И. С. Сергеев, Вентильные схемы ограниченной глубины, *Дискретн. анализ и исслед. опер.*, 2018, том 25, номер 1, 120–141

DOI: https://doi.org/10.17377/daio.2018.25.577

ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Январь–март 2018. Том 25, № 1. С. 120–141

УДК 519.7 DOI: 10.17377/daio.2018.25.577

ВЕНТИЛЬНЫЕ СХЕМЫ ОГРАНИЧЕННОЙ ГЛУБИНЫ *)

И. С. Сергеев
ФГУП «НИИ „Квант“»,
4-й Лихачёвский пер., 15, 125438 Москва, Россия
E-mail: isserg@gmail.com

Аннотация. Получены асимптотически точные оценки сложности вычисления классов (m, n)-матриц с коэффициентами из множества $\{0, 1, \ldots, q-1\}$ вентильными схемами ограниченной глубины d при некоторых соотношениях между m, n и q. В наиболее важном случае $q = 2$ показано, что асимптотика сложности класса булевых (m, n)-матриц $\log n = o(m)$, $\log m = o(n)$, достигается на схемах глубины 3. Ил. 1, библиогр. 11.

Ключевые слова: вентильные схемы, сложность, глубина.

1. Введение

Вентильная схема — это ориентированный ациклический граф, некоторые вершины которого отмечены как входы, а некоторые вершины — как выходы, с условием: вершины-входы не имеют входящих рёбер, а вершины-выходы не имеют исходящих рёбер. Рёбра графа также называются *вентилями*. Вентильная схема с n входами и m выходами реализует целочисленную матрицу размера $m \times n$ (т. е. из m строк и n столбцов), которая определяется так: на пересечении i-й строки и j-го столбца стоит число ориентированных путей, соединяющих j-й вход и i-й выход схемы.

Вентильную схему допустимо также интерпретировать как схему из функциональных элементов сложения. Припишем входам вентильной схемы символы переменных $x_1, \ldots, x_n$. Функционирование схемы определим индуктивно по правилам: (1) ребро схемы распространяет функцию, приписанную вершине, из которой оно исходит; (2) вершине схемы приписывается сумма функций по всем входящим в неё рёбрам. Тогда на выходах схемы реализуется линейное преобразование переменных $x_1, \ldots, x_n$. Можно проверить, что матрица этого преобразования (при соответствии нумерации выходов) совпадает с матрицей из определения

*) Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований (проект 17–01–00485а).

вентильной схемы. Далее иногда будем обращаться к указанной интерпретации вентильной схемы.

Сложность вентильной схемы равна числу вентилей в ней, *глубина* — числу рёбер в длиннейшем ориентированном пути между входами и выходами схемы.

Вентильная схема — одна из простейших вычислительных моделей. Вопросы оптимального синтеза вентильных схем изучаются с 1950-х гг., начиная с работы О. Б. Лупанова [5]. Основной вопрос асимптотической теории — сложность реализации матриц из заданного класса (множества). Напомним, что сложность матрицы определяется как минимальная сложность схемы, реализующей её, а сложность класса — как максимальная сложность матрицы из класса (функция Шеннона). Дополнительно ставится вопрос о сложности класса матриц при реализации схемами с заданным ограничением на глубину.

Обозначим $E_q = \{0, 1, \ldots, q-1\}$. Пусть $L_d(A)$ означает сложность реализации матрицы A вентильными схемами глубины d, а $L_d(q, m, n)$ — функцию Шеннона сложности реализации класса матриц размера $m \times n$ с элементами из E_q с глубиной d; $L(q, m, n)$ — обозначение для функции Шеннона без ограничения на глубину.

В силу симметрии $L_d(q, m, n) = L_d(q, n, m)$ (схема для произвольной матрицы A превращается в схему для транспонированной матрицы A^T изменением ориентации рёбер). Поэтому при изложении следующих известных результатов полагаем $m \leqslant n$. Для компактности формулировок определим $\log x = \max\{\log_2 x, 1\}$.

