theta, \varphi_\omega : \mathbb{R}^D \rightarrow \mathbb{R}$ - DenseICNN (U)		N/A
[MM-B]	$\psi_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ - DenseICNN (U)	$f_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ - ResNet	
[QC]	$\psi_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ - DenseICNN (U)	$f_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ - ResNet	
[MM]	$\psi_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ - DenseICNN (U) $H_\omega : \mathbb{R}^D \rightarrow \mathbb{R}^D$ - $\nabla$ of DenseICNN (U)	$f_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ - ResNet $H_\omega : \mathbb{R}^D \rightarrow \mathbb{R}^D$ - UNet	
[MM:R]	$T_\theta : \mathbb{R}^D \rightarrow \mathbb{R}^D$ - $\nabla$ of DenseICNN (U) $\varphi_\omega : \mathbb{R}^D \rightarrow \mathbb{R}$ - DenseICNN (U)		$T_\theta : \mathbb{R}^D \rightarrow \mathbb{R}^D$ - UNet $g_\omega : \mathbb{R}^D \rightarrow \mathbb{R}$ - ResNet
[MMv1]	$\psi_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ - DenseICNN		N/A
[MMv2]	$\psi_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ - DenseICNN	$\psi_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$ - ConvICNN64	
[W2]	$H_\omega : \mathbb{R}^D \rightarrow \mathbb{R}^D$ - $\nabla$ of DenseICNN	$H_\omega : \mathbb{R}^D \rightarrow \mathbb{R}^D$ - $\nabla$ of ConvICNN64	
[MMv2:R]	$T_\theta : \mathbb{R}^D \rightarrow \mathbb{R}^D$ - $\nabla$ of DenseICNN	$T_\theta : \mathbb{R}^D \rightarrow \mathbb{R}^D$ - $\nabla$ of ConvICNN64	
[W2:R]	$\varphi_\omega : \mathbb{R}^D \rightarrow \mathbb{R}$ - DenseICNN	$\varphi_\omega : \mathbb{R}^D \rightarrow \mathbb{R}$ - ConvICNN64	

Таблица 15 – Архитектуры сетей, используемых для параметризации потенциала  $f$  (или  $\Psi$ ) и отображения  $H$  в тестируемых решателях. В реверсированных решателях мы параметризуем второй потенциал  $g$  (или  $\varphi$ ) и прямое транспортное отображение  $T$  при помощи нейронных сетей.

### 5.7.1 Архитектуры нейронной сети

В табл. 15 мы приводим все архитектуры нейронных сетей, используемых в решателях задачи оптимального транспорта. В каждом эксперименте мы предварительно обучаем сети для удовлетворения условий  $\nabla\psi_\theta(x) = x - \nabla f_\theta(x) \approx x$  и  $H_\omega(y) \approx y$  в начале оптимизации. Мы эмпирически проверяем, что такая стратегия приводит к более устойчивой оптимизации.

При **многомерном бенчмарке** мы используем архитектуру DenseICNN из п. 2.7. Эта сеть есть полносвязная нейронная сеть с дополнительными квадратичными по входу обходными связями. Эта архитектура может быть сделана выпуклой по входу при требовании, что некоторые веса неотрицательны. Такое ограничение накладывается на решатели [MMv1], [MMv2], [W2], в которых требуется, чтобы сети были выпуклыми по входу. В других случаях ограничения по сетям не накладываются, и мы обозначаем архитектуру через DenseICNN (U). В экспериментах мы используем реализацию DenseICNN из главы 2. Более точно, в экспериментах с вероятностными мерами на  $\mathbb{R}^D$  мы используем

$$\text{DenseICNN}[1; \max(2D, 64), \max(2D, 64), \max(D, 32)].$$

Здесь 1 – ранг квадратичных по входу обходных связей, при этом по другим значениям определяются размеры полносвязных слоев последующей части сети. Обозначения такие же, как в п. 2.7.2.

Следует отметить, что архитектура DenseICNN сети  $\Psi_\theta$  имеет дифференцируемые функции активации CELU [138]. Таким образом, градиент  $\nabla \Psi_\theta$  корректно определен. В частности, искусственно  $\beta \cdot \|x\|^2/2$  для  $\beta = 10^{-4}$  добавляется к выходу последнего слоя сети ICNN. Как следствие, сеть  $\Psi_\theta$  является  $\beta$ -сильно выпуклой. Это влечет то, что  $\nabla \Psi_\theta$  – биективная функция с константой Липшица, ограниченной снизу величиной  $\beta$ .

**В экспериментах с изображениями из набора данных CelebA** для параметризации потенциала  $f = f_\theta : \mathbb{R}^D \rightarrow \mathbb{R}$  в [MM], [QC], [MM-B] мы используем архитектуру ResNet из официального репозитория WGAN-QC [43]:

<https://github.com/harryliew/WGAN-QC>

Для параметризации отображения  $H = H_\omega : \mathbb{R}^D \rightarrow \mathbb{R}^D$  в решателе [MM] мы используем архитектуру UNet из

<https://github.com/milesial/Pytorch-UNet>

В решателях [MMv2], [W2] мы параметризуем  $\Psi = \Psi_\theta$  и  $H = H_\omega = \nabla \varphi_\omega$ , где  $\Psi_\theta, \varphi_\omega$  имеют архитектуру ConvICNN64 (см. рис. 5.7). Мы искусственно добавляем  $\beta \cdot \|x\|^2/2$  (для  $\beta = 10^{-4}$ ) к выходу сети ConvICNN64 для обеспечения биективности градиента.

В этой архитектуре слои *PosConv2D* являются обычными 2D-сверточными слоями с неотрицательными весами (за исключением смещений). *Conv2D-CQ* (выпуклые квадратичные) – полностью сверточные блоки, дающие на выходе тензор, в котором элементами являются квадратичные по входу функции входного тензора. На рис. 5.8 показана архитектура блока Conv2D-CQ. Здесь операция *GroupChannelSumPool* соответствует расщеплению тензора по каналам на  $n_{out}$  последовательных подтензоров (каждый из которых состоит из  $r$  каналов) и объединения каждого подтензора в один 1-канал путем суммирования тензоров  $r$  отображений каналов. Слой можно представить как сверточный аналог полного слоя *ConvexQuadratic*, предложенного в п. 2.7.2.

**В экспериментах по генерации изображений CelebA** мы также использовали архитектуру ResNet для генерации сети  $g$ . Использована имплементация из упомянутого выше репозитория WGAN-QC.

