Санкт-Петербургский государственный университет

Группа 23.М04-мм

Черников Артем Александрович

Применение градиентного спуска для поиска схем запутывающих преобразований в линейной квантовой оптике

Отчёт по преддипломной практике

Научный руководитель:
к.ф.-м.н., старший преподаватель кафедры системного программирования СПбГУ
С.С. Сысоев

Санкт-Петербург
2024

Оглавление

Введение

Квантовые вычислительные устройства уже долгое время вызывают интерес у научного сообщества благодаря своей способности решать некоторые задачи намного быстрее, чем классические компьютеры [9, 12], тем самым демонстрируя так называемое квантовое превосходство. Впервые идея квантовых вычислений была предложена независимо Юрием Маниным и Ричардом Фейнманом в начале 1980-х [14, 16], но исследования в этой области начались ещё раньше [11]. С тех пор и по сей день активно рассматриваются способы создания квантовых компьютеров, а также открываются и изобретаются новые квантовые алгоритмы, то есть алгоритмы, исполняемые квантовым вычислительным устройством.

Одним из первых квантовых алгоритмов, демонстрирующих квантовое превосходство, является алгоритм Дойча-Джозы [5]. Впоследствии был предложен также алгоритм Бернштейна-Вазирани [1]. Несмотря на то что их практическое применение может вызывать сомнения, само их существование бросило вызов научному сообществу искать новые алгоритмы, демонстрирующие квантовое превосходство и решающие при этом более животрепещущие задачи.

Квантовый алгоритм, решающий очень важную задачу факторизации целого числа за полиномиальное время, был открыт американским учёным Питером Шором в 1994-м году [12]. На предположении о том, что такая задача нерешаема за обозримое количество времени, основаны многие даже современные криптосистемы. Отсюда стало ясно, что с помощью квантового компьютера можно взломать любую такую систему, и это открытие вызвало сильный интерес к квантовым компьютерам.

Спустя время в 1996-м году американский математик Лов Гровер изобрёл алгоритм поиска в неотсортированной базе данных [9], имеющий квадратичное ускорение по сравнению с лучшими известными классическими алгоритмами, решающими эту задачу. Алгоритм Гровера может быть использован для решения широкого спектра задач, в

частности, NP-полных задач.

Впоследствии американским физиком-теоретиком Дэвидом П. ДиВинченцо были сформулированы критерии [4], которым должно соответствовать вычислительное устройство для того, чтобы по праву называться квантовым компьютером. Было предложено множество архитектур квантовых компьютеров: основанные на явлении ядерного магнитного резонанса; использующие электроны, запертые в квантовых точках; основанные на ядерных спинах идентичных молекул; использующие запутанные фотоны.

Идея создать квантовый компьютер на фотонах вызывает особый интерес благодаря тому, что практически все критерии ДиВинченцо выполнимы в этой архитектуре без особых усилий. Единственная трудность остаётся в осуществлении преобразований, запутывающих пару фотонов. Исследования в этой области на текущий момент привели к некоторым любопытным результатам [7, 8, 10], однако, к сожалению, все они не масштабируемы, и задача нахождения запутывающего преобразования остаётся открытой.

В связи с этим возникло предположение о том, что для поиска таких преобразований стоит использовать техники компьютерного поиска. Одним из классических таких подходов является градиентный спуск.

1. Постановка задачи

Целью данной работы является реализация алгоритма поиска оптимальных схем запутывающих преобразований с корректировкой состояния в линейной квантовой оптике.

Для достижения данной цели были поставлены следующие задачи.

- Реализовать алгоритм поиска с помощью градиентного спуска.
- Подобрать оптимальные гиперпараметры поиска.
- Проанализировать найденные схемы и сделать вывод.

2. Описание решения

В данном разделе представлено описание контекста решаемой задачи, технических подробностей и реализации решения. Код доступен в репозитории[1], размещённом на веб-сервисе GitHub.

2.1. Математическая модель

Согласно квантовой теории поля, фотон, как и некоторые другие элементарные частицы, способен пребывать в нескольких состояниях одновременно до тех пор, пока неопределённость относительно его состояния нисколько не влияет на окружающую его вселенную. Такое сложное состояние называется квантовой суперпозицией.

