Федеральное государственное автономное образовательное учреждение высшего образования
«Национальный исследовательский университет «Высшая школа экономики»

Факультет компьютерных наук
Основная образовательная программа
Прикладная математика и информатика

КУРСОВАЯ РАБОТА

Исследовательском проект на тему
Коммуникационная сложность

Выполнил студент группы 198, 3 курса,
Случ Дмитрий Борисович

Руководитель КР:
доцент д.ф.-м.н. Подольский Владимир Владимирович

Москва 2022

Содержание

Аннотация

Я изучаю коммуникационную сложность функций являющихся композицией с функциями EQ, IP_n и $DISJ$, а так же композицию с произвольной XOR функцией. Для EQ я привожу нижнюю оценку на одностороннюю коммуникационную сложность в детерминированном и квантовом точном случае, и верхнюю оценку в модели с общей монетой. Для композиции с IP_n и $DISJ$ я привожу нижнюю оценку в квантовой односторонней модели с ограниченной вероятностью ошибкой.

1 Введение

Композиции функций это интересный объект для анализа вычислительной сложности. Для функций $f \in \{0,1\}^n$ и $g \in \{0,1\}^b$ композиция определяется как $f \circ g : (\{0,1\}^b)^n \to \{0,1\}$, $f \circ g(x_{1,1}, \dots, x_{1,b}, \dots, x_{n,1}, \dots, x_{n,b}) = f(g(x_{1,1}, \dots, x_{1,b}), \dots, g(x_{n,1}, \dots, x_{n,b}))$. В частности я рассматриваю случай, когда $g \in \{0,1\}^b \times \{0,1\}^b \to \{0,1\}$ коммуникационная задача (их называют гаджетом). Рассмотрим следующую коммуникационную задачу: у Алисы есть строка $x \in \{0,1\}^{nb}$ у Боба $y \in \{0,1\}^{nb}$ они хотят вычислить $f \circ g(x, y)$. Естественно использовать следующий протокол - Алиса и Боб применяют эффективный протокол для f и когда протокол запрашивает очередной бит, применяют эффективный протокол для g. Теоремы поднятия выражают нижние границы на сложность вычисления композиций функций, через коммуникационную сложность гаджета, и сложность вычисления f, показывая в некоторых случаях, что такой наивный алгоритм оптимален. Я докажу несколько теорем поднятия для гаджетов EQ, $DISJ$ и IP_n.

1.1 Предварительная информация

1.1.1 Коммуникационная сложность

Модель коммуникационной сложности была впервые представлена Yao [8]. В детерминированном варианте, модель выглядит следующим образом. Есть два игрока Алиса и Боб, которые хотят вычислить функцию $f : X \times Y \to \{0,1\}$ на аргументах $x \in X, y \in Y$, при этом Алиса знает x, а Боб y. Им необходимо коммуницировать, чтобы вычислить функцию, и коммуникация происходит в соответствии с протоколом $\mathcal{P}$ зависящим от f. Протокол определяет завершилась ли передача, и если не завершилась, какой игрок отправляет бит следующим. Эта информация зависит только от уже переданных бит, т.к. только эта информация общая для Алисы и Боба. Также если ходит Алиса, протокол определяет, что она должна отправить в зависимости от x и уже переданной информации, если ходит Боб аналогично. После завершения коммуникации значение функции $f(x, y)$, должно быть возможно восстановить зная только переданные биты. Вычислительные ресурсы Алисы и Боба считаются неограниченными, нас интересует только количество коммуникации между ними. Стоимость протокола это максимальное по всем входам $x \in X, y \in Y$, число бит переданных протоколом. Коммуникационной сложностью (D_{cc}) называется минимальная стоимость протокола. Иногда рассматривают немного другие модели. Например можно ослабить требование, что $f(x, y)$ однозначно определяется зная только переданные биты, и требовать чтобы только один из игроков мог восстановить $f(x, y)$. От этого сложность изменится не более чем на 1 бит. Иногда требуют также, чтобы игроки отправляли биты по очереди(в оригинальной статье Yao было такое требование), сложность при этом изменится не более чем в 2 раза. На протокол можно смотреть как на бинарное дерево, внутренние вершины соответствующие ходам Алисы помечены функциями от x, ходам Боба - функциями от y, листовые помечены 0 или 1 - значением функции для пар (x, y) соответствующих листу. Переход в ребенка из вершины соответствует отправке бита, путь из корня до вершины однозначно соответствует уже переданным битам, зная этот путь мы можем понять листовая ли вершина, и если нет чьему ходу она соответствует, зная дополнительно к этому x или y мы определяем в какого ребенка пойдем. Оказывается [2], что множества пар соответствующих одной вершине (в том числе листу) обладают структурой, которая называется комбинаторным прямоугольником.

Определение 1. Комбинаторным прямоугольником называется подмножество $R \subseteq X \times Y$, такое что $\exists A \subseteq X, \exists B \subseteq Y, R = A \times B$. Комбинаторный прямоугольник R называется монохромным, если $\exists v \in \{0,1\} \forall (x,y) \in R, f(x,y) = v$.

Методы построения нижних оценок на D_{cc} основаны на том, что мы доказываем, что мы не сможем покрыть $X \times Y$ "маленьким"количеством монохромных прямоугольников, а значит листов в протоколе "много". Максимальное количество ребер на пути из корня в бинарном дереве, не меньше $\log_2(l)$, где l количество листьев в дереве, таким образом если мы докажем что требуется не менее l монохромных прямоугольников, чтобы покрыть $X \times Y$ это будет означать, что $D_{cc}(f) \geq \log_2(l)$. Одним из способов доказать, что прямоугольников "много"является техника fooling set [2]:

Теорема 2. *Пусть* $f :\in \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, $S \subset X \times Y, \exists z$ *такой что* $\forall (x', y'), (x'', y'') \in S, f(x', y') = f(x'', y'') = z$, *при этом* $f(x', y'') \neq z$ *или* $f(x'', y') \neq z$. *Тогда* $D_{cc}(f) \leq \log_2(|S|)$.

