Письма в ЖЭТФ, том 122, вып. 3, с. 177 – 183

Статистические модели барреновских плато и антиконцентрация наблюдаемых Паули

Н. А. Немков[1)]

Национальный исследовательский технологический университет "МИСиС", 119049 Москва, Россия

Поступила в редакцию 19 мая 2025 г.
После переработки 16 июня 2025 г.
Принята к публикации 17 июня 2025 г.

Мы вводим статистические модели для каждого из трех основных источников барреновских плато: нелокальность наблюдаемой, запутанность начального состояния и экспрессивность квантовой цепи. В частности, нелокальные наблюдаемые моделируются случайными операторами Паули, что приводит к барреновским плато с вероятностью, экспоненциально близкой к единице. Эти модели являются дополнением к традиционным детерминированным подходам и зачастую проще в анализе. В рамках предложенного формализма мы показываем, что в режиме барреновского плато любые два оператора Паули антиконцентрированы с высокой вероятностью в следующем смысле: хотя каждый из них локализован в экспоненциально малом подпространстве параметров, соответствующие области по существу независимы, так что их пересечение экспоненциально меньше, чем каждое из подпространств по отдельности. Это наблюдение важно для понимания структуры квантовых ландшафтов с барреновскими плато и подходов к их оптимизации, включая стратегии предварительной инициализации.

DOI: 10.31857/S0370274X25080103, EDN: WKSPQZ

Введение. Вариационные квантовые алгоритмы (ВКА) [1] и модели квантового машинного обучения (КМО) [2, 3] являются обобщением классической оптимизации и машинного обучения и привлекли значительное внимание в последние годы. Первоначально ожидалось, что ВКА будут полезными как простые "черные ящики" для различных задач, однако они оказались сопряжены со множеством трудностей. Одной из ключевых проблем является трудность обучения, в основном связанная с плохими локальными минимумами [4–6] и экспоненциально плоскими областями, называемыми барреновскими плато (БП) [7, 8].

В данной работе мы сосредотачиваемся на свойствах квантовых ландшафтов в режиме БП. Когда функция потерь обладает барреновским плато, стандартные методы оптимизации, как градиентные, так и безградиентные, не способны выйти из плоской области и найти хорошие решения [9]. Следует отметить, что ВКА без БП, как правило, могут быть эффективно симулированы на классических компьютерах [10, 11]. В качестве потенциального решения этой проблемы было предложено производить предварительную инициализацию в областях вне БП [12–16]. Однако этот подход был поставлен под сомнение в [17], где утверждалось, что большинство конфигураций параметров вне области БП все равно являются тривиальными, поскольку, как правило, только один член в функции потерь не экспоненциально мал. Это наблюдение уточняет структуру квантовых ландшафтов с БП и накладывает дополнительные ограничения на стратегии инициализации. В то же время аргументы в [17] основывались на ряде формальных предположений, которые трудно проверять в рамках стандартных моделей БП. В настоящей статье мы предлагаем количественную меру независимости разных слагаемых в функции потерь через простую корреляционную функцию. Что более важно, мы вводим новые статистические модели БП. Эти модели позволяют сделать аргументацию из [17] строгой и количественной. Мы также полагаем, что наше статистическое описание является полезным дополнением к традиционным моделям БП и может представлять самостоятельный интерес.

ВКА. ВКА определяется начальным состоянием ρ, наблюдаемой $\mathcal{O}$ и параметризованной квантовой цепью (ПКЦ) $U(\boldsymbol{\phi})$. Соответствующая функция потерь имеет вид:

$$L_{\rho,\mathcal{O}}(\boldsymbol{\phi}) = \operatorname{Tr} \rho\, U^{\dagger}(\boldsymbol{\phi})\mathcal{O}U(\boldsymbol{\phi})\,. \tag{1}$$

В дальнейшем мы иногда будем опускать ρ или $\mathcal{O}$ в обозначениях, когда они ясны из контекста.

[1)]e-mail: nnemkov@gmail.com

Функцию потерь удобно характеризовать ее средним и дисперсией по пространству параметров:

$$\mathbb{E}_{\boldsymbol{\phi}}[L_{\rho,\mathcal{O}}(\boldsymbol{\phi})] = \int d\boldsymbol{\phi}\, L_{\rho,\mathcal{O}}(\boldsymbol{\phi}) \ , \tag{2}$$

$$\mathbb{V}_{\boldsymbol{\phi}}[L_{\rho,\mathcal{O}}(\boldsymbol{\phi})] = \mathbb{E}_{\boldsymbol{\phi}}[L^2_{\rho,\mathcal{O}}(\boldsymbol{\phi})] - (\mathbb{E}_{\boldsymbol{\phi}}[L_{\rho,\mathcal{O}}(\boldsymbol{\phi})])^2 \ . \tag{3}$$

Считается, что в ВКА возникает барреновское плато, если дисперсия функции потерь экспоненциально мала по числу кубитов:

$$\mathbb{V}_{\boldsymbol{\phi}}[L_{\rho,\mathcal{O}}(\boldsymbol{\phi})] = O(2^{-n}) \ . \tag{4}$$

Чтобы это определение было корректным, необходимо ограничить рост нормы наблюдаемой $\mathcal{O}$ при увеличении n. Мы будем использовать условие $\mathrm{Tr}\,\mathcal{O}^2 \leq 2^n$, которое, например, выполняется, если $\mathcal{O}$ – это линейная комбинация строк Паули $\mathcal{O} = \sum_i c_i P_i$ с $\sum_i |c_i|^2 \leq 1$.