Как только состояние окружающей вселенной начинает зависеть от состояния фотона (например, состояние фотона детектируется датчиком), с точки зрения её жителей неопределенность состояния фотона исчезает. При этом фотон переходит из суперпозиции в какое-нибудь одно из возможных состояний с некоторой определённой вероятностью. Этот процесс называется коллапсом волновой функции частицы, или измерением состояния частицы.

В описанном выше примере фотон и его окружение до измерения представляли собой две независимые системы. После измерения одна система стала зависеть от другой, поэтому их теперь невозможно рассматривать по отдельности. Фотон и его окружение теперь являют собой единую неразделимую систему. С этой точки зрения коллапса волновой функции после измерения не произошло, вместо этого сама вселенная стала пребывать сразу в нескольких состояниях, причем жители вселенной из одного состояния никак не могут провзаимодействовать с ними же из другого состояния. Поэтому приверженцы этой теории, в парадигме которой в данной работе и рассматриваются физические явления, называют такии версии вселенных параллельными.

Аналогичным образом данные рассуждения переносимы с системы

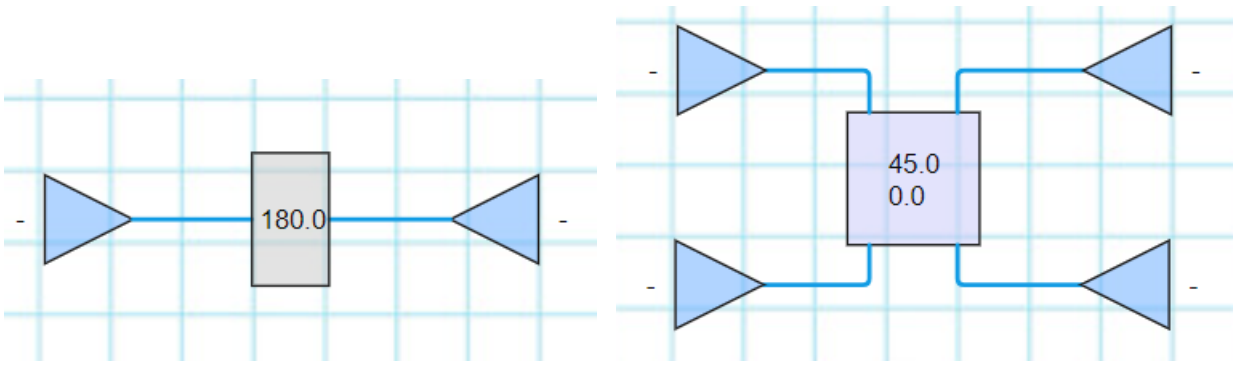
[1]Репозиторий проекта — https://github.com/sysoevss/galopy (дата обращения: 08.01.2024), пользователь artemgl

“фотон-окружение” на систему “фотон-фотон”. Если состояние двух фотонов не представимо в виде двух независимых систем (по одной для каждого фотона), то эти фотоны находятся в так называемом запутанном состоянии, в котором состояние одного фотона зависит от состояния второго. Существенным отличием данного примера от предыдущего является процесс запутывания, — на практике перевести два фотона из незапутанного состояния в запутанное оказывается нетривиальной задачей.

Квантовый алгоритм в достаточно широком смысле представляет собой некоторое обратимое преобразование над системой, способной находиться в N различных состояниях. Для решения содержательных задач зачастую число N должно быть катастрофически большим, поэтому часто рассматриваются системы, состоящие из n подсистем, называемых кубитами, каждая из которых способна находиться всего в двух состояниях — $|0\rangle$ или $|1\rangle$. Тогда общее количество состояний зависит экспоненциально от количества простых подсистем (кубитов) и равно 2^n. Доказано [3, 6], что в такой архитектуре произвольное преобразование над всей системой представимо в виде последовательности одно- и двухкубитных преобразований (гейтов), причём среди двухкубитных достаточно наличие всего одного запутывающего гейта.