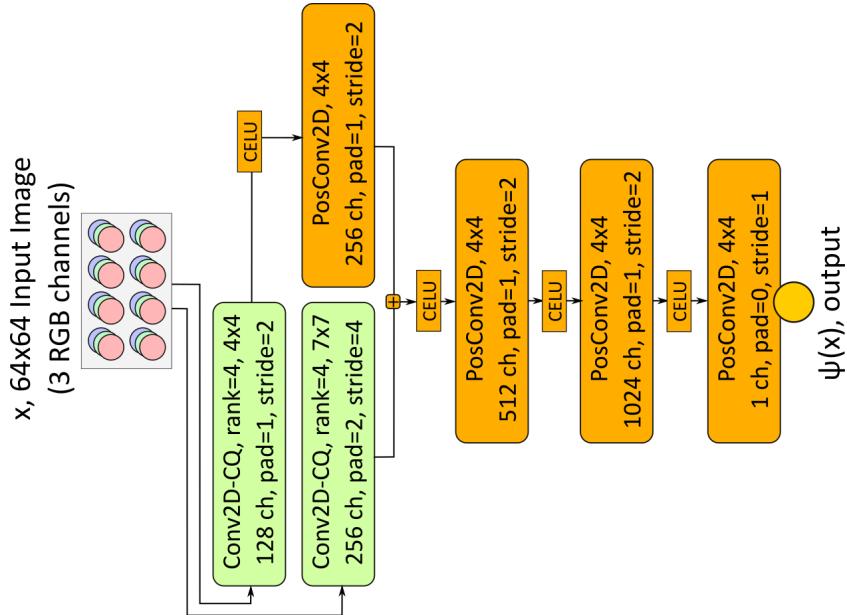


Рисунок 5.7 – Архитектура сверточной сети ICNN, используемой для обработки RGB-изображений размера  $64 \times 64$ .

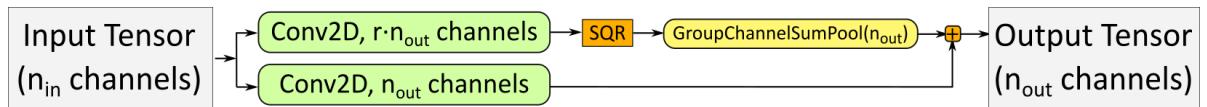


Рисунок 5.8 – 2D выпуклый квадратичный сверточный блок.

### 5.7.2 Гиперпараметры и детали имплементации

Оценка рассмотренных выше параметрических решателей не является тривиальной по двум следующим причинам. Во-первых, не для всех решателей имеется реализация на языке *Python*. Во-вторых, некоторые решатели не используются вне сетей GAN. Таким образом, для использования их в бенчмарке требуется правильное извлечение  $\mathbb{W}_2$ -решателя (части дискриминатора) из сети GAN.

Имплементация большинства решателей выполняется с нуля. Во всех случаях мы используем оптимизатор Adam [34] с заданными по умолчанию гиперпараметрами (за исключением скорости обучения). Для решателя [QC] из [43] мы использовали программный код, предоставленный авторами в официальном репозитории GitHub.

**Эталонные пары для большой размерности.** В табл. 16 приведены использованные нами гиперпараметры для эталонного тестирования в большой размерности. Столбец “общее число итераций” соответствует оптимизации

Решатель	Размер случайных выборок	Число итер.	Темп. обучения	Примечание
[LS]	1024	100000	$10^{-3}$	Квадратич. регуляризация с $\epsilon = 3 \cdot 10^{-2}$ , см. [2, (7)]
[MM-B]	1024	100000	$10^{-3}$	нет
[QC]	64	100000	$10^{-3}$	регуляризация задачи ОТ с $K = 1$ , $\gamma = 0.1$ , см. [43, (10)]
[MMv1]	1024	20000	$10^{-3}$	1000 итераций градиента ( $lr = 0.3$ ) для нахождения $\operatorname{argmin}$ в (5.8), см. [21, §6]. Ранняя остановка при норме градиента $< 10^{-3}$ .
[MM], [MMv2]	1024	50000	$10^{-3}$	15 внутр. циклич. итераций для обновления $H_\omega$ , ( $K = 15$ в обозначениях [7, Algorithm 1])
[W2]	1024	250000	$10^{-3}$	Циклически состоятельная регуляризация, $\lambda = D$ , см. главу 2

Таблица 16 — Гиперпараметры решателей, используемые при их тестировании на многомерном бенчмарке. Реверсированные решатели не представлены в таблице, поскольку в них используются те же гиперпараметры, что и в оригинальных версиях.

потенциала  $f_\theta$  (или  $\Psi_\theta$ ) с целью *максимизации* двойственной формы (5.8). В максиминных решателях также имеется внутренний цикл, соответствующий решению внутренней задачи *максимизации* в (5.8). Гиперпараметры выбирались эмпирически.

Для решателя [QC] большие размеры случайных выборок приводят к вычислительно нереализуемой подзадаче, поскольку при этом требуется решение линейной программы на каждом шаге оптимизации [43, §3.1]. Таким образом, мы используем размер выборки 64, как в оригинальной статье. Решатель [W2] используется с теми же гиперпараметрами для целей обучения и сравнения.

**Эталонные пары для изображений размера  $64 \times 64$  из набора данных CelebA.** Гиперпараметры изображений, использованных при тестировании решателей на бенчмарке, указаны в табл. 17.

**Эксперимент по генерации изображений размера  $64 \times 64$  набора данных CelebA.** Для обучения генеративной модели мы используем обучение по типу GAN: градиентные шаги сети генератора  $G_\alpha$  чередуются с шагами решателя для задачи оптимального транспорта (дискриминатора). Темп обучения для сети генератора равен  $3 \cdot 10^{-4}$ , общее число итераций генератора – 50000.

В решателе [QC] мы используем программный авторский код: выполняется одно обновление градиента решателя для задачи оптимального транспорта за

Решатель	Размер случайных выборок	Число итер.	Скор. обуч.	Примечание
[MM-B]	64	20000	$3 \cdot 10^{-4}$	нет
[QC]	64	20000	$3 \cdot 10^{-4}$	регуляризация задачи ОТ с $K = 1$ , $\gamma = 0.1$ , см. [43, (10)]
[MM]	64	50000	$3 \cdot 10^{-4}$	5 внутр. циклич. итераций для обновления $H_\omega$ , ( $K = 5$ в обозначениях [7, Algorithm 1])
[W2]	64	50000	$3 \cdot 10^{-4}$	Циклически-состоятельная регуляризация, $\lambda = 10^4$ , см. главу 2

Таблица 17 — Гиперпараметры решателей, использованные при их тестировании на бенчмарке с изображениями из CelebA.

одно обновление генератора. Во всех других методах одно обновление генератора чередуется с 10 обновлениями решателя для задачи оптимального транспорта (*итерациями* в обозначениях табл. 17). Все остальные гиперпараметры выбраны, как в предыдущем эксперименте.