Доказательство. Два элемента из S не могут лежать в одном монохромном прямоугольнике. Допустим противное, $\exists X_R \subseteq X, Y_R \subseteq Y, R = X_R \times Y_R$ - монохромный прямоугольник. $\exists (x', y'), (x'', y'') \in S, (x', y'), (x'', y'') \in R$. Тогда $x', x'' \in X_R, y', y'' \in Y_R$, а значит $(x', y''), (x'', y') \in R = X_R \times Y_R$. Но это значит, что R не монохромный. Противоречие. Значит невозможно покрыть M_f менее чем $|S|$ монохромными прямоугольниками, значит в дереве не менее $|S|$ листов и глубина дерева не менее $\log_2 S$. □

На основании метода fooling set доказывается следующая теорема [2]:

Теорема 3. *Пусть* $f :\in \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, *а* M_f *- коммуникационная матрица* f. *Тогда*

$$D_{cc}(f) \geq \log \operatorname{rk}(M_f).$$

Модель в которой передавать биты разрешается только Алисе, а после этого Боб должен назвать результат обозначим за $D_{cc}^{\rightarrow}$. Пусть сообщение, которое передает Алиса содержит не более k бит. Это сообщение зависит только от x, таким образом существует разбиение на X на 2^k классов, соответствующих каждому сообщению. Рассмотрим коммуникационную матрицу M_f, разбиение на x соответствует разбиению на строках этой матрицы. Если все строки относящиеся к одному классу совпадают, то Бобу достаточно взять любой x' из класса и $f(x', y)$ будет ответом. Если какие-то строки из класса не совпадают, то протокол не корректен, т.к. разбиение не зависит от x и y и мы можем подобрать y', x' и x'' такие, что $f(x', y') \neq f(x'', y')$ при этом x' и x'' лежат в одном классе, и Алиса передала одинаковое сообщение. По принципу Дирихле, если $2^k < \operatorname{nrows}(M_f)$ в каком-то классе окажутся не совпадающие строки и протокол не корректен. Отсюда следует следующая теорема: [2]

Теорема 4. *Пусть* $f \in \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, M_f *коммуникационная матрица* f. *Тогда*

$$D_{cc}^{\rightarrow}(f) = \lceil \log_2 \operatorname{nrows}(M_f) \rceil,$$

где $\operatorname{nrows}(M_f)$ *- количество различных строк в матрице* M_f.

Существует еще более ограниченная модель ($D_{cc}^{||}, SMP$), в которой Алисе и Бобу вообще не разрешается передавать информацию между собой, они должны передать какое-то сообщение рефери, который затем называет значение функции.
В рандомизированных моделях игрокам разрешается подбрасывать монету и вычислять функцию с ошибкой с константной вероятностью. Я приведу алгоритм в модели, где оба игрока видят результат подброшенный монеты (public coin, R^{pub}), также существует модель где у каждого игрока источник случайности свой (R^{priv}). Существует обобщение техники fooling set и на рандомизированную модель, такая техника называется discrepancy. Для того чтобы ее описать понадобится ввести еще одну модель $R_{\mu,\varepsilon}^{dist}$ - в этой модели на $X \times Y$ задано распределение вероятностей μ и вероятность входов (x, y) на которых алгоритм ошибается должна быть ограничена ε. Модели R_{ε}^{pub} и $R_{\mu,\varepsilon}^{dist}$ связывает следующая теорема: [2]

Теорема 5. *Пусть* $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, *тогда* $R_\varepsilon^{pub}(f) = \max_\mu R_{\mu,\varepsilon}^{dist}(f)$.

Идея техники discrepancy заключается в том, что мы доказываем что для всех "больших"по вероятностной мере комбинаторных прямоугольников, вероятность входов из этого прямоугольника, на которых функция принимает значение 0 приблизительно равна вероятности входов из этого прямоугольника на которых функция принимает значение 1, таким образом если этот прямоугольник будет соответствовать листу ошибка на этом листе будет большой, а значит нам надо брать прямоугольники "меньше"и их понадобится много. Формально discrepancy определяется так:

Определение 6. Пусть $f \in \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ функция, $R \subseteq X \times Y$ - прямоугольник. Определим

$$\mathrm{Disc}_\mu(R, f) = |\Pr_\mu[f(x,y) = 0, (x,y) \in R] - \Pr_\mu[f(x,y) = 1, (x,y) \in R]|.$$

Тогда discreptancy f при распределении μ:

$$\mathrm{Disc}_\mu(f) = \max_R \mathrm{Disc}_\mu(R, f).$$

Следующая теорема дает оценку на R_μ^{dist}, а значит и на R^{pub}: [2]

Теорема 7. *Пусть* $f \in \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ *- функция,* μ *- распределение вероятности, тогда*

$$\forall \varepsilon, R_{\mu,\varepsilon}^{dist} \geq \log_2(\frac{2\varepsilon}{\mathrm{Disc}_\mu(f)}).$$

Обобщение этих моделей на квантовый случай введено Yao [9], но более простое объяснение модели есть у R. de Wolf [7].

Определение 8. Квантовым состоянием называется линейная комбинация базисных векторов $\sum_{x\in\{0,1\}^n} \alpha_x |x\rangle$ где n обозначает число кубит. При этом вторая норма этой комбинации должна быть равна одному, т.е. $\sum_{x\in\{0,1\}^n} \alpha^2 = 1$.

Определение 9. Квантовым вентилем называется преобразование, которое действует на состояние умножением на унитарную матрицу. Пусть у вентиля l входов с номерами $1, \dots, l$ (без потери общности), а всего в состоянии n кубит. Тогда вентиль действует на состояние умножением на матрицу $U \otimes I_2^{\otimes n-l}$, где U унитарная матрица $2^l \times 2^l$, определяющая как вентиль действует на входы.