В данной работе рассматривается архитектура квантового компьютера на фотонах, называемая протоколом KLM [8]. Каждый кубит кодируется ровно одним фотоном и парой пространственных мод, в которых может находиться этот фотон. Если он находится в первой моде, то состояние кубита считается равным $|0\rangle$, если во второй — $|1\rangle$.

Однокубитные преобразования в протоколе KLM осуществимы с помощью таких оптических элементов как фазовая пластина и светоделитель. Фазовая пластина имеет единственную входную и выходную моду и параметризуется всего одним углом, отвечающим за сдвиг фазы волновой функции проходящего через эту пластину фотона (рис. 1а). Светоделитель имеет два входных и два выходных плеча, что позволяет обеспечить взаимодействие двух мод между собой. Его принято параметризовать парой углов, один из которых отвечает за коэффици-



(a) Фазовая пластина. (b) Светоделитель.

Рис. 1: Схематичные изображения оптических элементов, используемых в квантовой линейной оптике. Треугольниками обозначены источники и детекторы фотонов, линиями — моды, в которых могут находиться фотоны. Параметры оптических элементов (углы) указываются в градусах.

ент пропускания, а другой — за фазовый сдвиг после отражения или пропускания фотона. Так, если первый угол равен нулю, то светоделитель вырождается в идеальное зеркало, что эквивалетно отсутствию каких-либо оптических элементов; если же он равен 90°, то светоделитель наоборот пропускает все проходящие через него лучи, тем самым меняя фотоны в двух модах местами (рис. 1b).

Двухкубитный запутывающий гейт в протоколе KLM неосуществим при помощи указанных выше оптических элементов. Однако исследования в этой области на данный момент пришли к тому, что с использованием дополнительных фотонов запутывающее преобразование становится осуществимо [7, 8]. Дополнительные фотоны взаимодействуют посредством оптических элементов с сигнальными (то есть теми, над которыми требуется произвести преобразование), после чего на дополнительных детекторах ожидается измерение определённого паттерна (к примеру, два фотона на первом датчике, один на втором и ноль на третьем). Считается, что гейт сработал правильно, тогда и только тогда, когда заданный паттерн оказался на датчиках в результате измерения. Измерение на дополнительных модах называется оповещением о работе гейта, а удачное измерение — правильным или верным оповещением. При этом заданных паттернов, сигнализирующих об успешном сраба-

тывании схемы, может быть несколько. Также измерение паттерна происходит не одновременно на всех модах, а последовательно на каждой моде, причем после очередного измерения к оставшемуся состоянию применяется ещё одно аналогичное преобразование, зависящее от измеренного подпаттерна.

Неотъемлемой частью гейтов с вышеописанной логикой работы является ненулевая вероятность неверного оповещения, что мешает использовать их на практике, поскольку для работы всего алгоритма, состоящего из сотен или тысяч таких гейтов, требуется, чтобы сработали все гейты до единого, вероятность чего оказывается крайне мала. Максимально известная на данный момент вероятность правильного оповещения одного гейта составляет $2/27$ [7].

Рассмотрим данную архитектуру подробнее. На вход гейта подаётся пара фотонов в некотором двухкубитном состоянии и несколько дополнительных фотонов, которые запускаются в дополнительные моды. Состояние, в котором находятся дополнительные фотоны, задано заранее и не зависит от состояния сигнальных фотонов. Сама схема представляет собой некоторым образом расставленные на всех модах оптические элементы с заданными углами. Некоторые из этих элементов работают при условии измерения определённого количества фотонов на датчике, к которому они привязаны. Также к схеме прилагается несколько верных оповещений, ожидаемых на дополнительных модах. На выходе схемы в сигнальных модах ожидается пара фотонов в новом двухкубитном состоянии, получившемся в результате запутывающего преобразования при верном оповещении.

Итак, запутывающее преобразование в линейной квантовой оптике задаётся следующими параметрами.

- Количество дополнительных мод и состояние дополнительных фотонов, подаваемое на эти моды.
- Расстановка светоделителей и фазовых пластин на модах (далее — топология преобразования).
- Углы (параметры) расставленных светоделителей и фазовых пла-

стин.

- Измерения, считаемые за правильные оповещения.

