

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М. В. ЛОМОНОСОВА

На правах рукописи

Кузьмин Сергей Сергеевич

**Методы машинного обучения для характеристики
квантовых состояний и программирования
линейно-оптических интерферометров**

Специальность 1.3.19. Лазерная физика

Автореферат
диссертации на соискание учёной степени
кандидата физико-математических наук

Москва — 2026

Диссертация подготовлена на кафедре квантовой электроники физического факультета МГУ имени М. В. Ломоносова.

Научный руководитель: **Страупе Станислав Сергеевич**,
кандидат физико-математических наук

Официальные оппоненты: **Богданов Юрий Иванович**,
доктор физико-математических наук, Национальный исследовательский центр «Курчатовский институт», отделение Физико-технологических исследований имени К. А. Валиева, лаборатория физики квантовых компьютеров, главный научный сотрудник

Калачев Алексей Алексеевич,
доктор физико-математических наук, доцент, член-корреспондент РАН, Федеральный исследовательский центр «Казанский научный центр Российской академии наук», дирекция, директор

Оседец Иван Валерьевич,
доктор физико-математических наук, Институт искусственного интеллекта AIRI, дирекция, генеральный директор

Защита состоится 21 октября 2026 г. в 15 часов 00 минут на заседании диссертационного совета МГУ.013.4 Московского государственного университета имени М. В. Ломоносова по адресу: 119991, Москва, Ленинские горы, д. 1, стр. 62, Корпус нелинейной оптики, аудитория им. С. А. Ахманова.

E-mail: diss.sov.31@physics.msu.ru.

С диссертацией можно ознакомиться в отделе диссертаций научной библиотеки МГУ имени М. В. Ломоносова (Ломоносовский проспект, д. 27) и на портале: <https://dissovet.msu.ru/dissertation/4042>.

Автореферат разослан «__» сентября 2026 года.

Ученый секретарь
диссертационного совета
МГУ.013.4,
кандидат физико-математических
наук

Коновко Андрей Андреевич

Общая характеристика работы

Квантовые технологии быстро переходят от лабораторных демонстраций к инженерным системам, в которых ключевую роль играют воспроизводимость, программируемость и масштабируемость аппаратных платформ. Методы линейной оптики и интегральной фотоники используются во многих современных экспериментах и прототипах квантовых устройств. Многомодовые интерферометры позволяют реализовывать унитарные преобразования оптических мод и служат основой для обработки квантовой информации, генерации и анализа многомерных запутанных состояний, а также выполнения фотонных вычислений. Современные фотонные схемы, как правило, содержат большое число управляемых параметров, прежде всего фазовых сдвигов. Это обеспечивает программируемость таких устройств, однако одновременно повышает сложность их калибровки, настройки и последующей эксплуатации.

В общем случае программирование линейно-оптического интерферометра заключается в выборе управляющих параметров, при которых устройство реализует заданное преобразование, чаще всего представленное унитарной матрицей. Для некоторых «идеальных» универсальных архитектур разработаны аналитические алгоритмы программирования, основанные на декомпозиции унитарных матриц. Вместе с тем в реальных фотонных чипах и оптических системах возникают технологические отклонения, паразитные связи, потери и дрейфы параметров, вследствие чего реализуемое преобразование может значительно расходиться с проектным. Для более устойчивых к дефектам архитектур аналитическое преобразование «целевая матрица $\rightarrow$ набор фаз» как правило отсутствует. В этом случае программирование целевого преобразования превращается в итерационную оптимизацию на физическом устройстве, предполагающую многократную реконструкцию (томографию) преобразования, которая быстро становится ресурсоёмкой и нестабильной с увеличением размерности интерферометра.

Преодолеть эти ограничения можно с помощью построения «цифрового двойника» устройства: численной модели, согласованной с экспериментальными данными и используемой для быстрого расчёта управляющих параметров без обращения к физическому устройству (без томографии) на каждом шаге оптимизации. Инструментом для такого подхода служат методы машинного обучения, в которых параметры модели подбираются по обучающей выборке вида «управляющие параметры $\rightarrow$ наблюдаемый отклик». Для фотонных интерферометров это означает обучение модели на основе набора матриц преобразования, реконструированных при различных настройках фаз. Затем подбор фаз, реализующих целевое преобразование, осуществляется с использованием обученной модели.

Помимо задач программирования линейно-оптических устройств существует фундаментальная проблема эффективной характеристики (томографии) квантовых состояний. Полная квантовая томография требует экспоненциально растущего с размером системы числа измерений, что ограничивает применимость стандартных протоколов для квантовых устройств промежуточного масштаба с шумом (noisy intermediate-scale quantum, NISQ). Это мотивирует развивать методы неполной характеристики, которые используют априорные предположения о структуре состояния и нацелены на предсказание экспериментально наблюдаемых величин, а не на восстановление полного описания состояния. В этой области важную роль играет формализм тензорных сетей, позволяющий компактно параметризовать широкий класс физических состояний, а также используются подходы, близкие к парадигме вероятно приближённо корректного обучения (probably approximately correct, PAC) и методам теневой томографии, в которых для предсказания результатов будущих измерений достаточно сравнительно небольшого набора измерений.

Диссертация посвящена применению методов машинного обучения к двум связанным задачам квантовых технологий: построению и обучению моделей для линейно-оптических интерферометров и алгоритмов их программирования, а также обучению компактных тензорно-сетевых моделей квантовых состояний по ограниченному набору результатов локальных измерений для предсказания вероятностей исходов будущих измерений. Общей методологической основой работы является перенос ресурсоёмких процедур настройки и характеристики в пространство обучаемых моделей, использующих априорные структурные предположения о физической системе и обеспечивающих более эффективное масштабирование с ростом числа параметров системы.

В главе 1 представлен обзор работ, посвящённых линейно-оптическим интерферометрам и их универсальным архитектурам, а также реконструкции реализуемых ими преобразований. В главе 2 рассматривается обучение архитектурно-независимой модели линейно-оптического интерферометра по экспериментальным данным и её последующее использование для программирования экспериментального устройства. Приведены результаты экспериментальной проверки предложенного подхода, а также исследовано влияние факторов, существенных для его практического применения, включая паразитные взаимные влияния (кросс-токи) и ограничения схемы управления. В главе 3 предложены методы повышения эффективности обучения модели интерферометра и его программирования. Они основаны на использовании специальных признаков, зависящих от фаз в плечах интерферометра, и линеаризации зависимости реализуемого преобразования от этих признаков. В главе 4 представлен обзор тензорных сетей и рассмотрены их применения для моделирования квантовых состояний. В главе 5 предложен и исследован протокол обучения тензорно-сетевой

модели неизвестного квантового состояния из заданного класса состояний по результатам измерений, задаваемых операторами, представимыми в виде тензорных произведений однокубитных операторов. Обсуждаются условия применимости метода и его масштабируемость. В заключении сформулированы основные результаты работы и обозначены направления дальнейших исследований.