Наличие БП означает, что функция потерь может существенно отклоняться от своего среднего значения только в экспоненциально малом объеме пространства параметров, т.е. она сильно сконцентрирована. Выделяют три основные причины возникновения БП: экспрессивность квантовой цепи, нелокальность наблюдаемой и высокая запутанность начального состояния. Также известно, что квантовый шум может приводить к БП. Однако в случае достаточно сильного шума функция потерь оказывается экспоненциально плоской во всем пространстве параметров. Индуцированные шумом БП являются тривиальными в контексте настоящей работы и далее не будут рассматриваться.

Мы сосредоточим внимание на специальном классе ВКА, которые мы называем ВКА с операторами Клиффорда и вращениями Паули (КПР). ПКЦ $U(\boldsymbol{\phi})$ таких ВКА содержит только постоянные операторы Клиффорда и параметризованные вращения Паули $e^{-\frac{i\phi_k P_k}{2}}$. Начальное состояние ρ – стабилизаторное, а наблюдаемая $\mathcal{O}$ – это линейная комбинация $\mathrm{poly}(n)$ наблюдаемых Паули. Кроме того, мы предполагаем отсутствие коррелированных параметров, т.е. все $\phi_k \in \boldsymbol{\phi}$ являются независимыми. За исключением последнего условия, эти требования едва ли ограничительны, поскольку большинство широко изучаемых ВКА им удовлетворяют. Отсутствие коррелированных параметров также характерно для многих случаев (но не всегда, как, например, в алгоритме QAOA [18]).

Для ВКА типа КПР среднее любой функции потерь с одиночной наблюдаемой Паули $L_{\rho,P}$ либо равно нулю, либо является постоянным (в последнем случае ее можно отбросить). Более того, дисперсия такой функции потерь может быть выражена как среднее по конечному множеству точек $\boldsymbol{\phi}_c$ [17, 19]

$$\mathbb{V}_{\boldsymbol{\phi}}[L_{\rho,P}(\boldsymbol{\phi})] = \mathbb{E}_{\boldsymbol{\phi}_c}[L^2_{\rho,P}(\boldsymbol{\phi}_c)] = \\ = \frac{1}{|\boldsymbol{\phi}_c|}\sum_{\boldsymbol{\phi}_c}\left(\mathrm{Tr}\,\rho\, U^\dagger(\boldsymbol{\phi}_c) P U(\boldsymbol{\phi}_c)\right)^2 . \tag{5}$$

Здесь $\boldsymbol{\phi}_c$ – множество таких точек, что каждый параметр ϕ_k кратен $\pi/2$, т.е. $\boldsymbol{\phi}_c = \{0, \pi/2, \pi, 3\pi/2\}^{|\boldsymbol{\phi}|}$, где $|\boldsymbol{\phi}|$ – общее число параметров в цепи. Всего таких точек $|\boldsymbol{\phi}_c| = 4^{|\boldsymbol{\phi}|}$. Мы будем называть их точками Клиффорда, поскольку при $\boldsymbol{\phi} = \boldsymbol{\phi}_c$ оператор $U(\boldsymbol{\phi}_c)$ является оператором Клиффорда. Подчеркнем, что соотношение (5) выполняется для любого ВКА типа КПР и не требует дополнительных предположений вроде принадлежности $U(\boldsymbol{\phi})$ к 2-дизайну.

Статистическая модель барреновских плато. Описанные выше источники БП качественно характеризуют причины возникновения БП, но не всегда являются точными. Например, чтобы строго доказать связь между нелокальностью и БП, необходимо определить, что именно считается нелокальной наблюдаемой, а также наложить дополнительные ограничения на квантовую цепь, например, потребовать, чтобы она была локальным 2-дизайном [20]. Здесь мы предлагаем иной, статистический подход к описанию БП. Мы полагаем, что он отражает те же качественные свойства, что и стандартные модели, но имеет иные границы применимости.

Предположим, что начальное состояние ρ и наблюдаемая $\mathcal{O}$ выбираются случайно из некоторых распределений, и рассмотрим дисперсию функции потерь, усредненную по ρ и $\mathcal{O}$. Если это среднее экспоненциально мало

$$\mathbb{E}_\rho \mathbb{E}_{\mathcal{O}}\left[\mathbb{V}_{\boldsymbol{\phi}}[L_{\rho,\mathcal{O}}(\boldsymbol{\phi})]\right] = O(2^{-n}), \tag{6}$$

то для типичного выбора ρ и $\mathcal{O}$ соответствующая функция потерь будет иметь БП с вероятностью, экспоненциально близкой к единице. В то же время экспоненциально малая доля таких функций может не иметь БП.