Градиент генератора по  $\alpha$  на случайной выборке  $z_1, \dots, z_N \sim \mathbb{S}$  имеет вид

$$\begin{aligned} \partial \mathbb{W}_2^2(\mathbb{P}_\alpha, \mathbb{Q}) / \partial \alpha = \\ \int_z \mathbf{J}_\alpha G_\alpha(z)^T \nabla f^*(G_\alpha(z)) d\mathbb{S}(z) \approx \frac{1}{N} \sum_{n=1}^N \mathbf{J}_\alpha G_\alpha(z_n)^T \nabla f_\theta(G_\alpha(z_n)), \end{aligned} \quad (5.10)$$

где  $\mathbb{S}$  – мера на латентном пространстве и  $f_\theta$  – текущий потенциал (дискриминатор) решателя задачи оптимального транспорта. Отметим, что в решателе [MM:R] потенциал  $f$  не вычисляется; вместо этого параметризуется прямое отображение задачи оптимального транспорта  $T_\theta$ . В этом случае мы оцениваем градиент (5.10) на случайной выборке через  $\frac{1}{N} \sum_{n=1}^N \mathbf{J}_\alpha G_\alpha(z_n)^T (\text{id}_{\mathbb{R}^D} - T_\theta)$ .

## 5.8 Обсуждение

Наш метод генерирует пары непрерывных мер с эталонным отображением оптимального транспорта с квадратичной стоимостью. Это позволяет устранить имеющийся пробел в эталонном тестировании параметрических решателей задачи оптимального транспорта. Полученные в настоящей работе результаты позволяют нам оценить работу решателей задачи оптимального

транспорта с квадратичной стоимостью в задачах, близких к задачам оптимального транспорта. Помимо сравнения решателей транспортной задачи наш метод генеративного моделирования открывает неожиданное свойство: плохие решатели задачи оптимального транспорта могут тем не менее иметь хорошие генеративные функциональные характеристики, а простая реверсия несимметричных решателей может существенно повлиять на их производительность.

**Ограничения.** Мы рассматриваем градиенты ICNN как  $\mathbb{W}_2$ -отображения оптимального транспорта для генерации пар эталонных мер. В настоящий момент неясно, можно ли применить аналогичные конструкции для других функций стоимости (таких, например, как  $\mathbb{W}_1$ ). В проведенном исследовании найдено несоответствие в функциональных характеристиках некоторых решателей в применении к задачам оптимального транспорта и задачами генеративного моделирования. Например, требуется более глубокое исследование для нахождения ответа на вопрос о том, чем объясняются хорошие генеративные свойства [QC], и в тоже же время, его плохие свойства  $\mathbb{W}_2$ -аппроксимации.

**Возможные применения.** Мы ожидаем, что наши эталонные пары станут стандартным набором тестов для задач оптимального транспорта. Как результат, проведенная работа может улучшить качество и надежность решателей в прикладных задачах ОТ. Одним из негативных потенциальных эффектов может проявиться в том, что использование нашего бенчмарка может ограничить оценку качества будущих решателей задач оптимального транспорта на наборы данных нашего бенчмарка. Чтобы избежать этого, помимо конкретных наборов данных, в п. 5.4 мы описали общий метод для генерации новых контрольных пар.

## Заключение

Существующие численные методы ОТ в основном предназначены для данных, описываемых маломерными *дискретными* вероятностными распределениями. Это ограничение значительно сужает класс задач математического моделирования, для решения которых эти методы могут быть применены. В то же время существующие методы ОТ для *непрерывных* вероятностных распределений либо используют минимаксную оптимизацию, которая зачастую не стабильна и усложняет процесс обучения, либо вносят систематическую ошибку в решение (смещение), что ухудшает качество полученного решения.

Для устранения вышеупомянутых проблем в данной диссертационной работе представлены новые масштабируемые численные методы ОТ на основе выпуклых по входу нейронных сетей: предложены алгоритмы для вычисления отображений-/расстояний ОТ между непрерывными вероятностными распределениями (глава 2), барицентров Васерштейна семейств непрерывных распределений (глава 3) и Васерштейн градиентных потоков функционалов на пространстве вероятностных распределений (глава 4). Разработанные методы позволяют находить решения задач Вассерштейн-2 ОТ для непрерывных вероятностных распределений в пространствах большой размерности, не прибегая к минимаксной оптимизации или введению систематической ошибки в решение. Как следствие, они могут быть применены к задачам математического моделирования, где существующие численные методы ОТ работают неудовлетворительно или вовсе не применимы, например, из-за большой вычислительной сложности.

Важно также отметить, что для задачи Васерштейн-2 ОТ не существует нетривиального эталонного многомерного набора данных (бенчмарка). Таким образом, разработанная методология создания эталонных пар распределений с известным Васерштейн-2 ОТ отображением (глава 5) заполняет важный пробел в области вычислительного ОТ, тем самым позволяя проводить количественную оценку методов ОТ. Ожидается, что разработанные эталонные пары станут стандартным бенчмарком для оценки непрерывных (параметрических) алгоритмов ОТ.

Дальнейшим направлением исследований является разработка эффективных алгоритмов для задач непрерывного ОТ для **общих** транспортных стоимостей (помимо квадратичных).

## Публикации автора по теме диссертации

### **В изданиях, индексируемых Scopus:**

- A1 Korotin, A., Li, L., Genevay, A., Solomon, J. M., Filippov, A., & Burnaev, E. Do neural optimal transport solvers work? a continuous wasserstein-2 benchmark // Advances in Neural Information Processing Systems. — 2021. — Vol. 34. — P. 14593—14605.
- A2 Mokrov, P., Korotin, A., Li, L., Genevay, A., Solomon, J. M., & Burnaev, E. Large-scale wasserstein gradient flows // Advances in Neural Information Processing Systems. — 2021. — Vol. 34. — P. 15243—15256.

### **В трудах конференций:**

- A3 Korotin, A., Egiazarian, V., Asadulaev, A., Safin, A., & Burnaev, E. Wasserstein-2 Generative Networks // International Conference on Learning Representations. — 2021. [Электронный ресурс] URL: [https://openreview.net/pdf?id=bEoxzW\\_Exsa](https://openreview.net/pdf?id=bEoxzW_Exsa) (дата обращения 29.11.2022)
- A4 Korotin, A., Li, L., Solomon, J., & Burnaev, E. Continuous Wasserstein-2 Barycenter Estimation without Minimax Optimization // International Conference on Learning Representations. — 2021. [Электронный ресурс] URL: <https://openreview.net/pdf?id=3tFAs5E-Pe> (дата обращения 29.11.2022)
- A5 Rout, L., Korotin, A., & Burnaev, E. Generative Modeling with Optimal Transport Maps // International Conference on Learning Representations. — 2022. [Электронный ресурс] URL: <https://openreview.net/pdf?id=5JdLZg346Lw> (дата обращения 29.11.2022)