Топология схемы может быть представлена как набор оптических элементов, поставленных последовательно друг за другом, где для каждого из них указаны моды, на которых они действуют (один для фазовой пластины и два для светоделителя). Однако такое представление оказывается неудобным для использования на практике, а также избыточным. Эквивалентными преобразованиями любая схема может быть приведена к виду последовательности лишь светоделителей, после которых следуют фазовые пластины по одной на каждую моду. В связи с этим для удобства представления топологии схемы рассматривается только расстановка светоделителей, а фазовые пластины в конце преобразования всегда подразумеваются.

Вышеперечисленных параметров достаточно для однозначного определения оптической схемы, среди которых требуется найти такую, которая реализует запутывающий гейт с максимально возможной вероятностью правильного оповещения. Для решения этой задачи был применён градиентный спуск. После этого были проанализированы результаты.

2.2. Расчёт верности

Верность преобразования — это мера, отражающая, насколько одно преобразование похоже на другое. Во время поиска гейта верность требуется для оценки того, насколько точно имеющаяся схема реализует заданное искомое преобразование.

Способы задания верности гейтов основаны на верности, определённой на отдельных состояниях. Мерой схожести (верностью) двух состояний между собой принято считать квадрат их скалярного произведения.

$$F(|\varphi\rangle, |\psi\rangle) = |\langle\varphi|\psi\rangle|^2$$

Эта мера может быть перенесена на гейты с помощью идеи, описан-

ной в статье [13] и заключающейся в подсчёте средней верности между ожидаемым и полученным состоянием по всему пространству. Формула для её вычисления выглядит следующим образом.

$$F(U,V) = \int\limits_{S^{2n-1}} |\langle\psi|V^\dagger U|\psi\rangle|^2 dV =$$

$$= \frac{1}{n(n+1)}\Big(Tr(MM^\dagger) + |Tr(M)|^2\Big), \text{ где } M = V^\dagger U$$

Записью $M^\dagger$ обозначена эрмитово сопряжённая к M матрица; число n — количество столбцов или строк квадратной матрицы M. Из вида формулы можно заметить, что основная её вычислительная сложность заключается в двух матричных произведениях.

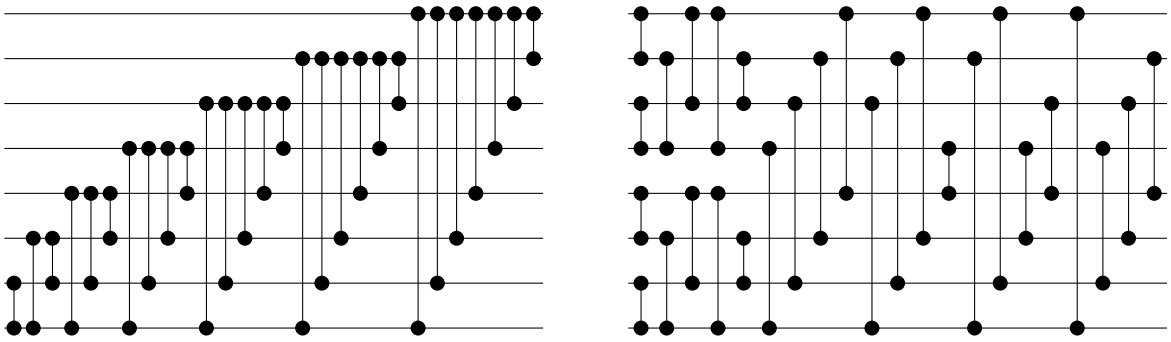
2.3. Расчёт оповещения

Состояние на выходе схемы описывает полную суперпозицию расположений фотонов, получившуюся в результате их прохождения через оптические элементы. Для того, чтобы расчитать преобразование, реализуемое схемой, и вероятность срабатывания гейта, нужно вычислить состояние в сигнальных модах, оказавшееся в результате коллапса волновой функции при верном оповещении. Это происходит путём редуцирования всего пространства на соответствующее подпространство.