Нелокальность. Рассмотрим явный пример. Пусть дана ВКА типа КПР с наблюдаемой в виде одного оператора Паули $\mathcal{O} = P$, где P – случайный n-кубитный оператор Паули. Тогда:

$$\mathbb{E}_P\left[\mathbb{V}_{\boldsymbol{\phi}}[L_P(\boldsymbol{\phi})]\right] = \mathbb{E}_{\boldsymbol{\phi}}\mathbb{E}_P[L^2_P(\boldsymbol{\phi})]. \tag{7}$$

Используя формулу (5), получим

$$\mathbb{E}_{\boldsymbol{\phi}}\mathbb{E}_P[L^2_P(\boldsymbol{\phi})] = \mathbb{E}_{\boldsymbol{\phi}_c}\mathbb{E}_P\left(\mathrm{Tr}\,\rho\, U(\boldsymbol{\phi}_c)^\dagger P U(\boldsymbol{\phi}_c)\right)^2 = \\ = \mathbb{E}_P\left(\mathrm{Tr}\,\rho P\right)^2 = 2^{-n}. \tag{8}$$

На втором шаге мы отбросили $U(\boldsymbol{\phi}_c)$, так как при фиксированном операторе Клиффорда $U(\boldsymbol{\phi}_c)$ оператор $U(\boldsymbol{\phi}_c)^\dagger P U(\boldsymbol{\phi}_c)$ также является случайным Паули-оператором, и при усреднении по P сопряжение $U(\boldsymbol{\phi}_c)$ не дает вклад. Последнее среднее – это вероятность того, что случайный оператор Паули имеет ненулевое среднее в заданном стабилизаторном состоянии (например, $\rho = |0\rangle^{\otimes n}$), и она равна 2^{-n}.

Таким образом, функция потерь ВКА типа КПР с наблюдаемой в виде случайного оператора Паули имеет БП с вероятностью, экспоненциально близкой к единице. В то же время возможны исключения. Например, в случае, когда ПКЦ имеет малую глубину, а наблюдаемая оказалась локальной (что происходит с экспоненциально малой вероятностью).

Мы рассматриваем ансамбль ВКА со случайными наблюдаемыми Паули как способ формализовать понятие нелокальности наблюдаемой, дополняющий существующую литературу. В нашем определении не накладываются дополнительные ограничения на ПКЦ, но при этом утверждение о наличии БП касается типичных наблюдаемых из заданного распределения. Важно отметить, что модель случайных операторов Паули позволит нам компактно доказать результаты об антиконцентрации, что и является основной целью настоящей работы.

Запутанность. По аналогии с предыдущим примером, мы предлагаем моделировать сильно запутанные начальные состояния с помощью случайных стабилизаторных состояний. Например, выбрав в качестве опорного состояния $\rho_0 = |0\rangle^{\otimes n}$, случайное стабилизаторное состояние ρ можно получить как $\rho = \mathcal{C}\rho_0\mathcal{C}^\dagger$, где $\mathcal{C}$ – случайный n-кубитный оператор Клиффорда. Проведя аналогичные вычисления, можно подтвердить наличие БП в этой модели

$$\mathbb{E}_\rho \mathbb{E}_\phi L_\rho^2(\boldsymbol{\phi}) = \mathbb{E}_{\boldsymbol{\phi}_c}\mathbb{E}_{\mathcal{C}}\left(\operatorname{Tr}\rho_0\mathcal{C}^\dagger U^\dagger(\boldsymbol{\phi}_c) P U(\boldsymbol{\phi}_c)\mathcal{C}\right)^2 = \tag{9}$$

$$\mathbb{E}_{\mathcal{C}}\left(\operatorname{Tr}\rho_0\mathcal{C}^\dagger P\mathcal{C}\right)^2 = 2^{-n}\ . \tag{10}$$

Здесь мы заменили усреднение по ρ на усреднение по группе Клиффорда, а затем отбросили $U(\boldsymbol{\phi}_c)$, так как они также являются операторами Клиффорда и не влияют на результат усреднения.

Экспрессивность. Наиболее распространенной мерой экспрессивности квантовой цепи является ее близость к унитарному t-дизайну. ПКЦ $U(\boldsymbol{\phi})$ называется t-дизайном, если ансамбль $U(\boldsymbol{\phi})$ совпадает по первым t моментам с мерой Хаара dU, т.е. выполнено следующее операторное равенство

$$\int d\boldsymbol{\phi}\ \left(U^\dagger(\boldsymbol{\phi})\right)^{\otimes t}\ U(\boldsymbol{\phi})^{\otimes t} = \int dU\ \left(U^\dagger\right)^{\otimes t}\ U^{\otimes t}. \tag{11}$$

Если квантовая цепь образует 2-дизайн, она обязательно приводит к БП [7], поскольку

$$\int dU\ \left(\operatorname{Tr}\rho U^\dagger \mathcal{O} U\right)^2 =$$

$$= \operatorname{Tr}\mathcal{O}^2\left(\frac{(\operatorname{Tr}\rho)^2}{4^n-1} - \frac{\operatorname{Tr}\rho^2}{2^n(4^n-1)}\right) = O(2^{-n}). \tag{12}$$

Если цепь реализует только приближенный 2-дизайн, как обычно бывает на практике, то это соотношение также выполняется приблизительно [21]. Для простоты будем рассматривать только точные t-дизайны.