В [5] О. Б. Лупанов получил первые результаты о сложности класса булевых матриц. Он доказал, что при $m = \omega(\log n)$[1)] имеет место соотношение

$$L_2(2, m, n) \sim \frac{mn}{\log n},$$

а при дополнительном ограничении $\log m = o(\log n)$ —

$$L(2, m, n) \sim L_2(2, m, n).$$

В случае матриц полиномиально эквивалентных размеров, т. е. если $\log m \asymp \log n$, асимптотики $L(2, m, n)$ и $L_2(2, m, n)$ расходятся. Мощностное рассуждение (см., например, [7]) показывает, что при $m = \Omega\big(\log^2 n\big)$

[1)] Для сравнения порядков роста используем следующие обозначения: $f = \omega(g)$ равносильно $g = o(f)$, $f = \Omega(g)$ равносильно $g = O(f)$, $f \asymp g$ означает $f = \Theta(g)$. Обозначение $f \sim g$ используется для асимптотического равенства.

справедлива нижняя оценка

$$L(2,m,n) \geqslant \frac{mn}{\log(mn)}\left(1+\Theta\left(\frac{\log\log n}{\log n}\right)\right). \tag{1}$$

Э. И. Нечипорук [6, 7] установил, что в ряде случаев асимптотика достигается на схемах глубины 3. А именно, если $\log m \sim c_{p,r}\log n$, где $c_{p,r} = \frac{p}{p(r-1)+r}$, $p, r \in \mathbb{N}$, то

$$L(2,m,n) \sim L_3(2,m,n) \sim \frac{mn}{\log(mn)}. \tag{2}$$

Пиппенджер [10] провёл доказательство в общем случае, получив при $H = mn\log q \to \infty$ универсальную верхнюю оценку

$$L(q,m,n) \leqslant 3m\log_3(q-1) + (1+\tau(H))\frac{H}{\log H} + O(n), \tag{3}$$

где $\tau(H) \asymp \sqrt{\frac{\log\log H}{\log H}}$, асимптотически точную в случае $\log n = o(m\log q)$, как показывает обобщающая (1) нижняя оценка, приведённая в [10]:

$$L(q,m,n) \geqslant 3m\log_3(q-1) + \left(1-\Theta\left(\frac{\log\log H}{\log H}\right)\right)\frac{H}{\log H}. \tag{4}$$

В частности, этот результат Пиппенджера закрывает вопрос об асимптотике сложности класса булевых матриц размера $m \times n$, $m = \omega(\log n)$. Однако формально схемы Пиппенджера имеют растущую глубину, поэтому в ряде работ (например, [4, 9]) поставлен вопрос о достижимости асимптотики сложности класса булевых матриц схемами ограниченной глубины.

В настоящей работе получен утвердительный ответ на этот вопрос. Показано, что при $m = \Omega(\log^{3/2} n)$ оценка

$$L_d(q,m,n) \leqslant (1+\tau(n))\frac{mn}{\log_q(mn)} \tag{5}$$

в случае $q = n^{o(1)}$ выполняется при условии $\tau(n) = o(1)$ и $d = 3$, в случае $q = \log^{O(1)} n$ — при $\tau(n) \asymp \sqrt{\frac{\log\log n}{\log n}}$ и $d = 3$, а в случае $q = o(n/\log^2 n)$ — при $\tau(n) \asymp \sqrt{\frac{\log\log n}{\log n}}$ и $d = 4$.

Более того, при наличии дополнительных ограничений $m = \Omega(\log^2 n)$ и $m \in n^{\mu}\log^{\pm O(1)} n$ для некоторой постоянной $\mu \in \mathbb{Q}$ оценка (5) доказывается с остаточным членом «правильного» порядка $\tau(n) \asymp \frac{\log\log n}{\log n}$ при $q = \log^{O(1)} n$ и $d = 3$, а также при $q = o(n/\log^2 n)$ и $d = 4$.

В заключение приводим универсальную оценку в виде

$$L(q,m,n) \leqslant 3m\log_3(q-1) + (1+\tau(H))\frac{H}{\log H} + n$$

при $H = mn\log q \to \infty$ и $\tau(H) \asymp \sqrt{\frac{\log\log H}{\log H}}$. Эта оценка при $q = n^{O(1)}$ достигается на схемах глубины $O(1)$, а при $n = \log^{O(1)} q$ остаточный член в ней можно уточнить до $\tau(H) \asymp \frac{\log\log H}{\log H}$.