Актуальность работы связана с двумя задачами. Первая состоит в разработке и исследовании методов машинного обучения для характеристики квантовых систем, в частности для построения компактных тензорно-сетевых моделей квантовых состояний по ограниченному набору измерений. Вторая задача состоит в калибровке программируемых линейно-оптических интерферометров с учётом технологических отклонений при изготовлении и паразитных взаимных влияний (кросс-токов). Решение этой задачи важно для создания и отладки квантовых фотонных вычислительных устройств.

Степень разработанности темы можно оценить как высокую. Для линейно-оптических интерферометров разработаны универсальные архитектуры, с помощью которых можно реализовывать произвольные целевые унитарные преобразования, а также практические методы их программирования и томографической реконструкции реализуемых ими преобразований.

Классические архитектуры универсальных интерферометров и соответствующие схемы реализации унитарных преобразований были предложены Реком (Reck) [1] в 1994 году и Клементсом (Clements) [2] в 2016 году.

Для экспериментальной реконструкции преобразований линейно-оптических сетей разработано множество методов томографии, включая подход «super-stable», предложенный Лэйнгом (Laing) и О’Брайеном (O’Brien) [3] в 2012 году, метод прямой характеристики, представленный Рахими-Кешари (Rahimi-Keshari) [4] в 2013 году, а также оптимизационные методы реконструкции, например подход PhaseLift, рассмотренный в работе Суэсса (Suess) [5], опубликованной в 2020 году.

Направление программируемых фотонных процессоров и универсальных интегральных фотонных схем также активно развивается, что отражено, например, в работе Харриса (Harris) [6], опубликованной в 2018 году. В этой области рассматриваются вопросы калибровки с учётом устойчивости к параметрам изготовления и воспроизводимости результатов реализации преобразований.

Традиционные методы настройки и томографии требуют значительного числа измерений и итераций оптимизации, что стимулирует развитие альтернативных подходов. В рамках таких подходов модель устройства обучается на экспериментальных данных и затем используется для подбора управляющих параметров без дорогостоящей итеративной оптимизации непосредственно на физическом чипе.

В задаче характеризации квантовых состояний степень разработанности темы также можно оценить как высокую, однако традиционная полная томография ограничена экспоненциальным ростом числа измерений с увеличением размерности системы. Существенный теоретический вклад в понимание возможности «обучаемой» характеризации внесла работа Ааронсона (Aaronson) [7] 2007 года, посвящённая РАС-обучаемости квантовых состояний. Дальнейшее развитие получили методы тензорной томографии и классических теней, позволяющие предсказывать множество свойств квантовой системы по относительно малому числу измерений; соответствующие результаты были получены Хуангом (Huang), Куэнгом (Kueng) и Прескиллом (Preskill) [8] в 2020 году. Параллельно сформировался мощный класс структурных моделей на основе тензорных сетей, широко применяемых для компактного описания и реконструкции квантовых состояний и процессов при наличии ограниченной эффективной сложности. Таким образом, в литературе хорошо разработаны как методы построения программируемых фотонных схем и их характеризации, так и подходы к обучению компактных моделей квантовых состояний; при этом остаётся актуальной разработка конструктивных протоколов, одновременно вычислительно эффективных и совместимых с экспериментальными ограничениями, такими как шум и неполные измерения.

Целью диссертационной работы является разработка, анализ и экспериментальная проверка методов машинного обучения для (i) идентификации и программирования линейно-оптических интерферометров по экспериментальным данным и (ii) эффективной характеризации квантовых состояний на основе компактных тензорно-сетевых моделей, обучаемых по ограниченному набору локальных измерений.

В рамках поставленной цели были сформулированы следующие **задачи**:

1. Разработать архитектурно-независимый алгоритм обучения цифровой модели реконфигурируемого линейно-оптического интерферометра по обучающей выборке, содержащей пары «фазовые настройки – реконструированная матрица преобразования» (восстановленная по экспериментальным измерениям).
2. Выполнить экспериментальную проверку подхода к программированию фотонного процессора на основе обученной модели для случая малого числа мод ($N = 4$), исследовав влияние реалистичных факторов, включая кросс-токи, ограничение числа фазовых слоёв и особенности используемых измерительных процедур.
3. Разработать улучшенный алгоритм обучения модели интерферометра, использующий специально построенные признаки (линеаризацию зависимости преобразования от фазовых параметров), и исследовать его свойства, включая требования к объёму данных,

устойчивость к шуму реконструкции и области применимости (в частности, для неглубоких схем).

4. Разработать протокол обучения тензорно-сетевой модели неизвестного многокубитного квантового состояния, использующий только факторизованные однокубитные измерения, с целью предсказания вероятностей исходов будущих измерений.
5. Исследовать численную эффективность и масштабируемость тензорно-сетевого протокола, включая зависимость качества предсказаний от размера обучающей выборки и ранговых ограничений модели, и определить область применимости метода для состояний, возникающих в типичных для NISQ-устройств схемах.

Объектом исследования являются методы машинного обучения и вычислительные протоколы, применяемые для (i) идентификации, калибровки и программирования линейно-оптических интерферометров и фотонных процессоров по экспериментальным данным и (ii) эффективной характеристики квантовых состояний на основе обучаемых компактных моделей. **Предметом исследования** являются алгоритмы построения цифровых моделей интерферометров (включая архитектурно-независимое и линеаризованное обучение) и тензорно-сетевых моделей квантовых состояний, а также исследование их точности, устойчивости к экспериментальным неидеальностям (шум, ошибки реконструкции, кросс-токи, ограниченные измерения) и масштабируемости по числу мод или кубитов и объёму обучающей выборки.

Научная новизна заключается в следующих положениях:

1. Разработан архитектурно-независимый алгоритм обучения цифровой модели реконфигурируемого линейно-оптического интерферометра по обучающей выборке вида «фазовые настройки – реконструированная матрица преобразования».
2. Для указанного алгоритма получены аналитические выражения градиентов целевой функции (на основе нормы Фробениуса) по параметрам модели интерферометра, что позволяет эффективно применять градиентные методы оптимизации при обучении.
3. Предложена практическая схема проекции промежуточных оценок параметров модели на унитарное множество (через полярное разложение) в ходе обучения, обеспечивающая корректность модели при численной оптимизации.
4. Численно исследованы условия успешной идентификации модели интерферометра и выявлены ограничения «black-box» обучения при росте размерности, связанные с трудностями невыпуклой оптимизации.
5. Показано, что использование априорной информации о внутренней структуре устройства (инициализация, близкая к проектным

базисным преобразованиям) существенно расширяет область применимости метода по числу мод и повышает устойчивость обучения.