Хотя экспрессивность описывает способность квантовой цепи охватывать большую часть пространства состояний, свойство 2-дизайна также предполагает определенную равномерность покрытия. Чтобы ослабить это требование, можно ввести статистическую модель экспрессивности. Один из вариантов – рассматривать такие ПКЦ $U(\boldsymbol{\phi})$, что при подстановке случайных углов Клиффорда $\boldsymbol{\phi}_c$ они равномерно покрывают группу Клиффорда: $U(\boldsymbol{\phi}_c) \sim \mathcal{C}$. Наличие БП в этом случае подтверждается следующим вычислением

$$\mathbb{E}_{\boldsymbol{\phi}}\left[L^2_{\rho,\mathcal{O}}(\boldsymbol{\phi})\right] = \mathbb{E}_{\boldsymbol{\phi}_c}\left[\left(\operatorname{Tr}\rho_0 U(\boldsymbol{\phi}_c)^\dagger \mathcal{O} U(\boldsymbol{\phi}_c)\right)^2\right] =$$

$$= \mathbb{E}_{\mathcal{C}}\left[\left(\operatorname{Tr}\rho_0\mathcal{C}\mathcal{O}\mathcal{C})\right)^2\right] = O(2^{-n}). \tag{13}$$

Однако такое определение все еще требует равномерного распределения по группе Клиффорда. Другой вариант – рассматривать ансамбль квантовых цепей $\mathcal{U}$, в котором равномерное покрытие достигается только при совместном случайном выборе U и $\boldsymbol{\phi}_c$. Пока у нас нет естественного примера такого ансамбля. Основной результат об антиконцентрации в данной работе будет доказан в стандартной модели БП, индуцированных экспрессивностью.

Антиконцентрация наблюдаемых Паули. Как правило, наблюдаемая в ВКА может быть разложена в линейную комбинацию операторов Паули: $\mathcal{O} = \sum_i c_i P_i$, что означает, что полная функция потерь также является линейной комбинацией функций потерь для отдельных операторов Паули: $L_\mathcal{O} = \sum_i c_i L_{P_i}$.

Для любой непостоянной функции L_P, соответствующей одному оператору Паули, ее максимум и минимум равны ± 1 и достигаются, например, в некотором подмножестве точек Клиффорда. Если при этом L_P обладает БП, она должна быть сильно сконцентрирована в окрестностях этих точек Клиффорда. Важно понять, локализуются ли функции потерь, соответствующие разным операторам Паули, в одних и тех же или в разных точках Клиффорда. Последнее мы называем антиконцентрацией наблюдаемых Паули.

С точки зрения ВКА, антиконцентрация наблюдаемых Паули, вероятно, является нежелательным явлением. Как отмечалось в [17], она создает множество плохих локальных минимумов в ландшафте полной функции потерь $L_{\mathcal{O}}$. Более того, это означает, что даже в экспоненциально малой области, где сконцентрирована полная функция потерь, большинство точек тривиальны, поскольку соответствуют только одному ненулевому члену. Следовательно, подпространство содержательных решений экспоненциально меньше даже по сравнению с областью локализации полной функции потерь.

Чтобы количественно оценить, коррелированы ли точки Клиффорда, где L_{P_1} и L_{P_2} принимают ненулевые значения, рассмотрим следующую величину:

$$A_{P_1,P_2} = \mathbb{E}_{\boldsymbol{\phi}_c}[|L_{P_1}(\boldsymbol{\phi}_c)L_{P_2}(\boldsymbol{\phi}_c)|] = \mathbb{E}_{\boldsymbol{\phi}_c}[L^2_{P_1}(\boldsymbol{\phi}_c)L^2_{P_2}(\boldsymbol{\phi}_c)]. \tag{14}$$

Этот коррелятор измеряет долю точек Клиффорда, в которых обе функции L_{P_1} и L_{P_2} принимают ненулевые значения. Если функции полностью коррелированы, т.е. принимают ненулевые значения в одних и тех же точках, $A_{P_1,P_2} = O(2^{-n})$. В противоположном случае, при полной антиконцентрации, ожидается $A_{P_1,P_2} = O(4^{-n})$.

Заметим, что в определении (14) необходимо брать модуль или квадрат, поскольку знакопеременное среднее может обращаться в ноль даже при значительном пересечении подмножеств ненулевых точек. Корреляторы, подобные (14), рассматривались в недавней работе [22].

Нелокальность. Антиконцентрацию проще всего доказать в модели нелокальности. Рассмотрим среднее значение A_{P_1,P_2} по ансамблю, в котором P_1 и P_2 – независимые случайные операторы Паули:

$$\begin{aligned}\mathbb{E}_{P_1}\mathbb{E}_{P_2}[A_{P_1,P_2}] = \mathbb{E}_{\boldsymbol{\phi}_c}\big[\mathbb{E}_{P_1}[L^2_{P_1}(\boldsymbol{\phi}_c)]\ \mathbb{E}_{P_2}[L^2_{P_2}(\boldsymbol{\phi}_c)]\big] =\\ = \mathbb{E}_{\boldsymbol{\phi}_c}[2^{-n}\times 2^{-n}] = 4^{-n}.\end{aligned} \tag{15}$$

Мы использовали независимость P_1 и P_2, чтобы факторизовать среднее. Это простое вычисление показывает, что любая типичная пара случайных операторов Паули антиконцентрирована с вероятностью, экспоненциально близкой к единице.