Принципиальная возможность получения перечисленных результатов заложена в конструкции Пиппенджера [10], и требуется лишь взглянуть на неё под нужным углом зрения. По существу оптимальная схема глубины 3 получается из схемы в [10], если, оставив нетронутым средний слой, на котором сосредоточена основная сложность схемы, остальные слои выше и соответственно ниже него «склеить». Приём склеивания слоёв также применял Э. И. Нечипорук при выводе оценки (2).

Отметим, что задаче о сложности реализации матрицы вентильными схемами подобна задача о сложности реализации матрицы (векторными) аддитивными цепочками. В частности, методы оптимального синтеза вентильных схем, как правило, могут быть перестроены в методы оптимального синтеза аддитивных цепочек, и наоборот. Так, оценки, аналогичные (3) и (4), для аддитивных цепочек доказаны Пиппенджером в [11], а уточнение верхней оценки получено С. Б. Гашковым и В. В. Кочергиным в [1]. Впоследствии В. В. Кочергин в серии работ получил ряд обобщений, в частности, для класса матриц с индивидуальными ограничениями на размер каждого коэффициента (см., например, [3]).

Понятия глубины для вентильных схем и аддитивных цепочек не вполне аналогичны (глубина аддитивной цепочки соответствует глубине вентильной схемы с ограничением 2 на число входящих в каждую вершину рёбер), поэтому прямой переносимости результатов, связанных с глубиной, из одной модели в другую, как правило, нет. Решение задачи одновременной минимизации глубины и сложности в асимптотическом смысле для обычной аддитивной цепочки приведено в [2].

Более подробное освещение разнообразных аспектов теории вентильных схем, в том числе более полный перечень результатов приводятся в современных обзорах [4, 9].

Изложение построено следующим образом. В разд. 2 приводим вспомогательный результат о свойствах приближений специального вида, возникающих в конструкциях схем, подобных схемам из работы [10]. В разд. 3 доказывается основной результат о синтезе схем глубины 3, а в разд. 4 — расширение для глубины 4. В разд. 5 устанавливаются

оценки для класса матриц с ограничением общего вида на размер коэффициентов.

2. Приближение

В конструкции из [10] используются приближения действительных чисел из отрезка [1/2, 1] дробями вида

$$P(r_1,\dots,r_k)=\frac{r_1\cdots r_k-r_2\cdots r_k+\cdots+(-1)^{k-1}r_k+(-1)^k}{r_1\cdots r_k}$$
$$=1-\frac{1}{r_1}\biggl(1-\frac{1}{r_2}\biggl(1-\frac{1}{r_3}\biggl(\dots\biggl(1-\frac{1}{r_k}\biggr)\dots\biggr)\biggr)\biggr), \quad (6)$$

где $r_i\in\mathbb{N}$. Положим формально $P(\,)=1$.

Лемма 1. *Пусть* $\alpha\in[1/2,\,1]$.

(i) *При любом* $\delta\in(0,\,1/2]$ *существуют чётное число* $k\geqslant 0$ *и натуральные числа* $r_1,\dots,r_k$ *такие, что*

$$0\leqslant\varepsilon=P(r_1,\dots,r_k)-\alpha\leqslant\delta,\quad R=r_1\cdots r_k\leqslant\frac{2}{\delta},\quad k\leqslant\log_2(2/\delta).$$

(ii) *Если* $\alpha=\frac{u}{v}$, *где* $u,v\in\mathbb{N}$, *то при некотором чётном* k *имеет место представление*

$$\alpha=P(r_1,\dots,r_k),\quad R=r_1\cdots r_k\leqslant v^v.$$

Доказательство. Будем строить приближение (6) для числа α градиентным алгоритмом, как описано ниже.