Аналогично расчёту верности, описанному в разделе 2.2, было принято решение вычислять среднюю вероятность по всему пространству. Формула вероятности верного оповещения принимает следующий вид (средний скалярный квадрат вектора, полученного после преобразования).

$$P(U) = \int\limits_{S^{2n-1}} \langle\psi|U^\dagger U|\psi\rangle dV = \frac{1}{n} Tr(U^\dagger U)$$

Поскольку от произведения матриц $U^\dagger$ и U сразу берётся след, полностью считать это произведение не нужно. Для вычисления следа достаточно просуммировать все элементы матрицы, полученной в результате поэлементного умножения $U^\dagger$ на U, что и было использовано в



(a) Устойчивая топология (b) Параллельная топология

Рис. 2: Рассматриваемые универсальные топологии схем для восьми мод. Горизонтальными линиями обозначены моды. Каждый вертикальный отрезок представляет собой светоделитель, действующий на модах, отмеченных жирными точками.

реализации.

2.4. Градиентный спуск

Условия применимости градиентного спуска заключаются в дифференцируемости функции потерь, из чего возникает требование наличия только вещественных параметров. Препятствием к применению этого метода выступает то, что многие параметры, задающие оптическую схему, целочисленны.

Однако это препятствие удалось обойти путём избавления от таких параметров как топология схемы и верное оповещение. Оставшиеся целочисленные параметры были вынесены в так называемые гиперпараметры, то есть настраиваемые вручную, перед запуском алгоритма поиска.

Существуют топологии схем, в которых возможно задать любое унитарное преобразование, меняя только углы закреплённых на своих местах оптических элементов [2, 15]. Назовём такие топологии универсальными. Таким образом, если зафиксировать некоторую универсальную топологию схемы, пространство поиска нисколько не сузится. В качестве универсальных топологий рассматривались представленные на рис. 2.

После нескольких запусков функция потерь была подобрана сле-

дующим образом. Аргументы функции p и f — соответственно массив вероятностей верного оповещения и массив верностей преобразования. Под индексом i в каждом массиве находится информация об i-том верном оповещении. Параметры f_min и p_min означают соответственно минимальные приемлемые верность и вероятность работы гейта. Параметр n_measurements означает количество допустимых верных оповещений.

```
def loss(p, f):
  loss_value = 0
  for i in range(n_measurements):
    fidelity = f[i]
    probability = p[i]
    if fidelity > f_min:
      loss_value += 1 + probability
    else:
      loss_value += fidelity
  return loss_value
```

Таким образом, в процессе работы алгоритма в первую очередь настраиваются наиболее приемлемые верности для всех оповещений, после чего параметры обновляются в направлении повышения вероятностей всех оповещений. Также можно заметить, что алгоритм максимизирует функцию потерь, в отличие от классического подхода, в котором она минимизируется. Концептуально это изменение ни на что не влияет.

3. Результаты

В данном разделе представлены результаты применения градиентного спуска и сделанные на их основе выводы. Все эксперименты проводились в среде выполнения Google Colab[2] с поддержкой графического процессора.

В качестве запутывающего гейта был выбран CZ[3], поскольку он является одним из стандартных двухкубитных гейтов. Также он изменяет фазу волновой функции пары фотонов, не меняя друг с другом сами состояния, что проще осуществимо с помощью линейной оптики. На рис. 3 изображена лучшая известная на данный момент схема такого преобразования, описанная в статье Эмануэля Книлла [7].

Поиск гейта производился в двух архитектурах, перечисленных ниже.

1. На детекторах ожидается одно из нескольких заданных верных оповещений. В зависимости от измеренного паттерна к оставшемуся состоянию применяется ещё одно корректирующее преобразование.

2. На детекторах ожидается два оповещения. При первом схема срабатывает как CZ, при втором — как тождественное преобразование. Если датчики измерили первый паттерн, состояние в сигнальных модах отправляется для дальнейших вычислений; если же второй, к полученному состоянию предпринимается ещё одна попытка применить этот же гейт. Таким образом, вероятность срабатывания теоретически может быть увеличена.

Также был произведён поиск схемы, генерирующей состояние Белла.