## Список литературы

1. *Arjovsky M., Chintala S., Bottou L.* Wasserstein GAN // arXiv preprint arXiv:1701.07875. — 2017.
2. Large Scale Optimal Transport and Mapping Estimation / V. Seguy [et al.] // International Conference on Learning Representations. — 2018.
3. Optimal transport for domain adaptation / N. Courty [et al.] // IEEE transactions on pattern analysis and machine intelligence. — 2016. — Vol. 39, no. 9. — P. 1853—1865.
4. *Aguech M., Carlier G.* Barycenters in the Wasserstein space // SIAM Journal on Mathematical Analysis. — 2011. — Vol. 43, no. 2. — P. 904—924.
5. *Jordan R., Kinderlehrer D., Otto F.* The variational formulation of the Fokker–Planck equation // SIAM journal on mathematical analysis. — 1998. — Vol. 29, no. 1. — P. 1—17.
6. Computational optimal transport: With applications to data science / G. Peyré, M. Cuturi, [et al.] // Foundations and Trends® in Machine Learning. — 2019. — Vol. 11, no. 5/6. — P. 355—607.
7. Optimal transport mapping via input convex neural networks / A. Makkuvu [et al.] // International Conference on Machine Learning. — PMLR. 2020. — P. 6672—6681.
8. Continuous Regularized Wasserstein Barycenters / L. Li [et al.] // arXiv preprint arXiv:2008.12534. — 2020.
9. *Santambrogio F.* Optimal transport for applied mathematicians // Birkhäuser, NY. — 2015. — Vol. 55, no. 58—63. — P. 94.
10. Scalable computation of monge maps with general costs / J. Fan [et al.] // arXiv preprint arXiv:2106.03812. — 2021.
11. *Frogner C., Poggio T.* Approximate inference with wasserstein gradient flows // International Conference on Artificial Intelligence and Statistics. — PMLR. 2020. — P. 2581—2590.
12. Generative adversarial nets / I. Goodfellow [et al.] // Advances in neural information processing systems. — 2014. — P. 2672—2680.

13. Nowozin S., Cseke B., Tomioka R. f-GAN: Training generative neural samplers using variational divergence minimization // Advances in neural information processing systems. — 2016. — P. 271—279.
14. Stabilizing adversarial nets with prediction methods / A. Yadav [et al.] // arXiv preprint arXiv:1705.07364. — 2017.
15. MMD GAN: Towards deeper understanding of moment matching network / C.-L. Li [et al.] // Advances in Neural Information Processing Systems. — 2017. — P. 2203—2213.
16. Mroueh Y., Sercu T. Fisher GAN // Advances in Neural Information Processing Systems. — 2017. — P. 2513—2523.
17. Existence and uniqueness of monotone measure-preserving maps / R. J. McCann [et al.] // Duke Mathematical Journal. — 1995. — Vol. 80, no. 2. — P. 309—324.
18. Peyré G. Mathematical Foundations of Data Sciences // def. — 2018. — Vol. 1, 2π. — 2π.
19. Rockafellar R. Characterization of the subdifferentials of convex functions // Pacific Journal of Mathematics. — 1966. — Vol. 17, no. 3. — P. 497—510.
20. Amos B., Xu L., Kolter J. Z. Input convex neural networks // Proceedings of the 34th International Conference on Machine Learning-Volume 70. — JMLR.org. 2017. — P. 146—155.
21. Taghvaei A., Jalali A. 2-Wasserstein Approximation via Restricted Convex Potentials with Application to Improved Training for GANs // arXiv preprint arXiv:1902.07197. — 2019.
22. A gradient descent solution to the Monge-Kantorovich problem / R. Chartrand [et al.] // Applied Mathematical Sciences. — 2009. — Vol. 3, no. 22. — P. 1071—1080.
23. Progressive growing of GANs for improved quality, stability, and variation / T. Karras [et al.] // arXiv preprint arXiv:1710.10196. — 2017.
24. Mirza M., Osindero S. Conditional generative adversarial nets // arXiv preprint arXiv:1411.1784. — 2014.
25. Villani C. Optimal transport: old and new. Vol. 338. — Springer Science & Business Media, 2008.

26. *Brenier Y.* Polar factorization and monotone rearrangement of vector-valued functions // Communications on pure and applied mathematics. — 1991. — Vol. 44, no. 4. — P. 375—417.
27. *Villani C.* Topics in optimal transportation. — American Mathematical Soc., 2003.
28. A geometric view of optimal transportation and generative model / N. Lei [et al.] // Computer Aided Geometric Design. — 2019. — Vol. 68. — P. 1—21.
29. *Fenchel W.* On conjugate convex functions // Canadian Journal of Mathematics. — 1949. — Vol. 1, no. 1. — P. 73—77.
30. Unpaired image-to-image translation using cycle-consistent adversarial networks / J.-Y. Zhu [et al.] // Proceedings of the IEEE international conference on computer vision. — 2017. — P. 2223—2232.
31. *Chen Y., Shi Y., Zhang B.* Optimal control via neural networks: A convex approach // arXiv preprint arXiv:1805.11835. — 2018.
32. *Paty F.-P., d'Aspremont A., Cuturi M.* Regularity as Regularization: Smooth and Strongly Convex Brenier Potentials in Optimal Transport // arXiv preprint arXiv:1905.10812. — 2019.
33. *Kakade S., Shalev-Shwartz S., Tewari A.* On the duality of strong convexity and strong smoothness: Learning applications and matrix regularization // Unpublished Manuscript, <http://ttic.uchicago.edu/shai/papers/Kakade-ShalevTewari09.pdf>. — 2009. — Vol. 2, no. 1.
34. *Kingma D. P., Ba J.* Adam: A method for stochastic optimization // arXiv preprint arXiv:1412.6980. — 2014.
35. Improved training of Wasserstein GANs / I. Gulrajani [et al.] // Advances in Neural Information Processing Systems. — 2017. — P. 5767—5777.
36. A fixed-point approach to barycenters in Wasserstein space / P. C. Álvarez-Este-  
ban [et al.] // Journal of Mathematical Analysis and Applications. — 2016. — Vol. 441, no. 2. — P. 744—762.
37. *Rabin J., Ferradans S., Papadakis N.* Adaptive color transfer with relaxed optimal transport // 2014 IEEE International Conference on Image Processing (ICIP). — IEEE. 2014. — P. 4852—4856.