Введём вспомогательные величины h_i, связанные с (6) и определяемые из соотношений

$$h_0=\alpha,\quad h_{i-1}=1-\frac{1}{r_i}h_i,\quad i>0. \quad (7)$$

Градиентный метод заключается в том, что на каждом шаге выбирается максимальное r_i такое, что $h_i\in[0,\,1]$. Если $h_i=1$, то процесс заканчивается в силу того, что $\alpha=P(r_1,\dots,r_i)$.

Несложно проверить, что выбор всегда возможен. Справедливы формулы

$$r_i=\left\lfloor\frac{1}{1-h_{i-1}}\right\rfloor,\quad h_i=r_i(1-h_{i-1}), \quad (8)$$

из которых следует, что h_i имеет вид $\lfloor x \rfloor / x$ при $x \geqslant 1$, поэтому $h_i \in [0,\, 1]$. Кроме того, последовательность $\{h_i\}$ возрастающая, поскольку

$$1 - h_i = 1 - \frac{\lfloor x \rfloor}{x} < \frac{1}{x} = 1 - h_{i-1}.$$

Как следствие, если $h_0 \geqslant 1/2$, то $h_i > 1/2$ и $r_i \geqslant 2$ при всех $i > 0$.

Положим $R_0 = 1$ и обозначим $R_i = r_1 \cdots r_i$. В силу (7) и (8)

$$P(r_1, \ldots, r_i) - \alpha = (-1)^i \frac{1 - h_i}{R_i} = (-1)^i \frac{h_{i+1}}{R_{i+1}}. \tag{9}$$

Если $1 - \alpha \leqslant \delta$, то условия леммы выполнены при $k = 0$. Иначе определим t из условий $R_t < 1/\delta$ и либо $1/\delta \leqslant R_{t+1}$, либо $h_t = 1$. Теперь неравенство $0 \leqslant (-1)^t (P(r_1, \ldots, r_t) - \alpha) \leqslant \delta$ следует из (9).

Если t чётно, положим $k = t$. В случае нечётного t положим $k = t + 1$ и $r'_{t+1} = \lceil 1/(\delta R_t) \rceil$. При этом $1/\delta \leqslant R'_{t+1} = R_t r'_{t+1} < 2/\delta$ и $r'_{t+1} \geqslant 2$. Если $h_t \neq 1$, то $r'_{t+1} \leqslant r_{t+1}$, а точность приближения гарантируется оценкой

$$P(r_1, \ldots, r_t, r'_{t+1}) - \alpha = \frac{1 - \frac{r'_{t+1}}{r_{t+1}} \cdot h_{t+1}}{R'_{t+1}} \in \left[0,\, \frac{1}{R'_{t+1}}\right] \subset [0,\, \delta].$$

Если $h_t = 1$, то

$$P(r_1, \ldots, r_t, r'_{t+1}) - \alpha = P(r_1, \ldots, r_t, r'_{t+1}) - P(r_1, \ldots, r_t) = \frac{1}{R'_{t+1}} \leqslant \delta.$$

Неравенство $k \leqslant \log_2(2/\delta)$ тривиально в силу $r_i \geqslant 2$ для всех i. Утверждение (i) доказано.

Если α — рациональное число со знаменателем дроби v, то последовательность $\{h_i\}$ состоит из рациональных чисел со знаменателями дробей v. Действительно, обозначая $h_i = u_i / v$, с учётом того, что $u_0 = u \in \mathbb{N}$, по формуле (8) получаем $u_i = r_i(v - u_{i-1}) \in \mathbb{N}$ при всех i. Кроме того, из $u_i \leqslant v$ следует $r_i \leqslant v$.

Поскольку $h_i > h_{i-1}$, все u_i различны, поэтому последовательность $\{h_i\}$ содержит не более $v - 1$ членов.

Вместо представления $\alpha = P(r_1, \ldots, r_k)$ с нечётным k можно выбрать представление $\alpha = P(r_1, \ldots, r_{k-1}, r_k - 1, r_k)$. Равенство справедливо в силу тождества

$$\frac{1}{r} = \frac{1}{r-1}\left(1 - \frac{1}{r}\right).$$

Оценка на R в (ii) следует из $k \leqslant v - 1$ и $r_i \leqslant v$. Лемма 1 доказана.