У Алисы и Боба есть какое-то состояние, которое разбито на часть, которая принадлежит Алисе, и часть которая принадлежит Бобу, Алиса и Боб могут неограниченно взаимодействовать со своими частями (т.е. применять к ним неограниченное количество вентилей), а так же "отправлять"кубиты - после чего общее состояние не меняется но меняется разбиение на кубиты Алисы и Боба. Количество отправленных кубитов необходимо минимизировать. Алиса и Боб могут добавлять кубиты, при этом общее состояние тензорно умножается на добавленный кубит, и если состояние можно представить как тензорное произведение, то часть этого произведения можно убрать. Я рассматриваю только односторонний случай таких моделей, когда Алиса отправляет кубиты Бобу. Количество операций без ошибки и с ограниченной вероятностью ошибки обозначим $Q_E^{\rightarrow}$ и $Q_\varepsilon^{\rightarrow}$ соответственно. Модель в которой игрокам разрешается разделять запутанное состояние произвольного размера обозначим за $Q_{\varepsilon,*}^{\rightarrow}$. В такой модели можно эмулировать рандомизированный алгоритм с публичной монетой, используя запутанной состояние в качестве источника случайности. H. Klauck доказал теоремы позволяющие строить оценки на квантовую коммуникационную сложность снизу. [1]

Теорема 10. $f \in \{0,1\}^n \to \{0,1\}$. *Пусть* M_f *коммуникационная матрица* f. *Тогда* $Q_E^{\rightarrow}(f) = \lceil \log_2 \mathrm{nrows}(M_f) \rceil$.

Определение 11. Бинарная энтропия для $p \in (0,1)$, $\mathbb{H}_{bin}$ определяется как энтропия Шеннона случайной величины принимающей значение 1 с вероятностью p и 0 с вероятностью $1-p$.

$$\mathbb{H}_{bin}(p) = -p\ln(p) - (1-p)\ln(1-p).$$

Определение 12. VC-размерность матрицы M (VC-dim(M)) это наибольшее k, такое, что существует $2^k \times k$ подматрица M' матрицы M, все строки которой различны. Как было показано H.Klauck [1] VC-размерность дает нижние оценки на квантовую одностороннюю сложность с ограниченной ошибкой.

Теорема 13. $f \in \{0,1\}^n \to \{0,1\}$, M_f *коммуникационная матрица* f. *Тогда*

$$Q_\varepsilon^{\rightarrow}(f) \geq (1 - \mathbb{H}_{bin}(\varepsilon))\,\text{VC-dim}(M_f),$$

$$Q_{\varepsilon,*}^{\rightarrow}(f) \geq \frac{1}{2}(1 - \mathbb{H}_{bin}(\varepsilon))\,\text{VC-dim}(M_f).$$

1.1.2 Разложение Фурье

Некоторые оценки можно построить рассматривая разложение Фурье функции над группой $\mathbb{Z}_2$. Пусть $[n] := \{1, \dots, n\}$. Тогда для каждого натурального n определим множество 2^n функций четности $\chi_S(x) = (-1)^{\sum_{i \in S} x_i}$ (Это вообще говоря характеры $\mathbb{Z}_2$). Пусть $f : \{0,1\}^n \to \{0,1\}$ функция на булевом кубе. Тогда коэффициентами Фурье называются коэффициенты индексируемые подмножествами $[n]$:

$$\hat{f}(S) = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x)\chi_S(x).$$

Функция f представляется следующим образом(и коэффициенты для которых это представление существует единственны):

$$f(x) = \sum_{S \in [n]} \hat{f}(S)\chi_S(x).$$

Семейство множеств S для которых коэффициенты Фурье не нулевые, обозначаются $\mathrm{supp}(f)$, отождествим множества S с характеристическими векторами, размерность подпространства $\mathbb{F}_2^n$ натянутого на вектора для которых коэффициенты Фурье не нулевые называется размерностью Фурье ($\dim_f(f)$). Также для некоторых оценок нужны p-нормы Фурье. Они определяются следующим образом

$$||f||_p = \sqrt[p]{\sum_{S \subset [n]} |\hat{f}(S)|^p}$$

с особыми случаями 0 и ∞ нормами: $||f||_0 = |\,\mathrm{supp}(f)|$, $||f||_\infty = \max_{S \in [n]}(\hat{f}(S))$.

1.1.3 Композиция с гаджетами

Довольно хорошо изучена композиция с гаджетом XOR. Определим следующую модель вычислений:

Определение 14. Non-adaptive parity decision tree complexity($\mathrm{NADT}_\oplus$) - минимальное количество множеств $S \in \{0,1\}$, таких, что зная значения $\chi_S(x)$ можно однозначно восстановить f.

Легко заметить, что для функции $f : \{0,1\}^n \to \{0,1\}, g(x,y) = f(x \oplus y)$, $D_{cc}^{\rightarrow}(g) \le \mathrm{NADT}_\oplus$, действительно, пусть $S_1, \dots, S_p$ - минимальный набор множеств, такой что зная $\chi_{S_1}(x), \dots, \chi_{S_p}(x)$ можно однозначно восстановить $f(x)$. Тогда Алисе достаточно передать Бобу для каждого $j \in [p]$, $a_j = \bigoplus_{i \in S_j} x_i$, затем Боб вычислит $b_j = a_j \oplus \bigoplus_{i \in S_j} y_i = \bigoplus_{i \in S_j} x_i \oplus y_i$, а зная эти значения возможно однозначно восстановить g. Оказывается верно равенство $D_{cc}^{\rightarrow}(g) = Q_E^{\rightarrow}(g) = \mathrm{NADT}_\oplus$. Оно следует из следующих теорем:

Теорема 15 (Sanyal [6])**.** *Для булевой функции* $f : \{0,1\}^n \to \{0,1\}$, $\dim_f = \mathrm{NADT}_\oplus(f)$.