38. Joint distribution optimal transportation for domain adaptation / N. Courty [et al.] // Advances in Neural Information Processing Systems. — 2017. — P. 3730—3739.
39. Optimal transport for multi-source domain adaptation under target shift / I. Redko [et al.] // arXiv preprint arXiv:1803.04899. — 2018.
40. Deepjdot: Deep joint distribution optimal transport for unsupervised domain adaptation / B. Bhushan Damodaran [et al.] // Proceedings of the European Conference on Computer Vision (ECCV). — 2018. — P. 447—463.
41. Deep Learning Face Attributes in the Wild / Z. Liu [et al.] // Proceedings of International Conference on Computer Vision (ICCV). — 12/2015.
42. GANs trained by a two time-scale update rule converge to a local nash equilibrium / M. Heusel [et al.] // Advances in neural information processing systems. — 2017. — P. 6626—6637.
43. *Liu H., Gu X., Samaras D.* Wasserstein GAN with quadratic transport cost // Proceedings of the IEEE International Conference on Computer Vision. — 2019. — P. 4832—4841.
44. *Burt P., Adelson E.* The Laplacian pyramid as a compact image code // IEEE Transactions on communications. — 1983. — Vol. 31, no. 4. — P. 532—540.
45. *Johnson J., Alahi A., Fei-Fei L.* Perceptual losses for real-time style transfer and super-resolution // European conference on computer vision. — Springer. 2016. — P. 694—711.
46. *Mroueh Y.* Wasserstein style transfer // arXiv preprint arXiv:1905.12828. — 2019.
47. Wasserstein barycenter and its application to texture mixing / J. Rabin [et al.] // International Conference on Scale Space and Variational Methods in Computer Vision. — Springer. 2011. — P. 435—446.
48. Convolutional Wasserstein distances: Efficient optimal transportation on geometric domains / J. Solomon [et al.] // ACM Transactions on Graphics (TOG). — 2015. — Vol. 34, no. 4. — P. 1—11.
49. *Korotin A., V'yugin V., Burnaev E.* Mixability of integral losses: A key to efficient online aggregation of functional and probabilistic forecasts // Pattern Recognition. — 2021. — Vol. 120. — P. 108175.

50. WASP: Scalable Bayes via barycenters of subset posteriors / S. Srivastava [et al.] // Artificial Intelligence and Statistics. — 2015. — P. 912—920.
51. *Srivastava S., Li C., Dunson D. B.* Scalable Bayes via barycenter in Wasserstein space // The Journal of Machine Learning Research. — 2018. — Vol. 19, no. 1. — P. 312—346.
52. *Fan J., Taghvaei A., Chen Y.* Scalable Computations of Wasserstein Barycenter via Input Convex Neural Networks // arXiv preprint arXiv:2007.04462. — 2020.
53. *Cohen S., Arbel M., Deisenroth M. P.* Estimating Barycenters of Measures in High Dimensions // arXiv preprint arXiv:2007.07105. — 2020.
54. *Cuturi M.* Sinkhorn distances: Lightspeed computation of optimal transport // Advances in neural information processing systems. — 2013. — P. 2292—2300.
55. *Cuturi M., Doucet A.* Fast computation of Wasserstein barycenters. — 2014.
56. Decentralize and randomize: Faster algorithm for Wasserstein barycenters / P. Dvurechenskii [et al.] // Advances in Neural Information Processing Systems. — 2018. — P. 10760—10770.
57. Parallel streaming Wasserstein barycenters / M. Staib [et al.] // Advances in Neural Information Processing Systems. — 2017. — P. 2647—2658.
58. Stochastic optimization for large-scale optimal transport / A. Genevay [et al.] // Advances in neural information processing systems. — 2016. — P. 3440—3448.
59. On Scalable and Efficient Computation of Large Scale Optimal Transport / Y. Xie [et al.] // . Vol. 97 / ed. by K. Chaudhuri, R. Salakhutdinov. — Long Beach, California, USA : PMLR, 09–15 Jun/2019. — P. 6882—6892. — (Proceedings of Machine Learning Research). — URL: <http://proceedings.mlr.press/v97/xie19a.html>.
60. Gradient descent algorithms for Bures-Wasserstein barycenters / S. Chewi [et al.] // arXiv preprint arXiv:2001.01700. — 2020.
61. Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance / J. Weed, F. Bach, [et al.] // Bernoulli. — 2019. — Vol. 25, 4A. — P. 2620—2648.

62. *Dowson D., Landau B.* The Fréchet distance between multivariate normal distributions // Journal of multivariate analysis. — 1982. — Vol. 12, no. 3. — P. 450—455.
63. Stan: A probabilistic programming language / B. Carpenter [et al.] // Journal of statistical software. — 2017. — Vol. 76, no. 1.
64. *Sobczyk K.* Stochastic differential equations: with applications to physics and engineering. Vol. 40. — Springer Science & Business Media, 2013.
65. *El Karoui N., Peng S., Quenez M. C.* Backward stochastic differential equations in finance // Mathematical finance. — 1997. — Vol. 7, no. 1. — P. 1—71.
66. *Platen E., Bruti-Liberati N.* Numerical solution of stochastic differential equations with jumps in finance. Vol. 64. — Springer Science & Business Media, 2010.
67. Population stochastic modelling (PSM)—an R package for mixed-effects models based on stochastic differential equations / S. Klim [et al.] // Computer methods and programs in biomedicine. — 2009. — Vol. 94, no. 3. — P. 279—289.
68. *Kalman R. E., Bucy R. S.* New results in linear filtering and prediction theory. — 1961.
69. *Doucet A., Johansen A. M.* A tutorial on particle filtering and smoothing: Fifteen years later // Handbook of nonlinear filtering. — 2009. — Vol. 12, no. 656—704. — P. 3.
70. *Welling M., Teh Y. W.* Bayesian learning via stochastic gradient Langevin dynamics // Proceedings of the 28th international conference on machine learning (ICML-11). — Citeseer. 2011. — P. 681—688.
71. *Chang J., Cooper G.* A practical difference scheme for Fokker-Planck equations // Journal of Computational Physics. — 1970. — Vol. 6, no. 1. — P. 1—16.
72. *Pareschi L., Zanella M.* Structure preserving schemes for nonlinear Fokker-Planck equations and applications // Journal of Scientific Computing. — 2018. — Vol. 74, no. 3. — P. 1575—1600.

73. *Burger M., Carrillo J. A., Wolfram M.-T.* A mixed finite element method for nonlinear diffusion equations // *Kinetic & Related Models*. — 2010. — Vol. 3, no. 1. — P. 59.
74. *Carrillo J. A., Chertock A., Huang Y.* A finite-volume method for nonlinear nonlocal equations with a gradient flow structure // *Communications in Computational Physics*. — 2015. — Vol. 17, no. 1. — P. 233—258.
75. Dynamical optimal transport on discrete surfaces / H. Lavenant [et al.] // *ACM Transactions on Graphics (TOG)*. — 2018. — Vol. 37, no. 6. — P. 1—16.
76. *Peyré G.* Entropic approximation of Wasserstein gradient flows // *SIAM Journal on Imaging Sciences*. — 2015. — Vol. 8, no. 4. — P. 2323—2351.
77. *Kloeden P. E.* Numerical solution of stochastic differential equations / Peter E. Kloeden, Eckhard Platen. — Berlin : Springer, 1992. — (Applications of mathematics; v. 23).
78. *Kushner H.* Approximations to optimal nonlinear filters // *IEEE Transactions on Automatic Control*. — 1967. — Vol. 12, no. 5. — P. 546—556.
79. *Julier S. J., Uhlmann J. K., Durrant-Whyte H. F.* A new approach for filtering nonlinear systems // *Proceedings of 1995 American Control Conference-ACC'95*. Vol. 3. — IEEE. 1995. — P. 1628—1632.
80. *Sarkka S.* On unscented Kalman filtering for state estimation of continuous-time nonlinear systems // *IEEE Transactions on automatic control*. — 2007. — Vol. 52, no. 9. — P. 1631—1641.
81. A novel Gaussian Sum Filter Method for accurate solution to the nonlinear filtering problem / G. Terejanu [et al.] // *2008 11th International Conference on Information Fusion*. — IEEE. 2008. — P. 1—8.
82. *Ala-Luhtala J., Särkkä S., Piché R.* Gaussian filtering and variational approximations for Bayesian smoothing in continuous-discrete stochastic dynamic systems // *Signal Processing*. — 2015. — Vol. 111. — P. 124—136.
83. *Sutter T., Ganguly A., Koeppl H.* A variational approach to path estimation and parameter inference of hidden diffusion processes // *The Journal of Machine Learning Research*. — 2016. — Vol. 17, no. 1. — P. 6544—6580.