Запутанность. Теперь рассмотрим среднее значение коррелятора A_{P_1,P_2} по ансамблю случайных стабилизаторных состояний для двух произвольных операторов Паули:

$$\begin{aligned}&\mathbb{E}_\rho[A_{P_1,P_2}] =\\ &= \mathbb{E}_{\boldsymbol{\phi}_c}\mathbb{E}_{\mathcal{C}}[(\mathrm{Tr}\,\rho_0\mathcal{C}^\dagger U_c^\dagger P_1 U_c\mathcal{C})^2(\mathrm{Tr}\,\rho_0\mathcal{C}^\dagger U_c^\dagger P_2 U_c\mathcal{C})^2] =\\ &= \mathbb{E}_{\mathcal{C}}[(\mathrm{Tr}\,\rho_0\mathcal{C}^\dagger P_1\mathcal{C})^2(\mathrm{Tr}\,\rho_0\mathcal{C}^\dagger P_2\mathcal{C})^2].\end{aligned} \tag{16}$$

Здесь мы сперва заменили усреднение по стабилизаторным состояниям на усреднение по группе Клиффорда $\mathcal{C}$, а затем отбросили $U_c = U(\boldsymbol{\phi}_c)$, поскольку они тоже являются операторами Клиффорда и не влияют на среднее.

Рассмотрим подмножество операторов Клиффорда $\mathcal{C}_1$, таких что $(\mathrm{Tr}\,\rho_0\mathcal{C}^\dagger P_1\mathcal{C})^2 = 1$. Для $\rho_0 = |0\rangle^{\otimes n}$ это означает, что $\mathcal{C}^\dagger P_1\mathcal{C} = Z^\alpha = Z_1^{\alpha_1}\dots Z_n^{\alpha_n}$, т.е. преобразованный оператор Паули содержит только Z-компоненты. Тогда среднее можно записать как:

$$\mathbb{E}_\rho[A_{P_1,P_2}] = \frac{|\mathcal{C}_1|}{|\mathcal{C}|}\mathbb{E}_{\mathcal{C}_1}(\mathrm{Tr}\,\rho_0\mathcal{C}^\dagger P_2\mathcal{C})^2. \tag{17}$$

Отношение мощностей

$$\frac{|\mathcal{C}_1|}{|\mathcal{C}|} = \frac{2^n-1}{4^n-1} \tag{18}$$

получается как доля всех операторов Паули без X-компонент (исключая единичный), достижимых из P_1, к числу всех достижимых операторов Паули.

Обозначим через $\mathcal{C}_{1,\alpha}$ подмножество $\mathcal{C}_1$, переводящее P_1 в конкретный Z^α. Из свойств группы Клиффорда следует, что орбита $\mathcal{P}_\alpha = \mathcal{C}^\dagger_{1,\alpha}P_2\mathcal{C}_{1,\alpha}$ равномерно покрывает все операторы Паули, коммутирующие с Z^α при $[P_1,P_2]=0$, либо антикоммутирующие при $[P_1,P_2]\neq 0$. Следовательно, если P_1 и P_2 антикоммутируют, то среднее (17) обращается в ноль, поскольку любой оператор, антикоммутирующий с Z^α, обязательно содержит X-компоненты и имеет нулевое среднее в ρ_0.

Рассмотрим теперь среднее $\mathbb{E}_{\mathcal{C}_{1,\alpha}}[(\mathrm{Tr}\,\rho_0\mathcal{C}^\dagger P_2\mathcal{C})^2]$. Оно равно доле операторов Паули из орбиты $\mathcal{P}_\alpha$, имеющих ненулевое среднее в ρ_0. Для любого α существует ровно 2^{n-1} операторов вида X^γ, коммутирующих с Z^α, т.е. все элементы $\mathcal{P}_\alpha$ имеют вид $Z^\beta X^\gamma$. Всего таких операторов (исключая Z^α и единичный оператор) $2^{2n-1}-2$. Из них только 2^n-2 операторов вида Z^β имеют ненулевое среднее в ρ_0 (снова исключая Z^α и единицу). Поэтому

$$\mathbb{E}_{\mathcal{C}_{1,\alpha}}(\mathrm{Tr}\,\rho_0\mathcal{C}^\dagger P_2\mathcal{C})^2 = \frac{2^n-2}{2^{2n-1}-2}. \tag{19}$$

Так как все $\mathcal{C}_{1,\alpha}$ равномощны, получаем

$$\mathbb{E}_\rho[A_{P_1,P_2}] = \frac{2^n-1}{4^n-1}\cdot\frac{2^n-2}{2^{2n-1}-2} = O(4^{-n}) \tag{20}$$

для любых двух операторов Паули P_1 и P_2, отличных от единичного и друг от друга. Это доказывает антиконцентрацию в модели БП, индуцированных запутанностью.