Теорема 16 (Montanaro и Osbourne [5])**.** *Пусть* $f : \{0,1\}^n \to \{0,1\}$ *функция на булевом кубе. Определим* $g(x,y) = f(x \oplus y)$. *Тогда*

$$D_{cc}^{\rightarrow}(g) = Q_E^{\rightarrow}(g) = \dim_f(f).$$

При доказательстве этой теоремы, доказывается утверждение, которое потребуется мне в дальнейшем.

Утверждение 17 (Montanaro и Osbourne [5])**.** *Пусть* $f : \{0,1\}^n \to \{0,1\}$ *функция на булевом кубе. Определим* $g(x,y) = f(x \oplus y)$. *Тогда*

$$\mathrm{nrows}(M_g) = 2^{\dim_f(f)}.$$

В модели с общей монетой, Montanaro и Osbourne [5] показали следующие оценки:

Теорема 18. *Пусть* $f : \{0,1\}^n \to \{0,1\}, g(x,y) = f(x \oplus y)$, *тогда* $R^{||,pub}(g) = O(||f||_1^2)$.

Теорема 19. *Пусть* $f : \{0,1\}^n \to \{0,1\}, g(x,y) = f(x \oplus y)$, *тогда* $R^{||,pub}(g) = O(2^{n-1}(1 - ||f||_\infty))$.

Некоторым аналогом $\mathrm{NADT}^\oplus$ для композиции с гаджетом AND является следующая модель:

Определение 20. Non-adaptive AND decision tree complexity(NAADT) - минимальное количество множеств $S \in \{0,1\}$, таких, что зная значения $\mathrm{AND}_S(x)$ можно однозначно восстановить f.

По аналогии с XOR функциями, для AND функций можно доказать, что для $f:\{0,1\}^n \to \{0,1\}, g(x,y) = f(x \wedge y), D^{\rightarrow}_{cc}(g) \le \mathrm{NAADT}(f)$, однако промежуток между оценкой снизу и оценкой сверху экспоненциальный. Mande et. al [4] показали, что

$$\log_2(\mathrm{NAPDT}(f)) \le D_{cc}(g) \le \mathrm{NAPDT}(f)$$

при этом нижняя оценка достигается с точностью до константы.
Для композиции с гаджетом скалярного произведения IP верна следующая нижняя оценка, которую я обобщу в разд. 2:

Теорема 21 (Mande et.al [4])**.** *Пусть $f:\{0,1\}^n \to \{0,1\}$ функция зависящая от всех аргументов, $IP:\{0,1\}^b \times \{0,1\}^b \to \{0,1\}$ скалярное произведение в $\mathbb{F}_2$, тогда*

$$Q^{\rightarrow}_{\varepsilon}(f \circ IP) \ge (1 - \mathbb{H}_{bin}(\varepsilon))n(b-1), Q^{\rightarrow}_{\varepsilon,*}(f \circ IP) \ge \frac{1}{2}(1 - \mathbb{H}_{bin}(\varepsilon))n(b-1).$$

Я изучаю композицию с гаджетом EQ в случае односторонней коммуникационной сложности, существуют нижние оценки на эту композицию в случае двусторонней модели.

Определение 22. $\mathrm{DT}(f)$ - минимальное количество запросов к битам x, необходимое чтобы вычислить $f(x)$.

Определение 23. $\mathrm{L}(f)$ - минимальное количество листьев в решающем дереве $f(x)$.

Является открытой проблемой верно ли, следующее утверждение:

Предложение 24. $D_{cc}(f) \ge b\,\mathrm{DT}(f)$.

Если бы это было так, то для любой зависящей от всех бит функции было бы верно, что $D_{cc}(f) \ge b(\log_2(n)-1)$, т.к. в дереве должна быть хотя бы одна вершина зависящая от каждого бита, т.е. как минимум n вершин. Оценки асимптотически лучшей чем эта, получить нельзя, поскольку композицию с функцией $ADDR:\{0,1\}^{m+2^m} \to \{0,1\}, ADDR(i,x) = x_i, n = m + 2^m$, можно вычислить тривиальным алгоритмом за $O(b\log(n))$. Действительно, Алиса отправляет Бобу первые m b-элементных блоков. Боб сравнивает их со своими блоками, и получает адрес $i \in \{0,1\}^m$, который отправляет Алисе. Алиса отправляет Бобу блок с номером $m+i$, Боб сравнивает его со своим $m+i$ -тым блоком и отправляет Алисе ответ. Для такого алгоритма нужно $O(b\log n)$ бит коммуникации. Loff и Mukhopadhyay доказали близкую к предложению 24 нижнюю оценку [3]

Теорема 25. *Для $f \in \{0,1\}^n \to \{0,1\}, EQ:\{0,1\}^b \times \{0,1\}^b \to \{0,1\}, b > 100\log_2 n$*

$$D_{cc}(f \circ EQ) = \Omega\left(b\frac{\log \mathrm{L}(f)}{\log n}\right).$$

1.2 Результаты

В разд. 2 я доказываю нижнюю оценку на квантовую одностороннюю сложность с ограниченной ошибкой композиции произвольной функции, зависящей от всех аргументов и некоторого класса гаджетов, к которому принадлежат $DISJ$ и IP_n. В разд. 3 я рассматриваю композицию с гаджетом EQ. В разд. 3.1 я рассматриваю детерминированную модель и квантовую модель без ошибки, и доказываю в них, что для любой зависящей от всех аргументов функции нельзя достичь сложности лучшей, чем у тривиального алгоритма передающего все биты. В разд. 3.2 я рассматриваю рандомизированную модель, и доказываю, что в ней композицию функции $f:\{0,1\}^n \to \{0,1\}$ с гаджетом EQ можно вычислить за $O(n\log n)$ в не зависимости от размера блока, а так же что для некоторой f не существует алгоритма работающего быстрее, чем за $O(n)$. В разд. 4 я обобщаю оценку из разд. 3.1 на гаджеты являющиеся XOR-функциями. А именно я доказываю, что если $f:\{0,1\}^n \to \{0,1\}$ функция зависящая от всех аргументов, $h:\{0,1\}^b \to \{0,1\}, b \ge 2$ $D^{\rightarrow}_{cc}(f \circ h \circ XOR) = Q^{\rightarrow}_{E}(f \circ h \circ XOR) \ge n(\dim_f(h)-1)$