84. *Opper M.* Variational inference for stochastic differential equations // Annalen der Physik. — 2019. — Vol. 531, no. 3. — P. 1800233.
85. Stochastic optimization for large-scale optimal transport / G. Aude [et al.] // arXiv preprint arXiv:1605.08527. — 2016.
86. *Ambrosio L., Gigli N., Savaré G.* Gradient flows: in metric spaces and in the space of probability measures. — Springer Science & Business Media, 2008.
87. *Santambrogio F.* {Euclidean, metric, and Wasserstein} gradient flows: an overview // Bulletin of Mathematical Sciences. — 2017. — Vol. 7, no. 1. — P. 87—154.
88. Maximum mean discrepancy gradient flow / M. Arbel [et al.] // arXiv preprint arXiv:1906.04370. — 2019.
89. Sliced-Wasserstein flows: Nonparametric generative modeling via optimal transport and diffusions / A. Liutkus [et al.] // International Conference on Machine Learning. — PMLR. 2019. — P. 4104—4113.
90. Deep generative learning via variational gradient flow / Y. Gao [et al.] // International Conference on Machine Learning. — PMLR. 2019. — P. 2093—2101.
91. *Ansari A. F., Ang M. L., Soh H.* Refining Deep Generative Models via Wasserstein Gradient Flows // arXiv preprint arXiv:2012.00780. — 2020.
92. *Richemond P. H., Maginnis B.* On Wasserstein reinforcement learning and the Fokker-Planck equation // arXiv preprint arXiv:1712.07185. — 2017.
93. Policy optimization as Wasserstein gradient flows / R. Zhang [et al.] // International Conference on Machine Learning. — PMLR. 2018. — P. 5737—5746.
94. *Maury B., Roudneff-Chupin A., Santambrogio F.* A macroscopic crowd motion model of gradient flow type // Mathematical Models and Methods in Applied Sciences. — 2010. — Vol. 20, no. 10. — P. 1787—1821.
95. *Santambrogio F.* Gradient flows in Wasserstein spaces and applications to crowd movement // Séminaire Équations aux dérivées partielles (Polytechnique). — 2010. — P. 1—16.
96. *Alvarez-Melis D., Fusi N.* Gradient Flows in Dataset Space // arXiv preprint arXiv:2010.12760. — 2020.
97. *Gao Y., Jin G., Liu J.-G.* Inbetweening auto-animation via Fokker-Planck dynamics and thresholding // arXiv preprint arXiv:2005.08858. — 2020.

98. Discretization of functionals involving the Monge–Ampère operator / J.-D. Benamou [et al.] // Numerische mathematik. — 2016. — Vol. 134, no. 3. — P. 611—636.
99. *Risken H.* Fokker-planck equation // The Fokker-Planck Equation. — Springer, 1996. — P. 63—95.
100. *Vatiwutipong P., Phewchean N.* Alternative way to derive the distribution of the multivariate Ornstein–Uhlenbeck process // Advances in Difference Equations. — 2019. — Vol. 2019, no. 1. — P. 1—7.
101. *Caluya K. F., Halder A.* Proximal Recursion for Solving the Fokker-Planck Equation. — 2019.
102. *Liu Q., Wang D.* Stein variational gradient descent: A general purpose bayesian inference algorithm // arXiv preprint arXiv:1608.04471. — 2016.
103. Fisher discriminant analysis with kernels / S. Mika [et al.] // Neural networks for signal processing IX: Proceedings of the 1999 IEEE signal processing society workshop (cat. no. 98th8468). — Ieee. 1999. — P. 41—48.
104. *Linnainmaa S.* The representation of the cumulative rounding error of an algorithm as a Taylor expansion of the local rounding errors // Master’s Thesis (in Finnish), Univ. Helsinki. — 1970. — P. 6—7.
105. Analyzing inverse problems with invertible neural networks / L. Ardizzone [et al.] // arXiv preprint arXiv:1808.04730. — 2018.
106. *Jacobsen J.-H., Smeulders A., Oyallon E.* i-revnet: Deep invertible networks // arXiv preprint arXiv:1802.07088. — 2018.
107. *Blei D. M., Kucukelbir A., McAuliffe J. D.* Variational inference: A review for statisticians // Journal of the American statistical Association. — 2017. — Vol. 112, no. 518. — P. 859—877.
108. *Rezende D. J., Mohamed S.* Variational inference with normalizing flows // arXiv preprint arXiv:1505.05770. — 2015.
109. Advances in variational inference / C. Zhang [et al.] // IEEE transactions on pattern analysis and machine intelligence. — 2018. — Vol. 41, no. 8. — P. 2008—2026.

110. *Ubaru S., Chen J., Saad Y.* Fast estimation of  $\text{tr}(f(a))$  via stochastic lanczos quadrature // SIAM Journal on Matrix Analysis and Applications. — 2017. — Vol. 38, no. 4. — P. 1075—1099.
111. *Han I., Malioutov D., Shin J.* Large-scale log-determinant computation through stochastic Chebyshev expansions // International Conference on Machine Learning. — PMLR. 2015. — P. 908—917.
112. Convex Potential Flows: Universal Probability Distributions with Optimal Transport and Convex Optimization / C.-W. Huang [et al.] // arXiv preprint arXiv:2012.05942. — 2020.
113. *Amos B., Kolter J. Z.* Optnet: Differentiable optimization as a layer in neural networks // International Conference on Machine Learning. — PMLR. 2017. — P. 136—145.
114. *Hashimoto T., Gifford D., Jaakkola T.* Learning population-level diffusions with generative RNNs // International Conference on Machine Learning. — PMLR. 2016. — P. 2417—2426.
115. JKOnet: Proximal Optimal Transport Modeling of Population Dynamics / C. Bunne [et al.] // arXiv preprint arXiv:2106.06345. — 2021.
116. *Alvarez-Melis D., Schiff Y., Mroueh Y.* Optimizing Functionals on the Space of Probabilities with Input Convex Neural Networks // arXiv preprint arXiv:2106.00774. — 2021.
117. Molecular Sets (MOSES): A Benchmarking Platform for Molecular Generation Models / D. Polykovskiy [et al.] // Frontiers in Pharmacology. — 2020. — Vol. 11. — P. 1931. — URL: <https://www.frontiersin.org/article/10.3389/fphar.2020.565644>.
118. Wasserstein divergence for GANs / J. Wu [et al.] // Proceedings of the European Conference on Computer Vision (ECCV). — 2018. — P. 653—668.
119. *Liu H., Xianfeng G., Samaras D.* A two-step computation of the exact GAN Wasserstein distance // International Conference on Machine Learning. — PMLR. 2018. — P. 3159—3168.
120. Multi-marginal Wasserstein GAN / J. Cao [et al.] // arXiv preprint arXiv:1911.00888. — 2019.