6. Исследована устойчивость обучения модели интерферометра к ошибкам реконструкции матриц преобразования (шуму томографии) и определены режимы, в которых точность данных становится лимитирующим фактором.
7. Выполнена экспериментальная реализация и проверка подхода к программированию фотонного процессора на основе обученной модели для случая $N = 4$.
8. Экспериментально исследовано влияние кросс-токов между фазовыми элементами на качество программирования и корректность обучаемой модели интерферометра.
9. Продемонстрирована работоспособность подхода при ограниченной управляемости схемы (один фазовый слой) и при использовании различных экспериментальных измерительных процедур, отличающихся от «идеализированной» полной реконструкции.
10. Предложен линейаризованный (признаковый) алгоритм обучения модели линейно-оптического интерферометра, в котором зависимость преобразования от фазовых параметров переводится в форму, допускающую обучение методами линейной алгебры.
11. Для линейаризованного подхода исследованы требования к объёму обучающих данных и влияние шума, а также выделены области применимости по глубине интерферометрической структуры.
12. Разработан вычислительно реализуемый протокол обучения тензорно-сетевой модели неизвестного многокубитного квантового состояния по данным факторизованных однокубитных измерений.
13. В рамках предложенного тензорно-сетевого протокола задача сводится к обучению модели с учителем по результатам доступных измерений и затем оцениванию вероятности исходов последующих измерений без восстановления полной матрицы плотности.
14. Проведено численное исследование точности и масштабируемости протокола, зависимости качества обучения от объёма обучающей выборки и ограничений на тензорные ранги модели. Для рассмотренных сценариев показано, что необходимый объём данных может расти в ряде случаев линейно от числа кубитов.

Теоретическая и практическая значимость определяется двумя обстоятельствами. Во-первых, работа связана с разработкой и экспериментальной проверкой новых методов программирования линейно-оптических интерферометров и характеристики квантовых состояний. Во-вторых, полученные результаты могут быть использованы при

создании и отладке квантовых вычислительных устройств, а также практической настройке линейно-оптических интерферометров.

Методология диссертационного исследования включает в себя аналитическое моделирование, численные симуляции и экспериментальную проверку в линейно-оптическом эксперименте.

В задаче программирования линейно-оптических интерферометров рассматривалась схемы с блочной архитектурой, в которой чередовались слои фазовых элементов и блоки смешивания мод. Параметры модели такого интерферометра оценивались в парадигме задачи обучения с учителем по выборке вида «фазовые параметры — реконструированная матрица преобразования».

Для подбора параметров модели применялись градиентные методы оптимизации целевой функции, основанной на норме Фробениуса. Унитарность промежуточных оценок обеспечивалась их проекцией на множество унитарных матриц с помощью полярного разложения.

Экспериментальная часть включала формирование обучающей выборки на основе результатов томографии преобразования интерферометра при различных установленных фазах и оценку качества его программирования. Для оценки использовались выбранные метрики близости между реализуемым и целевым преобразованиями.

В линеаризованном варианте модели интерферометра использовались специально построенные признаки, позволяющие свести этап обучения модели к решению задачи наименьших квадратов.

В задачах характеристики квантовых состояний методология опиралась на тензорно-сетевое моделирование и постановку задачи обучения с учителем для предсказания вероятностей исходов измерений по данным факторизованных однокубитных измерений; оптимизация параметров тензорной модели выполнялась стохастическими градиентными методами по среднеквадратичной функции потерь, а качество обучения оценивалось на независимых тестовых выборках и по метрикам согласия предсказанных и истинных вероятностей.

Положения, выносимые на защиту:

1. Обучение цифровой модели реконфигурируемого линейно-оптического интерферометра по выборке «фазовые настройки — реконструированная матрица преобразования» осуществимо без априорного знания базисных матриц для размерностей до $N = 6$ включительно.
2. Программирование фотонного процессора на основе обученной модели экспериментально реализуемо для $N = 4$ в условиях, характерных для реальной аппаратуры, включая наличие перекрёстного влияния между фазовыми элементами, наличие только одного фазового слоя и использование ограниченного набора экспериментальных данных для обучения.

3. В линеаризованном (признаковом) подходе к обучению модели интерферометра режим обучения «по всем слоям» требует объёма обучающей выборки, масштабирующегося как $O(N^L)$, где N — число оптических мод, L — глубина интерферометрической структуры.
4. В линеаризованном (признаковом) подходе режим послойного обучения приводит к линейной по числу мод масштабируемости объёма обучающей выборки для одного слоя, что обеспечивает применимость метода для интерферометров с большим числом слоёв по сравнению с режимом «по всем слоям».
5. Обучение тензорно-сетевой модели многокубитного квантового состояния по данным факторизованных однокубитных измерений позволяет предсказывать вероятности исходов новых измерений без восстановления полной матрицы плотности. В численных экспериментах зависимость минимально необходимого объёма обучающих данных для достижения заданного качества предсказания от числа кубитов хорошо аппроксимируется линейной зависимостью в рассмотренном диапазоне количества кубитов.

Достоверность эксперимента обеспечивается использованием современных оптических компонентов, приборов и измерительного оборудования. Достоверность численных симуляций опирается на применение распространённых и протестированных программных средств (компиляторов, библиотек алгоритмов). Наблюдается сходство результатов эксперимента и численных расчётов, а также согласие с аналитически ожидаемым поведением. Результаты находятся в соответствии с результатами, полученными другими авторами.

Личный вклад. Все результаты работы получены автором лично или при его непосредственном участии. Автор принимал участие в постановке задач, разработке алгоритмов, программной реализации методов, обработке результатов экспериментов, представленных в диссертационной работе.