Экспрессивность. Если принять статистическую модель БП, вызванных экспрессивностью, описанную выше – т.е. предположить, что при случайном выборе квантовой цепи и углов Клиффорда $U(\boldsymbol{\phi}_c)$ равномерно покрывают группу Клиффорда, – тогда среднее значение коррелятора (14) можно свести к уже рассмотренному случаю модели с запутанностью

$$\mathbb{E}_U[A_{P_1,P_2}] = \\ = \mathbb{E}_{\mathcal{C}}[(\operatorname{Tr}\rho_0 \mathcal{C}^\dagger P_1 \mathcal{C})^2 (\operatorname{Tr}\rho_0 \mathcal{C}^\dagger P_2 \mathcal{C})^2] = O(4^{-n}). \quad (21)$$

В этом случае можно также рассмотреть непрерывный аналог выражения (14)

$$A'_{P_1,P_2} = \mathbb{E}_{\boldsymbol{\phi}}\left[L^2_{P_1}(\boldsymbol{\phi}) L^2_{P_2}(\boldsymbol{\phi})\right] = \quad (22)$$
$$\mathbb{E}_{\boldsymbol{\phi}}\left[\left(\operatorname{Tr}\rho\, U^\dagger(\boldsymbol{\phi}) P_1 U(\boldsymbol{\phi})\right)^2 \left(\operatorname{Tr}\rho\, U^\dagger(\boldsymbol{\phi}) P_2 U(\boldsymbol{\phi})\right)^2\right].$$

Если $U(\boldsymbol{\phi})$ реализует 4-дизайн, то это среднее можно вычислить с помощью усреднения по мере Хаара, используя формализм Вайнгартена (см., например, [23])

$$\int dU\, U_{i_1,j_1} \dots U_{i_t,j_t}\, \overline{U}_{i'_1,j'_1} \dots \overline{U}_{i'_t,j'_t} = \quad (23)$$
$$\sum_{\sigma,\tau} \delta_{i_1,i_{\sigma(1)}} \cdots \delta_{i_n,i_{\sigma(n)}}\, \delta_{j_1,j_{\tau(1)}} \cdots \delta_{j_n,j_{\tau(n)}} \mathrm{Wg}(\tau\sigma^{-1}, 2^n).$$

Здесь $U_{i,j}$ – матричные элементы U, $\overline{U}_{i,j}$ – их комплексные сопряжения, а σ и τ – перестановки t элементов. Функция Wg – это функция Вайнгартена, задающая вес соответствующего члена суммы.

Применение формулы (23) к выражению (22) дает множество членов, представляющих собой произведения полностью свернутых тензоров $\rho^{\otimes 4}$ и $P_1^{\otimes 2} P_2^{\otimes 2}$. Большинство из них обнуляется, поскольку след большинства комбинаций операторов Паули равен нулю, например: $\operatorname{Tr} P_1 = \operatorname{Tr} P_1^2 P_2 = \operatorname{Tr} P_1 P_2 = 0$. Есть только два ненулевых члена: $\operatorname{Tr} P_1^2 P_2^2 = 2^n$, соответствующий перестановке $\tau = (1234)$, и $\operatorname{Tr} P_1^2 \operatorname{Tr} P_2^2 = 4^n$, соответствующий $\tau = (12)(34)$. Каждый из них умножается на определенные свертки $\rho^{\otimes 4}$, но для наших целей не требуется их полный перечень. Функция Вайнгартена четвертого порядка, асимптотически доминирующая при $n \to \infty$, соответствует тождественной перестановке и равна:

$$\mathrm{Wg}(\mathrm{id}, 2^n) = \frac{N^4 - 8N^2 + 6}{N^2(N^2-1)(N^2-4)(N^2-9)} = \\ = 16^{-n}\left(1 + O(2^{-n})\right), \quad (24)$$

где $N = 2^n$. Таким образом, ведущий член в асимптотическом разложении выражения (22) имеет вид

$$A'_{P_1,P_2} = \\ = \operatorname{Tr} P_1^2 \operatorname{Tr} P_2^2 (\operatorname{Tr}\rho^2)^2 \mathrm{Wg}(\mathrm{id}, 2^n)\left(1 + O(2^{-n})\right) = \\ = 4^{-n}\left(1 + O(2^{-n})\right), \quad (25)$$

что устанавливает непрерывную версию антиконцентрации для 4-дизайнов.

Хотя данное доказательство не применимо к 2-дизайнам, антиконцентрация может быть справедлива и в этом случае. Отметим, что мы установили аналогичное усреднение по группе Клиффорда в (16), несмотря на то, что группа Клиффорда не является 4-дизайном [24–26]. Если же 2-дизайны в самом деле не обязаны приводить к антиконцентрации, возникает интересная возможность: цепи, являющиеся 2-дизайнами, но не 4-дизайнами, находятся в тонком балансе между возможностью классической симуляции и избыточной экспрессивностью.

Численные эксперименты. Мы исследуем этот вопрос численно, используя ресурсоэффективный анзац (HEA) [27] (см. рис. 1a), который, как показано в [28], интерполирует между 2-дизайном и дизайнами более высокого порядка при увеличении числа слоев L. Для цепей с числом кубитов до $n = 10$ и $L = 30$ слоев мы оцениваем дисперсии функции потерь, выбирая 500 случайных точек для оценки как дискретного A_{P_1,P_2}, так и непрерывного A'_{P_1,P_2} корреляторов, для всех двухкубитных локальных операторов Паули (всего их $9(n-1)$).