121. Wasserstein distance guided representation learning for domain adaptation / J. Shen [et al.] // Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 32. — 2018.
122. WGAN domain adaptation for EEG-based emotion recognition / Y. Luo [et al.] // International Conference on Neural Information Processing. — Springer. 2018. — P. 275—286.
123. *Mallasto A., Montúfar G., Gerolin A.* How Well Do WGANs Estimate the Wasserstein Metric? // arXiv preprint arXiv:1910.03875. — 2019.
124. *Pinetz T., Soukup D., Pock T.* On the estimation of the Wasserstein distance in generative models // German Conference on Pattern Recognition. — Springer. 2019. — P. 156—170.
125. Wasserstein GANs Work Because They Fail (to Approximate the Wasserstein Distance) / J. Stanczuk [et al.] // arXiv preprint arXiv:2103.01678. — 2021.
126. *Schrieber J., Schuhmacher D., Gottschlich C.* Dotmark—a benchmark for discrete optimal transport // IEEE Access. — 2016. — Vol. 5. — P. 271—282.
127. *Kantorovitch L.* On the translocation of masses // Management Science. — 1958. — Vol. 5, no. 1. — P. 1—4.
128. On the convergence and robustness of training GANs with regularized optimal transport / M. Sanjabi [et al.] // arXiv preprint arXiv:1802.08249. — 2018.
129. *Milgrom P., Segal I.* Envelope theorems for arbitrary choice sets // Econometrica. — 2002. — Vol. 70, no. 2. — P. 583—601.
130. (q, p)-Wasserstein GANs: Comparing Ground Metrics for Wasserstein GANs / A. Mallasto [et al.] // arXiv preprint arXiv:1902.03642. — 2019.
131. Adversarial computation of optimal transport maps / J. Leygonie [et al.] // arXiv preprint arXiv:1906.09691. — 2019.
132. Learning High Dimensional Wasserstein Geodesics / S. Liu [et al.] // arXiv preprint arXiv:2102.02992. — 2021.
133. Large-Scale Optimal Transport via Adversarial Training with Cycle-Consistency / G. Lu [et al.] // arXiv preprint arXiv:2003.06635. — 2020.
134. Are GANs created equal? a large-scale study / M. Lucic [et al.] // Advances in neural information processing systems. — 2018. — P. 700—709.

135. Threeplayer Wasserstein GAN via amortised duality / Q. H. Nhan Dam [et al.] // Proc. of the 28th Int. Joint Conf. on Artificial Intelligence (IJCAI). — 2019.
136. A gradual, semi-discrete approach to generative network training via explicit Wasserstein minimization / Y. Chen [et al.] // International Conference on Machine Learning. — PMLR. 2019. — P. 1071—1080.
137. The cramer distance as a solution to biased Wasserstein gradients / M. G. Bellemare [et al.] // arXiv preprint arXiv:1705.10743. — 2017.
138. *Barron J. T.* Continuously differentiable exponential linear units // arXiv preprint arXiv:1704.07483. — 2017.

## Список рисунков

1.1	Схематически изображенная задача 1 оптимального транспорта между распределениями $\mathbb{P}, \mathbb{Q}$ с квадратичной функцией стоимости. . . . .	15
1.2	Схематически изображенные задачи 2, 3 и схема JKO [5]. . . . .	18
2.1	Два возможных генеративных отображения, переводящих распределение $\mathbb{P}$ в распределение $\mathbb{Q}$ . . . . .	22
2.2	Смесь 100 гауссовых распределений $\mathbb{Q}$ и распределение $\nabla\psi_\theta\#\mathbb{P}$ , вычисленное нашим алгоритмом. . . . .	41
2.3	Синтетические распределения, обученные нашим алгоритмом. . . . .	41
2.4	Сравнение скорости сходимости для методов W2GN, MM-1 и MM-2 для размерностей $D = 64, 256, 1024, 4096$ . . . . .	43
2.5	Результаты передачи цветов между изображениями с высоким разрешением (размера $\approx 10$ мегапикселей) при помощи циклически монотонных отображений. . . . .	45
2.6	Результаты передачи цветов между изображениями с помощью циклически монотонного отображения. . . . .	46
2.7	Пара главных компонент пространства признаков. Слева направо: пространство признаков MNIST; признаки USPS, отображенные W2GN; оригинальные признаки USPS. Главные компоненты получены из признаков MNIST. Цвета соответствуют различным классам цифр 0 – 9. . . . .	47
2.8	Схема транспорта массы в латентном пространстве. . . . .	49
2.9	Изображения, декодированные из стандартного латентного гауссовского шума (1-я строка) и из того же шума, перенесенного нашим циклическим монотонным отображением (2-я строка). . . . .	49
2.10	Схема переноса стилей между наборами изображений с помощью пары ConvICNN обученных нашим методом. . . . .	50
2.11	Результаты переноса стиля между изображениями, полученными при использовании сети ConvICNN для изображений $128 \times 128$ пикселей.	50
2.12	Дополнительные результаты переноса стиля между изображениями для наборов данных Winter2Summer and Photo2Monet. . . . .	51
2.13	Общая архитектура выпуклой по входу нейронной сети. . . . .	52
2.14	Полносвязная выпуклая по входу нейронная сеть. . . . .	54