Апробация работы. Основные результаты, представленные в диссертационной работе, докладывались на следующих международных и всероссийских конференциях:

1. HOLOEXPO 2023 20-я Международная конференция по голографии и прикладным оптическим технологиям, Сочи, Россия, 2023 год,
2. HOLOEXPO 2024 21-я Международная конференция по голографии и прикладным оптическим технологиям, Казань, Россия, 2024 год,
3. VIII International Conference on Quantum Technologies (ICQT), Москва, Россия, 2025 год,

4. НОЛОЕХРО 2025 22-я Международная конференция по голографии и прикладным оптическим технологиям, Уфа, Россия, 2025 год,
5. 7-ая Школа молодых учёных в рамках Российского форума «Микроэлектроника», Сочи, Россия, 2025 год,

а также на семинарах кафедры квантовой электроники физического факультета МГУ имени М. В. Ломоносова и семинарах Российского Квантового Центра.

Публикации. Основные результаты работы опубликованы в 4 печатных работах, в том числе в 4 статьях (общим объёмом 3.1 п.л.) в рецензируемых научных журналах, удовлетворяющих «Положению о присуждении учёных степеней в МГУ имени М. В. Ломоносова».

Основное содержание работы

Первая глава посвящена обзору теоретических и экспериментальных основ линейно-оптических интерферометров, рассматриваемых в диссертации в качестве платформы для реализации многомодовых линейно-оптических преобразований. В начале главы вводится математическое описание схем, состоящих из светоделителей и фазовых задержек. Обсуждается связь унитарных матриц, которые задают преобразование входных мод в выходные, с действием таких схем на многомодовые когерентные и фоковские состояния света.

Далее рассматриваются основные архитектуры универсальных линейно-оптических интерферометров. Описаны стандартные архитектуры универсальных интерферометров — треугольная схема Река и прямоугольная схема Клементса, реализуемые с помощью сеток интерферометров Маха–Цандера с управляемыми фазовыми задержками. Обсуждаются аналитические алгоритмы программирования интерферометров, обладающих такими архитектурами, а также различия между ними по оптической глубине и устойчивости к потерям. Отмечено, что прямоугольная архитектура Клементса обладает меньшей глубиной и поэтому предпочтительна для экспериментальной реализации.

Центральное место в этой главе занимает блочная архитектура линейно-оптических схем, основанная на чередовании фиксированных смешивающих моды преобразований и управляемых фазовых слоёв. Такая архитектура обеспечивает универсальность линейно-оптического интерферометра при отсутствии жёстких требований к точности изготовления элементов. Подчёркивается, что рассматриваемое в диссертации представление интерферометра как последовательности фиксированных многомодовых смесителей и управляемых фазовых слоёв носит общий характер, и любая архитектура линейно-оптической схемы может быть представлена в таком виде.

Отдельный раздел посвящён технологиям изготовления интегрально-оптических интерферометров. Рассмотрены два основных подхода: литография и фемтосекундная лазерная печать.

Глава завершается описанием методов томографии линейно-оптических преобразований. Для томографии когерентными состояниями описан метод прямого восстановления комплексной матрицы преобразования по измерениям выходных интенсивностей с учётом контролируемого фазового сдвига между входными состояниями. Для томографии фоковскими состояниями рассмотрен метод Super Stable Tomography (SST), основанный на счёте одиночных фотонов и двухфотонной интерференции Хонга–У–Мандела.

Первая глава формирует теоретическую и технологическую основу для дальнейшего изложения оригинальных результатов диссертации, связанных с программированием, настройкой и томографией линейно-оптических интерферометров.

Во **второй главе** рассматривается задача обучения цифровой модели реконфигурируемого линейно-оптического интерферометра по экспериментальной выборке, состоящей из пар «управляющие параметры — реализуемое унитарное преобразование». Актуальность данной задачи обусловлена тем, что в реальных фотонных интегральных схемах параметры элементов, смешивающих моды, отличаются от проектных вследствие технологических погрешностей изготовления. В результате прямое использование известных аналитических процедур программирования, разработанных для идеализированных универсальных интерферометрических архитектур [1; 2], оказывается затруднительным.

В работе используется представление интерферометра в виде чередования фиксированных блоков, смешивающих моды, и слоёв управляемых фазовых сдвигов. В этом случае программирование устройства может быть выполнено в два этапа: сначала по экспериментальным данным восстанавливается цифровая модель интерферометра, после чего целевая конфигурация управляющих параметров численно оптимизируется для обученной модели. В диссертационной работе основное внимание уделено первому этапу — задаче обучения модели по набору экспериментально восстановленных унитарных матриц.

Пусть преобразование, реализуемое интерферометром, представляется в виде

$$U = \Phi_{N+1} U_N \Phi_N \dots U_2 \Phi_2 U_1 \Phi_1, \quad (1)$$

где $U_1, \dots, U_N$ — фиксированные унитарные матрицы, описывающие модосмешивающие элементы, а $\Phi_\ell = \text{diag}(e^{i\varphi_{\ell 1}}, \dots, e^{i\varphi_{\ell N}})$ — диагональные матрицы фазовых сдвигов. Задача обучения состоит в восстановлении матриц $U_1, \dots, U_N$ по обучающей выборке $\mathcal{T} = \{(\bar{\Phi}^{(m)}, \bar{U}^{(m)})\}_{m=1}^M$, где $\bar{\Phi}^{(m)}$ соответствует заданной конфигурации управляющих фаз, а $\bar{U}^{(m)}$ —

унитарной матрице, восстановленной для реального устройства экспериментальными методами томографии [3; 4].

В качестве функционала невязки используется нормированный квадрат нормы Фробениуса

$$J_{\text{FR}}(U, \bar{U}) = \frac{1}{N} \sum_{i,j=1}^N |u_{ij} - \bar{u}_{ij}|^2. \quad (2)$$

Для данного функционала в главе получены аналитические выражения для градиентов как по элементам базисных матриц, так и по управляющим фазам. Предложенный подход является архитектурно-независимым и не требует наличия явной аналитической процедуры декомпозиции целевого унитарного преобразования в параметры конкретной схемы.