2 Композиция с $DISJ$ и IP_n

Доказательство теор. 21 легко обобщается на случай $f \circ g \circ AND$, где $g:\{0,1\}^b \to \{0,1\}$ при этом g принимает значение 0 на строке из одних нулей, и 1 на строках в которых ровно одна единица в произвольной позиции. $DISJ = \bigvee_{i=1}^{b} x_i \wedge y_i, IP_p = (\sum_{i=1}^{b} x_i \wedge y_i)^p \pmod p$, т.е. эти гаджеты представимы в виде $g \circ AND$, и g удовлетворяет условию. Сформулирую и докажу обобщенную теорему

Теорема 26. *Пусть $f : \{0,1\}^n \to \{0,1\}$ функция зависящая от всех аргументов, а $g : \{0,1\}^b \to \{0,1\}$ принимает значение 0, на строке из одних нулей, и 1 на строках с ровно одной единицей. Значения на других строках могут быть произвольными. Тогда для $\varepsilon \in (0, \frac{1}{2}), Q^{\to}_{\varepsilon}(f \circ g \circ AND) \geq (1 - \mathbb{H}_{bin}(\varepsilon))n(b - 1), Q^{\to}_{\varepsilon,*}(f \circ g \circ AND) \geq \frac{1}{2}(1 - \mathbb{H}_{bin}(\varepsilon))n(b-1)$*

Доказательство. По теор. 13, достаточно показать, что $VC(f \circ g) \geq n(b-1)$. Обозначим за M коммуникационную матрицу $f \circ g \circ AND$. Т.к. f зависит от всех своих аргументов, $\forall i \exists z_1^{(i)}, \dots, z_n^{(i)}, \exists v_i, f(z_1^{(i)}, \dots, z_{i-1}^{(i)}, 0, z_{i+1}^{(i)}, \dots, z_n^{(i)}) = v_i, f(z_1^{(i)}, \dots, z_{i-1}^{(i)}, 1, z_{i+1}^{(i)}, \dots, z_n^{(i)}) = 1 - v_i$.
Для каждого $i \in 1, \dots, n$ и $j \in 2, \dots, b$ определим строку $y^{(i,j)}$ следующим образом. Для каждого $k \in 1, \dots, n$ и $l \in 1, \dots b$

$$y_{k,l}^{(i,j)} = \begin{cases} z_k^{(i)} & \textit{если } i \neq k, l = 1 \\ 1 & \textit{если } i = k, l = j \\ 0 & \textit{иначе} \end{cases}$$

Таким образом, для $k \neq i$ k-тый блок $y^{(i,j)}$ имеет вид $(z^{(i)}, 0^{b-1})$, i-тый блок $y^{(i,j)}$ имеет вид $(0^{j-1}, 1, 0^{n-j})$. Рассмотрим $n(b-1)$ столбцов M соответствующих $y^{(i,j)}$ и докажем что на этих столбцах существует подматрица $2^{n(b-1)} \times n(b-1)$ все строки которой различаются. Рассмотрим произвольную строку $c \in \{0,1\}^{n(b-1)}$. Я покажу что существует строка коммуникационной матрицы, ограничение которой на столбцы y совпадает с строкой c. Определим $x \in \{0,1\}^{nb}$ следующим образом. Для всех $i \in 1, \dots, n$ и $j \in 2, \dots, b$, $x_{i,1} = 1$ и

$$x_{i,j} = \begin{cases} c_{i,j} & \textit{если } v_i = 0 \\ 1 - c_{i,j} & \textit{если } v_i = 1 \end{cases}$$

Т.е. первый элемент в блоке x_i равен 1, а остальные совпадают с c_i или с отрицанием c_i, в зависимости от значения v_i. Строка коммуникационной матрицы соответствующая x при ограничении на столбцы y дает строку c. Чтобы заметить это зафиксируем $i \in 1, \dots, n$ и $j \in 2, \dots, b$ и рассмотрим $f \circ g \circ AND(x, y^{(i,j)})$. Для каждого $k \neq i$ после применения AND только первый бит в блоке может остаться не нулевым, он равен $z_k^{(i)}$ (действительно первый элемент блока x_k равен 1, первый элемент блока $y_k^{(i,j)}$ равен z_i, а остальные 0 по определению), а значит $g \circ AND(x_k, y_k^{(i,j)}) = z_k^{(i)}$. Для $k = i$, только в j-том элементе i-того блока y не ноль, а единица, более того $x_{i,j} = c_{i,j}$ если $v_i = 0$, и 1 если $v_i = 1$, значит после применения конъюнкции останется только один ненулевой бит $x_{i,j}$, значит $g \circ AND(x_i, y_i^{(i,j)}) = x_{i,j}$. Таким образом после применения $g \circ AND$ строка имеет вид $z_1^{(i)}, \dots, z_{i-1}^{(i)}, c_i, z_{i+1}^{(i)}, \dots, z_n^{(i)}$, если $v_i = 0$, и $z_1^{(i)}, \dots, z_{i-1}^{(i)}, 1 - c_i, z_{i+1}^{(i)}, \dots, z_n^{(i)}$ если $v_i = 1$. По определению $z^{(i)}$ и v_i значение f примененной к этим входам равно $c_{i,j}$. Это завершает доказательство. □