Результаты показаны на рис. 1b и не дают однозначного ответа. Видно, что дисперсия функции потерь стремится к ожидаемому значению 2^{-n} как при равномерной, так и при Клиффордовой выборке. Дискретные A_{P_1,P_2} и непрерывные A'_{P_1,P_2} корреляционные функции также сходятся к ожидаемому значению 4^{-n} примерно с той же скоростью, что и дисперсия. Это указывает на то, что антиконцентрация возникает одновременно с барреновскими плато.

Сходимость дискретных корреляторов A_{P_1,P_2} при $n = 8$ и $n = 10$ выглядит несколько медленнее, однако пока не ясно, является ли это статистически значимым. Проведение подобных экспериментов при больших n быстро становится крайне трудным, поскольку увеличивается не только сложность симуляции, но и значения целевых величин становятся экспоненциально малы. Соответственно, для сохранения точности нужно экспоненциально увеличивать объем выборки. Проведение таких вычислений для больших n требует специализированных симуляторов и выходит за рамки настоящей работы.

Наш код написан с использованием библиотеки PennyLane [29] и доступен в [30]. Отметим, что полученные результаты схожи с приведенными в [17], но имеют два важных отличия. Во-первых, здесь мы напрямую исследуем корреляционные функции A_{P_1,P_2} и A'_{P_1,P_2}, тогда как в [17] использовалась косвенная мера антиконцентрации. Во-вторых, мы использовали специальный HEA-анзац, для которого было строго доказано, что он интерполирует между 2- и 4-дизайнами [28].

Заключение. В этой работе мы предложили несколько статистических моделей барреновских плато, дополняющих существующие детерминированные подходы. На основе этих моделей мы показали, что во всех трех сценариях возникновения БП – а именно, при нелокальности наблюдаемой, запутанности начального состояния и высокой экспрессивности квантовой цепи – операторы Паули антиконцентрированы с вероятностью, экспоненциально близкой к единице. Это означает, что типичные ландшафты с БП могут быть даже более сложными, чем считалось ранее.

Не только отдельные операторы Паули сконцентрированы в экспоненциально малом подпространстве параметров, но и их области локализации, как правило, существенно независимы. Это означает, что области параметров, где полная функция потерь имеет ненулевые значения, в основном соответствуют случаям, когда лишь один член не экспоненциально мал. Иными словами, конфигурации параметров, в которых ненулевыми оказываются сразу несколько слагаемых, составляют лишь экспоненциально малую и без того экспоненциально малого пространства локализации.

Хотя это наблюдение не означает принципиальную невозможность находить нетривиальные решения, оно подчеркивает, что стандартные методы, такие как градиентный спуск, в этих условиях крайне неэффективны. Представляется интересной идея искать нетривиальные решения в окрестностях тех точек Клиффорда, где одновременно наблюдается большое число ненулевых операторов Паули. Это превращает задачу предварительной инициализации в задачу дискретной оптимизации, заслуживающую дальнейшего изучения.

Мы полагаем, что дальнейшее исследование квантовых ландшафтов с использованием статистических моделей может оказаться плодотворным.

Финансирование работы. Работа поддержана Российским научным фондом, проект # 23-71-01095 (теоретические результаты). Численные эксперименты выполнены при поддержке программы