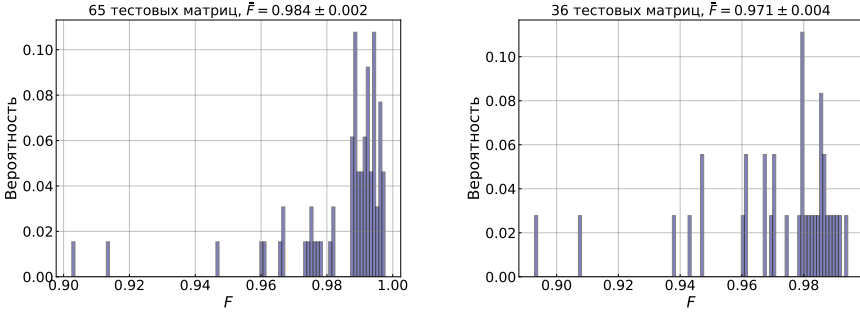
Проведено исследование в численных симуляциях предложенного алгоритма для многомодовых интерферометров различной размерности. Показано, что при случайной инициализации базисных матриц задача обучения быстро усложняется с ростом числа мод: сходимость оптимизационного процесса замедляется, а значение функционала невязки выходит на плато. Успешное обучение с высокой точностью на тестовой выборке удалось продемонстрировать только при $N < 7$. Использование априорной информации о структуре устройства, то есть близкого к проектной конфигурации начального приближения, существенно повышает эффективность обучения и позволяет успешно восстанавливать модели интерферометров большей размерности. Дополнительно исследована устойчивость метода к шумам экспериментальной реконструкции матриц, присутствующим в обучающей выборке. Установлено, что при умеренном уровне шума алгоритм работоспособен и обеспечивает приемлемую точность восстановления модели.

Экспериментальная апробация алгоритма выполнена на интегральном фотонном чипе размерности 4×4 , изготовленном методом фемтосекундной лазерной записи. В данном эксперименте рассматривались два типа обучающих данных: множество пар «фазовые сдвиги — унитарная матрица» и множество пар «управляющие токи — унитарная матрица». Во втором случае учитывалось также, что термооптические фазовращатели влияют друг на друга перекрёстно. Это приводит к тому, что фазовый сдвиг в каждом канале зависит от токов, подаваемых на все нагреватели. Для описания этого эффекта использовалась аналитическая зависимость такого вида:

$$\varphi_i(\vec{x}) = \varphi_i^{(0)} + \sum_j \alpha_{ij} x_j^2. \quad (3)$$

Результаты тестирования обученной модели представлены на рисунке 1. Показано, что предложенный метод позволяет построить цифровую

модель интерферометра как в случае, когда обучающая выборка задаётся непосредственно фазовыми сдвигами, так и в случае, когда доступны только управляющие токи.



а) Обучающая выборка $\vec{\varphi} \rightarrow U$

б) Обучающая выборка $\vec{x} \rightarrow U$

Рисунок 1 — Гистограммы распределения фиделити для обученной цифровой модели интерферометра на тестовой выборке: а) при обучении на наборе данных «фазовые сдвиги — унитарная матрица»; б) при обучении на наборе данных «управляющие токи — унитарная матрица».

Для тестовой выборки получено высокое соответствие между экспериментально восстановленными унитарными матрицами и матрицами, предсказанными обученной моделью: значение фиделити превышает 0.90, а максимум распределения расположен вблизи 0.99. Тем самым продемонстрирована возможность использования методов градиентного обучения для калибровки и моделирования реальных реконфигурируемых линейно-оптических устройств.

Полученные результаты показывают, что обучение цифровой модели по экспериментальным данным является эффективным инструментом для последующего программирования фотонных интегральных схем в условиях производственных отклонений и неполной априорной информации о параметрах устройства.

В третьей главе предлагается метод программирования реконфигурируемого линейно-оптического интерферометра, основанный на переходе от исходного пространства фазовых параметров к специальному признаковому пространству. Необходимость такого перехода обусловлена тем, что задача обучения цифровой модели интерферометра сводится к сложной невыпуклой оптимизации с большим числом параметров. Предложенный подход использует структуру линейно-оптической схемы для построения признаков, в которых зависимость элементов матрицы преобразования от фазовых переменных принимает линейный вид. Это позволяет заменить

прямую нелинейную процедуру обучения решением задачи наименьших квадратов.

Рассматривается многомодовый интерферометр, матрица преобразования которого представляется в виде чередования фазовых слоёв и фиксированных модосмешивающих блоков:

$$U = \Phi_L U_{L-1} \dots U_\ell \Phi_\ell \dots \Phi_2 U_1 \Phi_1, \quad (4)$$

где $\Phi_\ell = \text{diag}(e^{i\varphi_1^{(\ell)}}, \dots, e^{i\varphi_N^{(\ell)}})$ — диагональные матрицы фазовых сдвигов, а U_ℓ — базисные унитарные матрицы, описывающие перераспределение излучения между модами. Поэлементное разложение матрицы U показывает, что каждый элемент u_{ij} может быть представлен в виде суммы по всем оптическим путям, соединяющим входную и выходную моды. При этом множители, зависящие только от фаз, могут быть выделены в отдельные признаки, а множители, зависящие от структуры устройства, — в коэффициенты линейной модели:

$$u_{ij} = \sum_{\vec{k}} w_{i\vec{k}j} \theta_{i\vec{k}j}, \quad (5)$$

где $\theta_{i\vec{k}j}$ представляют собой произведения фазовых экспонент вдоль соответствующих путей, а коэффициенты $w_{i\vec{k}j}$ определяются параметрами базисных матриц интерферометра. Таким образом, задача восстановления модели интерферометра по экспериментальной выборке сводится к линейной регрессии в специально построенном признаковом пространстве.

Для обучения полной модели по набору экспериментальных данных используется задача наименьших квадратов

$$\|\Theta W - U\|_F^2 \rightarrow \min, \quad (6)$$

где Θ — матрица признаков, W — вектор или тензор искомых коэффициентов модели, а U — набор измеренных элементов матрицы преобразования. Получено аналитическое решение через псевдообратную матрицу Мура–Пенроуза, что позволяет избежать численной невыпуклой оптимизации при обучении. Показано, что данный алгоритм эффективен для интерферометров малой глубины, когда число признаков остаётся умеренным и полная модель может быть восстановлена по обучающей выборке разумного объёма.

Результаты численного моделирования алгоритма обучения на всех слоях показаны на рисунке 2.

Показано, что успешное обучение модели достигается при превышении порогового размера обучающей выборки. При переходе через этот порог ошибка на тестовой выборке убывает скачком, что подтверждает корректность предложенной линейной постановки задачи обучения.

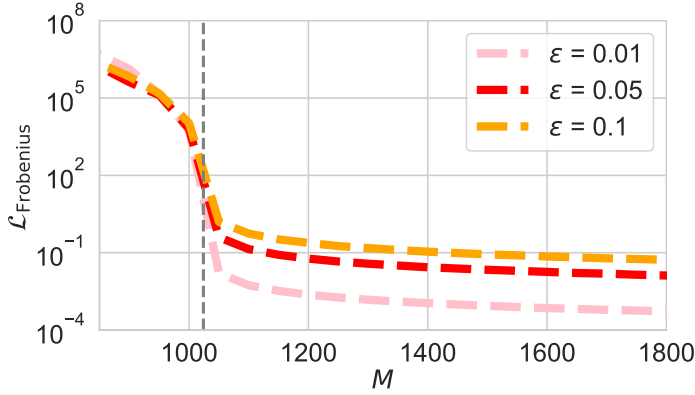


Рисунок 2 — Зависимость метрики Фробениуса $\mathcal{L}_{\text{Frobenius}}$ на тестовой выборке от размера обучающей выборки M для алгоритма обучения на всех слоях для различных уровней томографического шума. Вертикальная пунктирная показывает минимальный размер выборки, который необходим для того, чтобы модель могла с хорошей точностью работать на тестовой выборке.