[2]Стартовая страница с описанием среды выполнения Google Colab — https://colab.research.google.com (дата обращения: 08.01.2024)

[3]Статья про квантовые гейты, в том числе гейт CZ (Controlled Z) — https://en.wikipedia.org/wiki/Quantum_logic_gate (дата обращения: 08.01.2024)

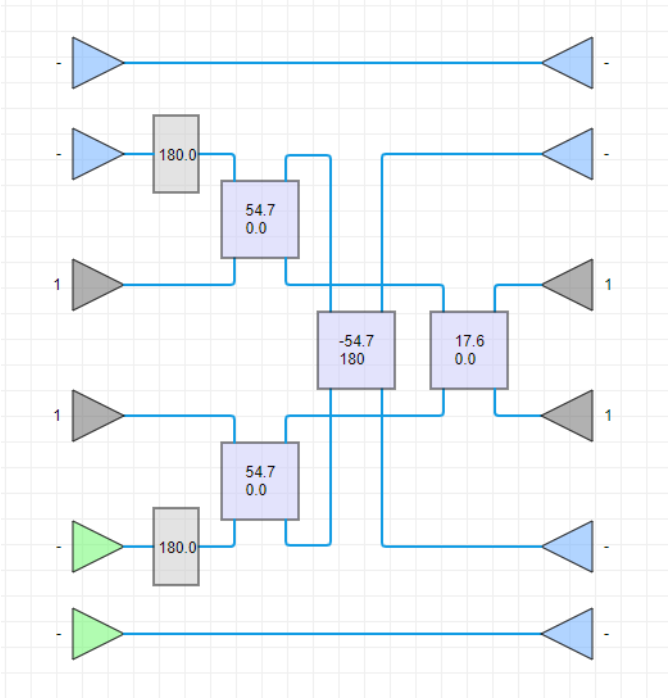


Рис. 3: Схема гейта CZ с вероятностью верного оповещения 2/27. Парой голубых треугольников в левой части схемы обозначены моды первого кубита, парой зелёных — моды второго. Серыми треугольниками обозначены входы и выходы вспомогательных мод. Число около каждого из треугольников слева означает количество фотонов, подаваемое на данную моду. Число справа означает количество фотонов в данной моде, которое ожидается измерить для верного оповещения.

3.1. Гейт CZ

Запуск градиентного спуска производился по 5 раз для каждой топологии из приведённых на рисунке 2 и описанных в разделе 2.4 со следующими параметрами.

- Количество градиентных шагов: 1000.
- Приемлемая вероятность верного оповещения: 2/27.
- Количество вспомогательных мод: 4.
- Количество вспомогательных фотонов: 2.

В обоих рассматриваемых архитектурах не нашлось гейта, который срабатывает с вероятностью лучше, чем у уже известных гейтов (2/27).

3.2. Генерация состояния Белла

Запуск градиентного спуска производился по 5 раз для каждой топологии со следующими параметрами.

- Количество градиентных шагов: 1000.
- Приемлемая вероятность верного оповещения: 2/27.
- Количество вспомогательных мод: 4.
- Количество вспомогательных фотонов: 2.

Поиск производился в обычной архитектуре без корректировок состояний. В результате поиска не нашлось генерации с вероятностью большей 2/27.

3.3. Выводы

В результате анализа экспериментов были сделаны следующие выводы.

- Увеличение количества верных оповещений не даёт выигрыш по вероятности даже при использовании корректировки состояния.
- Среди гейтов-гибридов, которые либо успешно срабатывают, либо не портят состояние сигнальных фотонов, вероятность срабатывания такая же, как и среди обычных гейтов.
- Среди гейтов, генерирующих состояние Белла, вероятность срабатывания не отличается от вероятности гейта CZ.

Таким образом, согласно полученным результатам, задача генерации состояния Белла в квантовой линейной оптике эквивалентна задаче применения гейта CZ к некоторому состоянию. Это утверждение контринтуитивно, поскольку сгенерировать запутанную пару фотонов кажется проще, чем запутать произвольную существующую. Поэтому этот вывод можно принять к рассмотрению для составления лучшего понимания устройства оптических гейтов.