3 Композиция с EQ

3.1 Сложность в случае без ошибки

Теорема 27. *Пусть $f : \{0,1\}^n \to \{0,1\}$ функция зависящая от всех аргументов, а $EQ : \{0,1\}^b \times \{0,1\}^b \to \{0,1\}$ функция равенства. Пусть $g(x,y) = f \circ EQ$ Тогда для $b \geq 2$, $Q^{\to}_{E}(g) = D^{\to}_{cc}(g) = nb$*

Доказательство. По теор. 4, теор. 10 $Q^{\to}_{E}(g) = D^{\to}_{cc}(g) = \lceil \log_2 \mathrm{nrows(g)} \rceil$, где nrows(g) обозначает количество различных строк в коммуникационной матрице g. Докажем что в коммуникационной матрице $g = f \circ EQ$ все строки различны. Рассмотрим $x^{(1)}, x^{(2)} \in \{0,1\}^{nb}$, докажем что $\exists y \in \{0,1\}^{nb} : g(x^{(1)}, y) \neq g(x^{(2)}, y)$. Дальше в записи нижние индексы обозначают номера b элементных блоков, а не элементов. Не теряя общности(разница только в порядке индексов), пусть $x^{(1)}_{1,\dots,l} = x^{(2)}_{1,\dots,l}, \forall i > l, x_i^{(1)} \neq x_i^{(2)}$ Найдем такие $z_1, \dots z_l$, что $\exists z_{l+1}, \dots, z_n$, $f(z_1, \dots, z_l, 0, \dots, 0) \neq f(z_1, \dots, z_l, z_{l+1}, \dots, z_n)$. Если таких $z_1, \dots, z_l$ не существует, то функция зависит только от l бит, а мы потребовали, что она зависит от n бит. Тогда положим для $i \in 1, \dots, l$

$$y_i = \begin{cases} x_i^{(1)} & \textit{если } z_i = 1 \\ t_i \in \{0,1\}^b, t \neq x_i^{(1)} & \textit{иначе} \end{cases}$$

для $i \in l + 1, \dots, n$

$$y_i = \begin{cases} x_i^{(2)} & \textit{если } z_i = 1 \\ t_i \in \{0,1\}^b, t \neq x_i^{(1)}, t \neq x_i^{(2)} & \textit{иначе} \end{cases}$$

Такое t_i найдется, т.к. в блоке $2^b \geq 4$ элемента, а мы запрещаем использовать максимум 2 из них. Но тогда для блока $i \leq l$, если $z_i = 1$ $EQ(x_i^{(2)}, y_i) = EQ(x_i^{(1)}, y_i) = EQ(x_i^{(1)}, x_i^{(1)}) = 1 = z_i$, если $z_i = 0$, $EQ(x_i^{(2)}, y_i) = EQ(x_i^{(1)}, y_i) = EQ(x_i^{(1)}, t) = 0 = z_i$. Для блока $i > l$ если $z_i = 1$, $EQ(x_i^{(1)}, y_i) = EQ(x_i^{(1)}, x_i^{(2)}) = 0, EQ(x_i^{(2)}, y_i) = EQ(x_i^{(2)}, x_i^{(2)}) = 1 = z_i$, если $z_i = 0$, $EQ(x_i^{(1)}, y_i) = EQ(x_i^{(1)}, t_i) = 0, EQ(x_i^{(2)}, y_i) = EQ(x_i^{(2)}, t_i) = 0 = z_i$. Таким образом $EQ(x^{(1)}, y) = z_1, \dots, z_l, 0, \dots, 0$, а $EQ(x^{(2)}, y) = z_1, \dots, z_l, z_{l+1}, \dots, z_n \Rightarrow f \circ EQ(x^{(1)}, y) \neq f \circ EQ(x^{(2)}, y)$ □

3.2 Сложность в модели с публичной монетой

В случае если мы разрешаем Алисе и Бобу ошибаться с ограниченной вероятностью, вычислить $f \circ EQ$ можно значительно быстрее. Известно, что в модели с ограниченной ошибкой, функцию равенства можно вычислить за $O(1)$. Обобщим этот алгоритм композицию с функцией равенства.

Теорема 28. $R_{1/4}^{||,pub}(f \circ EQ) = O(n \log n)$

Доказательство. Рассмотрим следующий алгоритм, для каждого блока Алиса генерирует k случайных строк $a_{i,1}^{(j)}, \dots, a_{i,b}^{(j)}$, и находит $c_i^{(j)} = \langle a_i^{(j)}, x_i \rangle$, где $\langle a_i^{(j)}, x_i \rangle$ - обозначает скалярное произведение в $\mathbb{F}_2$. Боб знает a и находит $d_i^{(j)} = \langle a_i^{(j)}, y_i \rangle$. Затем Алиса и Боб отправляют рефери строки c и d. Всего они передают $O(nk)$ бит. Для каждого блока i для всех j рефери сравнивает $c_i^{(j)}$ и $d_i^{(j)}$, и если все биты совпадают считает, что $z_i = 1$ иначе $z_i = 0$. Рефери отвечает что результат равен $f(z_1, \dots, z_n)$. Если $x_i = y_i$, то $c_i^{(j)} = d_i^{(j)}$ т.к. это просто одинаковые суммы, значит мы найдем этот бит правильно. Если строки x_i и y_i различаются, то пусть l первый бит, такой что $x_{i,l} \neq y_{i,l}$. Заметим, что $\langle x_i, a \rangle \oplus \langle y_i, a \rangle = \langle x_i \oplus y_i, a \rangle$ Пусть $a \in \{0,1\}$ такая что $\langle x_i \oplus y_i, a \rangle = v$ Тогда заменим в a l-тый бит на противоположный, полученную строку обозначим $a^{\oplus l}$. $\langle x_i \oplus y_i, a^{\oplus l} \rangle = 1 - v$. Таким образом количество a, таких что $\langle x_i \oplus y_i, a \rangle = 1$ равно 2^{b-1}, а вероятность, что для случайно выбранной строки a, $\langle x_i, a \rangle \neq \langle y_i, a \rangle$ равна $\frac{1}{2}$. Мы генерируем k строк, значит вероятность что мы неправильно выбрали z_i для фиксированного i равна 2^{-k}, а вероятность того что мы неправильно выбрали z_i хотя бы в одном блоке не больше чем $n2^{-k}$. Положим $k = \lceil \log_2 n \rceil + 1$, тогда $n2^{-k} \leq \frac{1}{4} = \varepsilon$, т.е. рефери ошибается не более чем в $\frac{1}{4}$ случаях. □