При $M \gg M_{\min}$ ошибка восстановления модели убывает асимптотически как $O(M^{-1})$, что согласуется с оценками на основе метода наименьших квадратов и теории статистического оценивания.

Для интерферометров большой глубины предложен второй алгоритм, который заключается в послойном обучении. В этом случае на каждом шаге алгоритма оптимизируются только фазы одного слоя, при этом параметры остальных слоёв фиксируются. На каждом таком шаге алгоритма интерферометр описывается выражением:

$$U = U_L^{(\ell)} \Phi^{(\ell)} U_R^{(\ell)}, \quad (7)$$

где $\Phi^{(\ell)}$ — фазовый слой, который оптимизируется на текущем шаге, а $U_L^{(\ell)}$ и $U_R^{(\ell)}$ задают преобразования части интерферометра слева и справа от рассматриваемого фазового слоя. После обучения локальной линейной модели фазы данного слоя численно оптимизируются, затем процедура повторяется для следующего слоя. Такой подход эквивалентен методу попеременных наименьших квадратов и позволяет существенно сократить объём данных, необходимый для настройки: число обучающих примеров для одного слоя растёт лишь линейно с числом мод, как $O(N)$.

Результаты моделирования послойного алгоритма представлены на рисунке 3.

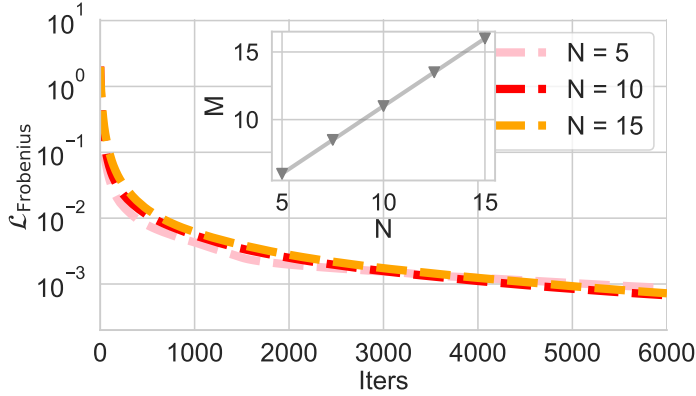


Рисунок 3 — Зависимость функции потерь Фробениуса от числа итераций при послойной настройке полнослойного линейно-оптического интерферометра для различных значений числа мод N .

Показано, что функция потерь монотонно уменьшается с числом итераций настройки и для полнослойных интерферометров устойчиво выходит на малые значения. Это свидетельствует о хорошей сходимости метода даже в случае большой глубины схемы.

С практической точки зрения предложенные методы ориентированы на использование совместно с процедурами унитарной томографии линейно-оптических сетей [4; 5]. Алгоритм обучения на всех слоях наиболее удобен для интерферометров малой глубины, где возможно однократное построение полной цифровой модели устройства. Послойный алгоритм, напротив, ориентирован на масштабируемую настройку полнослойных интерферометров и может применяться в устройствах с большим числом мод. Таким образом, в главе предложен универсальный подход к программированию линейно-оптических интерферометров, сочетающий аналитическую линейную реконструкцию модели и эффективные процедуры численной настройки.

В **четвёртой главе** излагается формализм тензорных сетей, используемый в дальнейшем для классической симуляции многокубитных квантовых систем. Необходимость привлечения данного аппарата обусловлена экспоненциальным ростом размерности гильбертова пространства с увеличением числа кубитов. Для чистого состояния N -кубитной системы

$$|\psi\rangle = \sum_{i_1, \dots, i_N=0}^1 \psi_{i_1 \dots i_N} |i_1\rangle \otimes \dots \otimes |i_N\rangle \quad (8)$$

число комплексных амплитуд $\psi_{i_1 \dots i_N}$ растёт как 2^N . В качестве основного формата представления квантовых состояний в работе используется формат тензорного поезда (tensor train, TT), предложенный в работе [9].

В TT/MPS-формате тензор амплитуд представляется в виде произведения трёхиндексных ядер (вагонов)

$$\psi_{i_1 \dots i_N} = \sum_{k_0, \dots, k_N} \psi_{k_0 i_1 k_1}^{(1)} \psi_{k_1 i_2 k_2}^{(2)} \dots \psi_{k_{N-1} i_N k_N}^{(N)}, \quad r_0 = r_N = 1, \quad (9)$$

где величины $r_1, \dots, r_{N-1}$ называются TT-рангами.

Если TT-ранги ограничены величиной χ , то число параметров в таком представлении растёт уже не экспоненциально, а как $O(N\chi^2)$, что делает возможным моделирование многих состояний, обладающих умеренной межчастичной запутанностью.

Глава носит методический характер. В ней кратко излагаются стандартные результаты, лежащие в основе TT/MPS-подхода.

Основным практическим результатом главы является разработка библиотеки **QTensor**, предназначенной для моделирования квантовых состояний и квантовых схем в формате тензорных поездов. В рамках библиотеки реализованы представление чистых квантовых состояний в TT/MPS-формате, применение одно- и двухкубитных преобразований, процедуры ортонормализации, а также вычисление скалярных произведений и локальных редуцированных матриц плотности. Для проверки корректности реализованных алгоритмов рассмотрено моделирование случайных квантовых схем, которые широко используются как тестовые задачи для классической симуляции квантовых вычислений [10].

Четвёртая глава формирует методическую и программную базу для изложения последующих результатов диссертации. В ней вводятся основные понятия и конструкции теории тензорных сетей, приводится описание применяемых стандартных алгоритмов, а также представляется разработанная автором библиотека **QTensor**. Дальше эта библиотека используется как основной инструмент для проведения численного моделирования.

В пятой главе исследуется задача обучения тензорно-сетевой модели квантового состояния на основании результатов локальных проективных измерений, выполненных над отдельными кубитами. Интерес к данной постановке обусловлен тем, что полная квантовая томография многокубитных состояний требует экспоненциального роста числа измерений [11]. По этой причине её применение оказывается непрактичным для квантовых NISQ-устройств, состоящих из нескольких десятков кубитов.