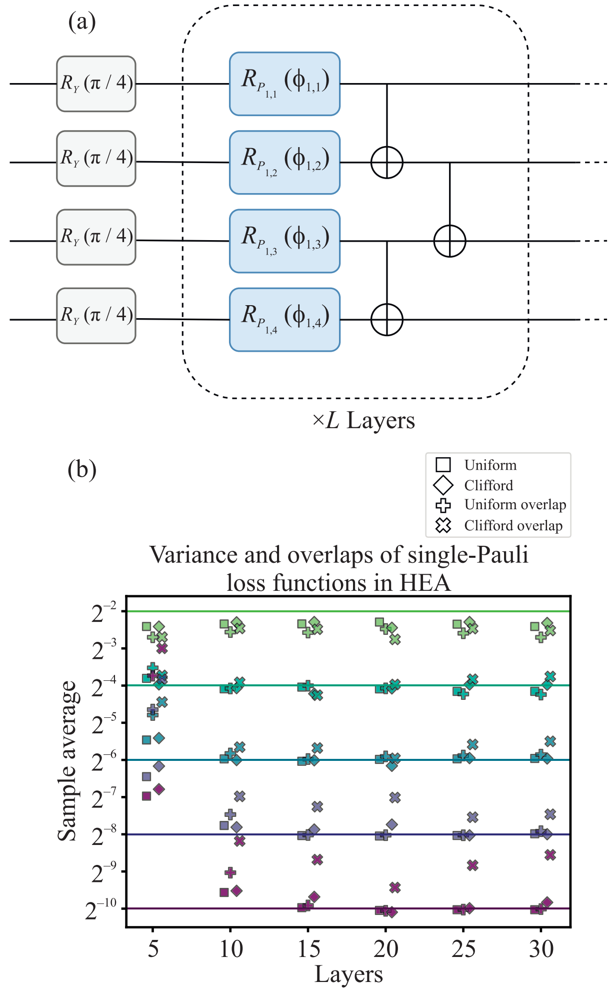

Рис. 1. (Цветной онлайн) (a) – Пример квантовой цепи HEA при $n = 4$ кубитах. Оси вращений Паули $R_{P_{i,j}}$ выбираются случайно. (b) – На графике показаны: (i) дисперсии функции потерь при равномерной ($\square = \mathbb{E}_{\boldsymbol{\phi}}[L_P^2(\phi)]$) и Клиффордовой ($\blacklozenge = \mathbb{E}_{\boldsymbol{\phi}_c}[L_P^2(\phi_c)]$) выборке; (ii) оценки непрерывного ($+ = A'_{P_1,P_2}$) и дискретного ($\times = A_{P_1,P_2}$) корреляционных коэффициентов. Значения L_{P_i} вычисляются по 500 точкам (равномерным или клиффордовым) для всех двухкубитных соседних операторов Паули, существующих при данном n. Усреднение производится по всем наблюдаемым и их попарным корреляциям. Корреляционные величины A_{P_1,P_2} и A'_{P_1,P_2} умножены на 2^n для сопоставимости с масштабом дисперсий

“Приоритет 2030” в Национальном исследовательском технологическом университете “МИСиС”, проект К1-2022-027.

Конфликт интересов. Автор данной работы заявляет, что у него нет конфликта интересов.

1. M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan,

L. Cincio, and P. J. Coles, Nature Reviews Physics **3**, 625 (2021); arXiv:2012.09265.

2. J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Nature **549**, 195 (2017); arXiv:1611.09347.
3. M. Schuld, R. Sweke, and J. J. Meyer, Phys. Rev. A **103**, 032430 (2021); arXiv:2008.08605v2.
4. L. Bittel and M. Kliesch, Phys. Rev. Lett. **127**, 120502 (2021); arXiv:2101.07267.
5. E. R. Anschuetz, arXiv:2109.06957 (2021).
6. E. R. Anschuetz and B. T. Kiani, Nat. Commun. **13**, 7760 (2022); arXiv:2205.05786.
7. J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Nat. Commun. **9**, 1 (2018); arXiv:1803.11173.
8. M. Larocca, S. Thanasilp, S. Wang, K. Sharma, J. Biamonte, P. J. Coles, L. Cincio, J. R. McClean, Z. Holmes, and M. Cerezo, Nature Reviews Physics **7**, 174 (2025); arXiv:2405.00781.
9. A. Arrasmith, M. Cerezo, P. Czarnik, L. Cincio, and P. J. Coles, Quantum **5**, 558 (2021); arXiv:2011.12245v2.
10. P. Bermejo, P. Braccia, M. S. Rudolph, Z. Holmes, L. Cincio, and M. Cerezo, arXiv:2408.12739 (2024).
11. M. Cerezo, M. Larocca, D. García-Martín, N. L. Diaz, P. Braccia, E. Fontana, M. S. Rudolph, P. Bermejo, A. Ijaz, S. Thanasilp, E. R. Anschuetz, and Z. Holmes, arXiv:2312.09121 (2023).
12. A. Skolik, J. R. McClean, M. Mohseni, P. van der Smagt, and M. Leib, Quantum Machine Intelligence **3**, 5 (2021); arXiv:2006.14904.
13. E. Grant, L. Wossnig, M. Ostaszewski, and M. Benedetti, Quantum **3**, 214 (2019); arXiv:1903.05076v3.
14. X.-M. Zhang, T. Li, and X. Yuan, arXiv:2201.11495 (2022).
15. S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles, Nat. Commun. **12**, 6961 (2021); arXiv:2007.14384.
16. M. S. Rudolph, J. Miller, D. Motlagh, J. Chen, A. Acharya, and A. Perdomo-Ortiz, arXiv:2208.13673 (2022).
17. N. A. Nemkov, E. O. Kiktenko, and A. K. Fedorov, Phys. Rev. A **111**, 012441 (2025); arXiv:2405.05332.
18. E. Farhi, J. Goldstone, and S. Gutmann, arXiv:1411.4028 (2014).
19. A. Letcher, S. Woerner, and Ch. Zoufal, arXiv:2309.12681 (2023).
20. M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Nat. Commun. **12**, 1 (2021); arXiv:2001.00550.
21. Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, PRX Quantum **3**, 010313 (2022); arXiv:2101.02138v2.
22. N. Dowling, P. Kos, and Xh. Turkeshi, arXiv:2408.16047.
23. Z. Puchała and J. A. Miszczak, Bulletin of the Polish Academy of Sciences Technical Sciences **65**, 21 (2017); arXiv:1109.4244v2.
24. Z. Webb, Quantum Information and Computation **16**, 1379 (2016); arXiv:1510.02769.
25. H. Zhu, Phys. Rev. A **96**, 062336 (2018); arXiv:1510.02619.
26. H. Zhu, R. Kueng, M. Grassl, and D. Gross, arXiv:1609.08172 (2016).
27. A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature **549**, 242 (2017); arXiv:1704.05018.
28. M. Liu, J. Liu, Y. Alexeev, and L. Jiang, npj Quantum Inf. **8**, 137 (2022); arXiv:2205.09900.
29. V. Bergholm, J. Izaac, M. Schuld et al. (Collaboration), arXiv:1811.04968 (2018).
30. N. Nemkov, https://github.com/idnm/barren_traps/tree/hea (2024).