Эта оценка точная(tight) с точностью до умножения на логарифм. Это можно доказать, рассмотрев композицию с функцией адресации $ADDR : \{0,1\}^{m+2^m} \to \{0,1\}, ADDR(i,x) = x_i, n = m + 2^m$, VC-размерность такой композиции будет $\Omega(n)$. Она дает оценку снизу на квантовую сложность, с запутанным состоянием произвольного размера, а значит и на сложность в модели с публичной монетой. По сути мы будем использовать только один бит в блоке и сведем задачу к оценке VC-размерности композиции $ADDR \circ XOR$.

Теорема 29. $ADDR : \{0,1\}^{m+2^m} \to \{0,1\}, n = m+2^m, EQ : \{0,1\}^b \times \{0,1\}^b \to \{0,1\}, g = ADDR \circ EQ, Q_\varepsilon^{\to}(g) = \Omega(n), Q_{\varepsilon,*}^{\to}(g) = \Omega(n)$

Доказательство. Для $i \in \{0,1\}^m$ положим

$$y_{k,l}^{(i)} = \begin{cases} 1 - i_k & \textit{если } k \leq m, l = 1 \\ 0 & \textit{иначе} \end{cases}$$

Для всех строк $c \in \{0,1\}^{2^m}$ существует x, такой что соответствующая x строка коммуникационной матрицы M_g ограниченная на строки y равна c. Действительно, положим

$$x_{k,l} = \begin{cases} 1 - c_{k-m+1} & \textit{если } k > m, l = 1 \\ 0 & \textit{иначе} \end{cases}$$

В каждом блоке ненулевой элемент может стоять только на первой позиции, при этом в строке y ненулевые элементы только в блоках с индексом не большим m, а в x с большим. Для $k \leq m$ мы сравниваем блок $(1 - i_k, 0^{b-1})$ с блоком 0^b, блоки будут равны если $i_k = 1$, и не равны если $i_k = 0$. Аналогично блоки с $k > m$ будут равны если $c_{k-m+1} = 1$ и не равны иначе. Таким образом $g(y^{(i)}, x) = ADDR(i, c) = c_i$. Значит VC-dim$(g) = \Omega(n)$, а значит по теор. 13 $Q_\varepsilon^{\to}(g) = \Omega(n), Q_{\varepsilon,*}^{\to}(g) = \Omega(n)$. □

4 Композиция с XOR функциями

Можно заметить, что функция EQ которую мы рассматривали представима в виде $\wedge_{i=1}^{b} x_i \oplus y_i \oplus 1$. Ее коммуникационная сложность равна сложности функции $\wedge_{i=1}^{b} x_i \oplus y_i$, которая является XOR функцией. Я обобщу результат для композиции $f \circ EQ$, на композицию $f \circ g \circ XOR$

Теорема 30. *Пусть* $f : \{0,1\}^n \to \{0,1\}$ *функция зависящая от всех аргументов,* $h : \{0,1\}^b \to \{0,1\}$. *Пусть* $g = h \circ XOR$. *Тогда*

$$Q_E^{\rightarrow}(f \circ g) = D_{cc}^{\rightarrow}(f \circ g) \geq n(\dim_f(h) - 1).$$

Доказательство. Рассмотрим коммуникационную матрицу функции g. Как следует из теор. 17 количество различных строк в этой матрице $2^{\dim_f(h)} = 2^p$. Выберем 2^p различных строк и если среди них есть строки все биты которых отличаются, выкинем из каждой пары таких строк одну. Осталось не менее 2^{p-1} строк, обозначим множество соответствующих им значений параметров A. Т.к. все эти строки различны для любых $a', a'' \in A$, $\exists q : g(a', q) \neq g(a'', q)$. Также, т.к. мы выкинули из каждой пары полностью различных строк одну, для любых $a', a'' \in A$ существует бит в котором a', a'' совпадают, т.е. $\exists t$, такой что $g(a', t) = g(a'', t)$. Я докажу, что для любого x такого, что $\forall i, x_i \in A$ соответствующие этим x строки коммуникационной матрицы $f \circ g$ различны. Всего не менее $2^{(p-1)n}$ строк удовлетворяют этому свойству, значит по теор. 4, теор. 10 $Q_{cc}^{\rightarrow}(f \circ g) = D_{cc}^{\rightarrow}(f \circ g) \geq n(p-1)$. Рассмотрим $x^{(1)}, x^{(2)} : \forall i, x_i \in A$. Не теряя общности $x_{1,\ldots,l}^{(1)} = x_{1,\ldots,l}^{(2)}, \forall i > l, x_i^{(1)} \neq x_i^{(2)}$. Для $i > l, \exists t_i$: $g(x_i^{(1)}, t_i) = g(x_i^{(2)}, t_i)$ т.к. $x_i^{(1)}, x_i^{(2)} \in A$. Положим $c_i := g(x_i^{(1)}, t_i)$ Найдем такие $z_1, \ldots z_l$, что $\exists z_{l+1}, \ldots, z_n$, $f(z_1, \ldots, z_l, c_{l+1}, \ldots, c_n) \neq f(z_1, \ldots, z_l, z_{l+1}, \ldots, z_n)$. Если таких $z_1, \ldots, z_l$ не существует, то f зависит только от l бит, а мы потребовали, что она зависит от n бит. $\exists \alpha, \beta : h(\alpha) = 0, h(\beta) = 1$. В противном случае h тождественное отображение, $\dim_f(h) = 1$ и утверждение теоремы тривиально. $\forall i > l, \exists q_i : g(x_i^{(1)}, q_i) \neq g(x_i^{(2)}, q_i)$ т.к. $x_i^{(1)}, x_i^{(2)} \in A$. Положим $d_i := g(x_i^{(1)}, q_i)$. Для $i \in 1, \ldots, l$ положим

$$y_i = \begin{cases} x_i^{(1)} \oplus \alpha & \textit{если } z_i = 0 \\ x_i^{(1)} \oplus \beta & \textit{иначе} \end{cases}$$