Альтернативный подход состоит в построении вычислительно эффективной модели состояния, которое принадлежит какому-то узкому классу состояний. Такая модель должна по ограниченному набору экспериментальных данных предсказывать вероятности результатов новых измерений. Рассматриваемая постановка задачи связана с парадигмой

РАС-обучения квантовых состояний [7] и теневой томографии [8; 12]. В данной работе рассматривается специальный случай: используются факторизованные локальные измерения, а модель состояния задаётся низкоранговым параметрическим представлением в тензорно-сетевом формате.

В диссертации рассматривается неизвестное смешанное состояние N -кубитной системы с оператором плотности $\hat{\rho}$. Обучающая выборка состоит из пар $\{\hat{E}_m, p_m\}$, где $\hat{E}_m$ — оператор измерения, а $p_m = \text{Tr}(\hat{E}_m \hat{\rho})$ — соответствующая вероятность положительного исхода. Рассматриваются только факторизованные проективные измерения вида

$$\hat{E} = \hat{E}_1 \otimes \cdots \otimes \hat{E}_N, \quad (10)$$

где каждый однокубитный оператор имеет вид $\hat{E}_i = |\psi_i\rangle\langle\psi_i|$. Эквивалентно, оператор измерения записывается как

$$\hat{E} = \hat{U}_1|0\rangle\langle 0|\hat{U}_1^\dagger \otimes \cdots \otimes \hat{U}_N|0\rangle\langle 0|\hat{U}_N^\dagger, \quad (11)$$

где $\hat{U}_i$ — случайные однокубитные унитарные преобразования. Такой выбор измерений соответствует экспериментально доступной схеме локальных измерений и не требует реализации сложных глобальных многокубитных операций.

В качестве обучаемой модели вводится положительный оператор

$$\hat{\sigma} = \hat{\Omega}\hat{\Omega}^\dagger, \quad (12)$$

где $\hat{\Omega}$ параметризуется в формате тензорного поезда. Такое представление соответствует классу локально-очищенных состояний (locally purified states, LPS) и автоматически гарантирует положительность модели. Предсказываемая вероятность для нового измерения $\hat{E}$ определяется выражением

$$\tilde{p} = \text{Tr}(\hat{E}\hat{\sigma}) = \text{Tr}(\hat{E}\hat{\Omega}\hat{\Omega}^\dagger). \quad (13)$$

Параметры тензорно-сетевой модели находятся из условия минимизации среднеквадратической ошибки между предсказанными и точными вероятностями на обучающей выборке:

$$\mathcal{L} = \frac{1}{M} \sum_{m=1}^M \left(\text{Tr}(\hat{E}_m \hat{\Omega}\hat{\Omega}^\dagger) - p_m \right)^2. \quad (14)$$

Таким образом, задача сводится к обучению регрессионной модели, которая по набору локальных измерений восстанавливает предсказатель вероятностей для новых, ранее не встречавшихся операторов измерения.

В численном эксперименте исследуется обучение модели на случайных смешанных состояниях, сгенерированных в формате МРО малых

рангов. Численно показано, что при ограниченном тензорном ранге состояния и модели существует режим, в котором зависимость минимального размера обучающей выборки, необходимого для достижения заданного качества предсказания, от числа кубитов хорошо аппроксимируется линейной функцией:

$$M_{\min} \sim O(N). \quad (15)$$

Эта зависимость показана на рисунке 4. Данный результат согласуется с общей идеологией РАС-обучения квантовых состояний [7] и показывает, что в рассматриваемом классе низкоранговых состояний возможно масштабируемое обучение предиктора вероятностей по данным локальных измерений.

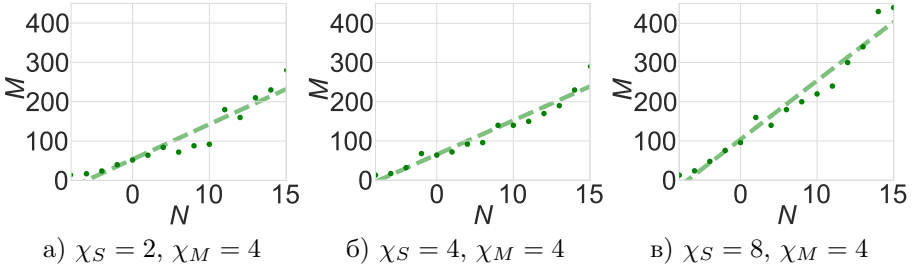


Рисунок 4 — Зависимость минимального размера обучающей выборки, необходимого для достижения заданного качества предсказания, от числа кубитов для тензорно-сетевой модели квантового состояния. Численные результаты демонстрируют линейный рост $M_{\min}$ при ограниченном ТТ-ранге.

Итак, в данной главе предложен масштабируемый алгоритм обучения тензорно-сетевой модели квантового состояния по данным локальных проективных измерений. В отличие от полной квантовой томографии, данный подход ориентирован не на восстановление полного описания состояния, а на построение эффективного предиктора вероятностей результатов последующих измерений. Показано, что для низкоранговых состояний, в том числе для состояний, характерных для NISQ-архитектур, объём обучающей выборки может расти лишь линейно с числом кубитов, а качество предсказания сохраняется на высоком уровне. Полученные результаты развивают идеи РАС-обучения квантовых состояний и теневой томографии.

В **заключении** приведены основные результаты работы, которые заключаются в следующем:

1. Рассмотрена и формализована задача программирования реконфигурируемых линейно-оптических интерферометров в общем