2.15	Выпуклая нейронная сеть со сверточным входом. Все сверточные слои имеют 128 каналов.	55
3.1	Барицентр линейно-разбросанной популяции “швейцарский рулет”, вычисленный по трем методам.	75
3.2	Результаты работы нашего метода, примененного к усреднению цветовых палитр изображений.	78
4.1	Симметрическая дивергенция Кульбака–Лейблера SymKL между вычисленной и стационарной мерой $D = 2, 4, \dots, 12$	92
4.2	Проекции на две первые РСА-компоненты (анализ главных компонент) истинной стационарной меры и меры, полученной нашим методом для размерностей $D = 13$ (слева) и $D = 32$ (справа).	92
4.3	Значения симметрической дивергенции Кульбака–Лейблера между вычисленной мерой и истинной мерой $\rho_t$ при $t = 0.5$ (слева) и $t = 0.9$ (справа) для размерностей $D = 1, 2, \dots, 12$ .	93
5.1	Пример создания эталонной пары для размерности $D = 16$ . Сначала инициализируются 3 случайные смеси гауссовских распределений $\mathbb{P}$ и $\mathbb{Q}_1, \mathbb{Q}_2$ и обучаются два приближенные отображения задачи оптимального транспорта $\nabla \Psi_i \sharp \mathbb{P} \approx \mathbb{Q}_i$ , $i = 1, 2$ . Усреднение потенциалов используется для определения выходной меры: $\frac{1}{2}(\nabla \Psi_1 + \nabla \Psi_2) \sharp \mathbb{P}$ . Каждый рисунок содержит 512 случайных точек, спроектированных на две принципиальные компоненты меры $\frac{1}{2}(\nabla \Psi_1 + \nabla \Psi_2) \sharp \mathbb{P}$ .	113
5.2	Методология создания эталонных пар на пространстве изображений. Мы используем 3 генеративных модели: $\mathbb{P}_{\text{Final}}^3$ (хорошо обученная) и $\mathbb{Q}_{\text{Cpkt}}^1, \mathbb{Q}_{\text{Cpkt}}^2$ (недообученные). Для $k = 1, 2$ мы обучаем приближенное ОТ отображение $\mathbb{P}_{\text{Final}}^3 \rightarrow \mathbb{Q}_{\text{Cpkt}}^k$ как $\nabla \Psi_{\text{Cpkt}}^k$ , т.е. градиент сети ICNN. Определим эталонную пару через $(\mathbb{P}_{\text{CelebA}}, \mathbb{Q}_{\text{Cpkt}}) \stackrel{\text{def}}{=} (\mathbb{P}_{\text{Final}}^3, \frac{1}{2}(\nabla \Psi_{\text{Cpkt}}^1 + \nabla \Psi_{\text{Cpkt}}^2) \sharp \mathbb{P}_{\text{Final}}^3)$ . На визуализации представлена эталонная пара Early.	114
5.3	Визуализация 64-мерной эталонной пары и ОТ отображений обученных решателями. Диаграммы рассеяния содержат 512 случайных выборок, спроектированных на две главные компоненты меры $\nabla \Psi^* \sharp \mathbb{P}$ .	117

5.4 Отображения, обученные решателями на эталонных парах ( $\mathbb{Q}_{\text{Cpkt}}, \mathbb{P}_{\text{CelebA}}$ ). В первой строке даны случайные $x \sim \mathbb{Q}_{\text{Cpkt}}$ . Вторая строка содержит выборки из $\mathbb{P}_{\text{CelebA}}$ , полученные прямым отображением $x$ посредством ОТ отображения $T^* = \nabla\psi^*$ . Далее показаны образы $x$ под действием отображений, обученных нашими решателями задачи оптимального транспорта. . . . .	118
5.5 Случайные изображения, полученные в результате обучения генеративной модели с помощью решателя. В первой строке показаны случайные сгенерированные изображения $x = G_\alpha(z) \sim \mathbb{P}_\alpha$ , $z \sim \mathbb{S}$ . Во второй строке показано вычисленное транспортное отображение из сгенерированного $x = G_\alpha(z) \sim \mathbb{P}_\alpha$ в $\mathbb{Q}$ . . . . .	121
5.6 Случайно сгенерированные смеси гауссовских распределений. Проекция на первые две размерности. . . . .	122
5.7 Архитектура сверточной сети ICNN, используемой для обработки RGB-изображений размера $64 \times 64$ . . . . .	126
5.8 2D выпуклый квадратичный сверточный блок. . . . .	126

## Список таблиц

1	Сравнение $\mathcal{L}^2$ -UVP (%) для методов LSOT, MM-1, MM-2 и (нашего) метода W2GN в размерностях $D = 2, 4, \dots, 2^{12}$ . . . . .	43
2	Точность классификации (больше = лучше) 1-NN на наборах данных USPS $\rightarrow$ MNIST в задаче адаптации домена. . . . .	47
3	Значения метрики FID для сгенерированных изображений. . . . .	48
4	Сравнение UVP в случае $\{\mathbb{P}_n\} \subset \mathcal{F}(\mathbb{P}_0)$ , $\mathbb{P}_0 = \mathcal{N}(0, I_D)$ , $N = 4$ . . . . .	76
5	Сравнение UVP в случае $\{\mathbb{P}_n\} \subset \mathcal{F}(\mathbb{P}_0)$ , $\mathbb{P}_0 = \text{Uniform}([- \sqrt{3}, + \sqrt{3}]^D)$ , $N = 4$ . . . . .	76
6	Сравнение UVP в случае $\{\mathbb{P}_n\} \subset \mathcal{F}(\mathbb{P}_0)$ , $\mathbb{P}_0 = \text{Uniform}([- \sqrt{3}, + \sqrt{3}]^D)$ , $N = 20$ . . . . .	76
7	Сравнение UVP для восстановленных барицентров в нашей задаче агрегации апостериорных распределений подмножеств. . . . .	77
8	Сравнение нашего метода с [SVGD] [102] для байесовской логистической регрессии. . . . .	95
9	Гиперпараметры, используемые в эксперименте. . . . .	96
10	Гиперпараметры, используемые в экспериментах по байесовской логистической регрессии. . . . .	97
11	Вычислительная сложность операций в нашем методе для вычисления шагов схемы JKO с помощью сетей ICNN. . . . .	98
12	Сравнение существующих параметрических двойственных решателей задач оптимального транспорта 1. . . . .	109
13	Значения метрик $\mathcal{L}^2$ -UVP (%), слева) и $\cos \in [-1, 1]$ (справа) для транспортных отображений, обученных при помощи решателей задачи оптимального транспорта многомерных на эталонных парах в размерностях $D = 2, 2^2, \dots, 2^8$ . Оранжевый цвет — $\mathcal{L}^2$ -UVP > 10% и $\cos < 0.95$ . Красный цвет — функционирование хуже, чем у бейзлайна [L]. . . . .	117
14	Значения метрик $\mathcal{L}^2$ -UVP (%), слева) и $\cos \in [-1, 1]$ (справа) для транспортных отображений $\mathbb{Q}_{\text{Cpkt}} \rightarrow \mathbb{P}_{\text{CelebA}}$ , обученных решателями оптимального транспорта на 3 разработанных CelebA64 $\mathbb{W}_2$ -эталонных наборах. . . . .	118

15	Архитектуры сетей, используемых для параметризации потенциала $f$ (или $\Psi$ ) и отображения $H$ в тестируемых решателях. В реверсированных решателях мы параметризуем второй потенциал $g$ (или $\varphi$ ) и прямое транспортное отображение $T$ при помощи нейронных сетей. . . . .	124
16	Гиперпараметры решателей, используемые при их тестировании на многомерном бенчмарке. Реверсированные решатели не представлены в таблице, поскольку в них используются те же гиперпараметры, что и в оригинальных версиях. . . . .	127
17	Гиперпараметры решателей, использованные при их тестировании на бенчмарке с изображениями из CelebA. . . . .	128