для $i \in l+1, \ldots, n$

$$y_i = \begin{cases} t_i & \textit{если } c_i = z_i \\ q_i & \textit{если } c_i \neq z_i, c_i = d_i \\ q_i \oplus x_i^{(1)} \oplus x_i^{(2)} & \textit{если } c_i \neq z_i, c_i \neq d_i \end{cases}$$

Но тогда для i-того блока, $i \leq l$, если $z_i = 0$, $g(x_i^{(2)}, y_i) = g(x_i^{(1)}, y_i) = h(x^{(1)} \oplus x^{(1)} \oplus \alpha) = h(\alpha) = 0 = z_i$, если $z_i = 1$, $g(x_i^{(2)}, y_i) = g(x_i^{(1)}, y_i) = h(x^{(1)} \oplus x^{(1)} \oplus \beta) = h(\beta) = 1 = z_i$. Для блоков $i > l$, если $c_i = z_i$, $g(x_i^{(1)}, y_i) = g(x_i^{(1)}, t_i) = c_i, g(x_i^{(2)}, y_i) = g(x_i^{(2)}, t_i) = c_i = z_i$. Если $c_i \neq z_i, c_i = d_i$, $g(x_i^{(1)}, y_i) = g(x_i^{(1)}, q_i) = d_i = c_i, g(x_i^{(2)}, y_i) = g(x_i^{(2)}, q_i) = 1 - d_i = 1 - c_i = z_i$. Если $c_i \neq z_i, c_i \neq d_i$, $g(x_i^{(1)}, y_i) = g(x_i^{(1)}, q_i \oplus x_i^{(1)} \oplus x_i^{(2)}) = h(x_i^{(1)} \oplus q_i \oplus x_i^{(1)} \oplus x_i^{(2)}) = g(x_i^{(2)}, q_i) = 1 - d_i = c_i, g(x_i^{(2)}, y_i) = g(x_i^{(2)}, q_i \oplus x_i^{(1)} \oplus x_i^{(2)}) = g(x_i^{(1)}, q_i) = d_i = 1 - c_i = z_i$ Таким образом $f \circ g(x^{(1)}, y) = f(z_1, \ldots, z_l, c_{l+1} \ldots, c_n) \neq f \circ g(x^{(2)}, y) = f(z_1, \ldots, z_l, z_{l+1}, \ldots, z_n)$. Это завершает доказательство ☐

Следствие 31. *Пусть* $f \in \{0,1\}^n \to \{0,1\}, h \in \{0,1\}^b \to \{0,1\}$ *булевы функции, тогда*

$$\dim_f(f \circ h) \geq n(\dim_f(h) - 1)$$

Доказательство. Рассмотрим $g(x, y) = h(x \oplus y)$. По теор. 30, теор. 16 $\dim_f(f \circ h) = D_{cc}^{\rightarrow}(f \circ g) \geq n(\dim_f(h) - 1)$ ☐

Следствие 32. *Пусть* $f \in \{0,1\}^n \to \{0,1\}, h \in \{0,1\}^b \to \{0,1\}$ *булевы функции, тогда*

$$\mathrm{NADT}_{\oplus}(f \circ h) \geq n(\mathrm{NADT}_{\oplus}(h) - 1)$$

Доказательство. Рассмотрим $g(x, y) = h(x \oplus y)$. По теор. 31, теор. 15 $\mathrm{NADT}_{\oplus}(f \circ h) = \dim(f \circ h) \geq n(\dim_f(h) - 1) = n(\mathrm{NADT}_{\oplus}(h) - 1)$ ☐

Список литературы

[1] H. Klauck. On quantum and probabilistic communication: Las vegas and one-way protocols. *In Proc. 32nd Annual ACM Symp. Theory of Computing*, pages 664–651, 2000.

[2] Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, 1996. doi:10.1017/CBO9780511574948.

[3] Bruno Loff and Sagnik Mukhopadhyay. Lifting Theorems for Equality. In Rolf Niedermeier and Christophe Paul, editors, *36th International Symposium on Theoretical Aspects of Computer Science (STACS 2019)*, volume 126 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 50:1–50:19, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. URL: http://drops.dagstuhl.de/opus/volltexte/2019/10289, doi:10.4230/LIPIcs.STACS.2019.50.

[4] Nikhil S. Mande, Swagato Sanyal, and Suhail Sherif. One-way communication complexity and non-adaptive decision trees. 2021. URL: https://arxiv.org/abs/2105.01963, doi:10.48550/ARXIV.2105.01963.

[5] Ashley Montanaro and Tobias Osborne. On the communication complexity of xor functions. 2009. arXiv:0909.3392.

[6] Swagato Sanyal. Sub-linear upper bounds on fourier dimension of boolean functions in terms of fourier sparsity, 2014. URL: https://arxiv.org/abs/1407.3500, doi:10.48550/ARXIV.1407.3500.

[7] Ronald Wolf. Quantum communication and complexity. *Theoretical Computer Science*, 287:337–353, 09 2002. doi:10.1016/S0304-3975(02)00377-8.

[8] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing(preliminary report). In *Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing*, STOC '79, page 209–213, New York, NY, USA, 1979. Association for Computing Machinery. URL: https://doi.org/10.1145/800135.804414, doi:10.1145/800135.804414.

[9] Andrew Chi-Chih Yao. Quantum circuit complexity. In *FOCS*, 1993.