- (архитектурно-независимом) виде как задача идентификации параметров модели устройства по экспериментальным данным и последующего подбора фазовых настроек для реализации заданного преобразования.
2. Исследован алгоритм обучения цифровой модели интерферометра по обучающей выборке вида «настройки фаз — реконструированная матрица преобразования». Для целевой функции, основанной на норме Фробениуса, использованы аналитические выражения градиентов и проекция промежуточных оценок на унитарное множество, что обеспечивает практическую реализуемость обучения.
 3. В численных экспериментах изучены условия сходимости архитектурно-независимого обучения модели при росте числа мод. Установлено, что в режиме «чёрного ящика», без априорного знания базисных матриц, обучение устойчиво достигается для размерностей до $N = 6$ включительно, тогда как при больших N для успешной сходимости требуется использование априорной информации (в виде начального приближения, близкого к реальным или проектным базисным преобразованиям).
 4. Экспериментально продемонстрирована применимость подхода к программированию фотонного процессора на основе обученной модели для случая $N = 4$ в условиях аппаратных неидеальностей, включая наличие кросс-токов между фазовыми элементами, ограничение числа фазовых слоёв (один фазовый слой) и использование отличающихся измерительных процедур; выполнен анализ влияния этих факторов на качество программирования.
 5. Предложен и исследован линеаризованный (признаковый) подход к обучению модели линейно-оптического интерферометра, в котором зависимость элементов матрицы преобразования от фазовых параметров представляется линейной по весам в специально выбранном признаковом пространстве.
 6. Показано, что в режиме обучения «по всем слоям» (all-layers) линеаризованный подход требует объёма обучающих данных, растущего как $O(N^L)$ по числу мод N и глубине L интерферометрической структуры, что ограничивает применимость этого режима для глубоких схем.
 7. Показано, что режим «последовательного обучения» (layer-wise) в линеаризованном подходе приводит к линейной по числу мод масштабируемости требуемого объёма данных ($O(N)$ на слой/шаг), что делает метод применимым для более глубоких интерферометрических структур при сохранении практической вычислительной сложности.

8. Разработан и исследован протокол обучения тензорно-сетевой модели квантового состояния по данным факторизованных однокубитных измерений, ориентированный на предсказание вероятностей исходов новых измерений без восстановления полной матрицы плотности.
9. В численных экспериментах для тензорно- сетевого протокола показано, что минимально необходимый объём обучающих данных для достижения заданного качества предсказания демонстрирует линейное масштабирование по числу кубитов ($O(N)$) в рассмотренных сценариях.

Статьи по теме диссертации, опубликованные в рецензируемых научных изданиях, рекомендованных для защиты в диссертационном совете МГУ по специальности и отрасли наук:

1. Kuzmin S.S., Dyakonov I.V., Kulik S.P. Architecture agnostic algorithm for reconfigurable optical interferometer programming // Optics Express. — 2021. — Vol. 29, № 23. — pp. 38429–38440. — Импакт-фактор 3,3 (JIF) / 0.75 п.л. / вклад соискателя: теоретическая разработка и программная реализация алгоритма обучения модели машинного обучения линейно-оптического интерферометра, выполнение численных симуляций, EDN: STTIRC.
2. Kuzmin S.S., Mikhailova V.I., Dyakonov I.V., Straupe S.S. Learning the tensor network model of a quantum state using a few single-qubit measurements // Physical Review A. — 2024. — Vol. 109, № 5. — p. 052616. — Импакт-фактор 2,9 (JIF) / 0.55 п.л. / вклад соискателя: теоретическая разработка и программная реализация алгоритма обучения модели машинного обучения квантового состояния в формате тензорной сети, выполнение численных симуляций, EDN: YRNJQN.
3. Kuzmin S.S., Dyakonov I.V., Straupe S.S. Leveraging a feature-based model for linear optical interferometer control // Physical Review A. — 2025. — Vol. 112, № 5. — p. 053515. — Импакт-фактор 2,9 (JIF) / 0.55 п.л. / вклад соискателя: теоретическая разработка и программная реализация алгоритма обучения модели машинного обучения линейно-оптического интерферометра, выполнение численных симуляций, EDN: SILNRE.
4. Kondratyev I.V., Urusova K.N., Argenchiev A.S., Klushnikov N.S., Kuzmin S.S., Skryabin N.N., Golikov A.D., Kovalyuk V.V., Goltsman G.N., Dyakonov I.V., Straupe S.S., Kulik S.P. Effective programming of a photonic processor with complex interferometric structure // Physical Review Applied. — 2026. — Vol. 25, № 3. — p. 034072. — Импакт-фактор 4,4 (JIF) / 1.25 п.л. / вклад соискателя: теоретическая разработка и программная реализация алгоритма

Список литературы

1. Experimental realization of any discrete unitary operator / M. Reck [и др.] // Physical review letters. — 1994. — Т. 73, № 1. — С. 58.
2. Optimal design for universal multiport interferometers / W. R. Clements [и др.] // Optica. — 2016. — Т. 3, № 12. — С. 1460–1465.
3. *Laing, A.* Super-stable tomography of any linear optical device / A. Laing, J. L. O'Brien // arXiv preprint arXiv:1208.2868. — 2012.
4. Direct characterization of linear-optical networks / S. Rahimi-Keshari [и др.] // Optics express. — 2013. — Т. 21, № 11. — С. 13450–13458.
5. Rapid characterisation of linear-optical networks via PhaseLift / D. Suess [и др.] // arXiv e-prints. — 2020. — Окт. — arXiv:2010.00517. — arXiv: 2010.00517 [physics.optics].
6. Linear programmable nanophotonic processors / N. C. Harris [и др.] // Optica. — 2018. — Т. 5, № 12. — С. 1623–1631.
7. *Aaronson, S.* The learnability of quantum states / S. Aaronson // Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences. — 2007. — Т. 463, № 2088. — С. 3089–3114.
8. *Huang, H.-Y.* Predicting many properties of a quantum system from very few measurements / H.-Y. Huang, R. Kueng, J. Preskill // Nature Physics. — 2020. — Т. 16, № 10. — С. 1050–1057.
9. *Oseledets, I. V.* Tensor-Train Decomposition / I. V. Oseledets // SIAM Journal on Scientific Computing. — 2011. — Т. 33, № 5. — С. 2295–2317. — eprint: <https://doi.org/10.1137/090752286>. — URL: <https://doi.org/10.1137/090752286>.
10. Quantum supremacy using a programmable superconducting processor / F. Arute [и др.] // Nature. — 2019. — Окт. — Т. 574, № 7779. — С. 505–510. — URL: <https://doi.org/10.1038/s41586-019-1666-5>.
11. Sample-optimal tomography of quantum states / J. Haah [и др.] // Proceedings of the forty-eighth annual ACM symposium on Theory of Computing. — 2016. — С. 913–925.
12. *Aaronson, S.* Shadow tomography of quantum states / S. Aaronson // Proceedings of the 50th annual ACM SIGACT symposium on theory of computing. — 2018. — С. 325–338.

Кузьмин Сергей Сергеевич

Методы машинного обучения для характеристики квантовых состояний и
программирования линейно-оптических интерферометров

Автореф. дис. на соискание ученой степени канд. физ.-мат. наук

Подписано в печать _____._____._____. Заказ № _____

Формат 60×90/16. Усл. печ. л. 1. Тираж 100 экз.

Типография _____