Ещё один контринтуитивный результат заключается в том, что после расширения пространства поиска вероятность срабатывания всё равно не была улучшена. С одной стороны, это может означать, что расмотренное расширение излишне, с другой — наоборот, недостаточно велико. Две рассмотренные архитектуры являются частными случаями более общей архитектуры гейта с корректировками после измерения в каждой моде. В данной работе не производился поиск схем в этой обобщённой архитектуре, но эта задача включена в дальнейшие планы.

Заключение

В ходе выполнения данной работы были достигнуты следующие результаты.

- Реализованы два вида поиска в двух разных архитектурах.
- Применены реализованные методы для поиска запутывающего преобразования, а также для генерации состояния Белла.
- Проанализированы результаты, сделаны выводы и сформулирован план дальнейших работ.

Код доступен в репозитории[4], размещённом на веб-сервисе GitHub. В дальнейшие планы включены следующие задачи:

- Реализовать алгоритм поиска в обобщённой архитектуре.
- Проанализировать результаты поиска и сделать выводы.

[4]Репозиторий проекта — https://github.com/sysoevss/galopy (дата обращения: 08.01.2024), пользователь artemgl

Список литературы

[1] Bernstein E. Vazirani U. Quantum complexity theory. — 1993. — Access mode: https://dl.acm.org/doi/pdf/10.1145/167088.167097 (online; accessed: 15.05.2023).

[2] Clements W.R. et al. Optimal design for universal multiport interferometers. — 2016. — Access mode: https://opg.optica.org/optica/fulltext.cfm?uri=optica-3-12-1460 (online; accessed: 15.05.2023).

[3] D.P. DiVincenzo. Two-bit gates are universal for quantum computation. — 1995. — Access mode: https://arxiv.org/pdf/cond-mat/9407022.pdf (online; accessed: 15.05.2023).

[4] D.P. DiVincenzo. The physical implementation of quantum computation. — 2000. — Access mode: https://arxiv.org/pdf/quant-ph/0002077.pdf (online; accessed: 15.05.2023).

[5] Deutsch D. Jozsa R. Rapid solution of problems by quantum computation. — 1992. — Access mode: https://www.isical.ac.in/~rcbose/internship/lectures2016/rt08deutschjozsa.pdf (online; accessed: 15.05.2023).

[6] Deutsch D.E. Barenco A. Ekert A. Universality in quantum computation. — 1995. — Access mode: https://arxiv.org/pdf/quant-ph/9505018.pdf (online; accessed: 15.05.2023).

[7] E. Knill. Quantum gates using linear optics and postselection. — 2002. — Access mode: https://journals.aps.org/pra/abstract/10.1103/PhysRevA.66.052306 (online; accessed: 15.05.2023).

[8] Knill E. Laflamme R. Milburn G.J. A scheme for efficient quantum computation with linear optics. — 2001. — Access mode: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=7955e6b51070395b97b70ddb8d7e2ae797528a81 (online; accessed: 15.05.2023).

[9] L.K. Grover. A fast quantum mechanical algorithm for database search. — 1996. — Access mode: https://dl.acm.org/doi/pdf/10.1145/237814.237866 (online; accessed: 15.05.2023).

[10] O'Brien J.L. et al. Demonstration of an all-optical quantum controlled-NOT gate. — 2003. — Access mode: https://arxiv.org/pdf/quant-ph/0403062.pdf (online; accessed: 15.05.2023).

[11] P. Benioff. The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines. — 1980. — Access mode: https://link.springer.com/article/10.1007/BF01011339 (online; accessed: 15.05.2023).

[12] P.W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. — 1994. — Access mode: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=2273d9829cdf7fc9d3be3cbecb961c7a6e4a34ea (online; accessed: 15.05.2023).

[13] Pedersen L.H. Møller N.M. Mølmer K. Fidelity of quantum operations. — 2007. — Access mode: https://arxiv.org/pdf/quant-ph/0701138.pdf (online; accessed: 15.05.2023).

[14] R.P. Feynman. Simulating physics with computers. — 1982.

[15] Reck M. et al. Experimental realization of any discrete unitary operator. — 1994. — Access mode: https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.73.58 (online; accessed: 15.05.2023).

[16] Ю.И. Манин. Вычислимое и невычислимое. — 1980.