Более подробно свойства приближения (6) исследуются в [10].

3. Схемы глубины 3

Как правило, схемы, рассматриваемые далее, имеют слоистую структуру. Это означает, что длина всех путей в схеме одинакова, а множество рёбер естественным образом распадается на слои: слой образуют рёбра, расположенные на одинаковом расстоянии от входов (или выходов).

В следующих двух леммах описываются простые приёмы синтеза схем, комбинация которых приводит к основному результату. Первый приём известен как метод разрезания на полосы Лупанова [5]. Второй является частью метода Нечипорука [7] синтеза схем глубины 3.

Лемма 2. *Пусть $s \in \mathbb{N}$. Произвольная (m,n)-матрица*[2)] *над E_q может быть реализована вентильной схемой глубины 2 с $q^s n/s$ внутренними вершинами, $q^{s+1}n$ вентилями на первом слое и $m(n/s+1)$ вентилями на втором слое*[3)].

Доказательство. Разобьём матрицу на вертикальные полосы ширины s. На первом уровне реализуем всевозможные суммы в полосах (напомним, что вентильную схему можно интерпретировать как схему из элементов сложения): всего для каждой полосы q^s сумм, для неполной полосы ширины $s' < s$ их менее $(s'/s)q^s$. Для реализации одной суммы требуется не более sq вентилей. На втором слое суммы в строках складываются из сумм в полосах. Лемма 2 доказана.

Далее, как обычно, под *весом* булевой матрицы понимается число единиц в ней.

Лемма 3. *Пусть $p, r \in \mathbb{N}$. Произвольная булева (m,n)-матрица веса V может быть реализована схемой глубины 2 с np^{r-1} внутренними вершинами, rnp^{r-1} вентилями на первом слое и $V/r + m(n/p+1)$ вентилями на втором слое.*

Доказательство. Разобьём матрицу на вертикальные полосы ширины p. На первом уровне реализуем всевозможные суммы не более чем r переменных из каждой полосы. Для одной полосы таких сумм не более p^r, для неполной полосы ширины $p' < p$ их менее $(p'/p)p^r$.

Каждую строку матрицы внутри каждой полосы представим в виде суммы подстрок веса r и, при необходимости, одной подстроки меньшего веса. На втором слое схемы выполняется сложение подстрок. Лемма 3 доказана.

[2)] Матрица размера $m \times n$.

[3)] Здесь и далее в верхних оценках опускаем округления.

Помимо схем из двух указанных лемм, будем применять двойственные к ним схемы. Это значит, что схема для матрицы A получена из схемы, построенной одним из способов из лемм 2 и 3, для транспонированной матрицы A^T с последующим обращением ориентации вентилей. Чтобы сформулировать двойственное утверждение, нужно в исходной формулировке каждой из лемм поменять ролями m и n.

Из леммы 2 вытекает решение асимптотической проблемы при малых значениях m и q.

Следствие 1. *Пусть* $qm = n^{o(1)}$. *Тогда*

$$L_2(q,m,n) \leqslant (1+\tau(q,m,n))\frac{mn}{\log_q(mn)} + n, \quad \tau(q,m,n) \asymp \frac{\log(qm\log n)}{\log n}.$$

Доказательство. Оценка достигается на схеме из леммы 2 (транспонированная версия) с параметром $s = \lfloor \log_q n - 2\log_q \log n - 1 \rfloor$. Следствие 1 доказано.

В доказательстве следующей теоремы фактически строится упрощённая схема Пиппенджера [10].

Теорема 1. (i) *Пусть* $m \leqslant n$, $m = \Omega(\log^{3/2} n)$ *и* $q = n^{o(1)}$. *Тогда*

$$L_3(q,m,n) \leqslant (1+\tau(q,n))\frac{mn}{\log_q(mn)}, \quad \tau(q,n) \asymp \sqrt{\frac{\log(q\log n)}{\log n}}.$$

(ii) *Пусть дополнительно выполнено* $q = \log^{O(1)} n$, $m = \Omega(\log^2 n)$ *и* $m \in n^{\mu}\log^{\pm O(1)} n$ *для некоторой постоянной* $\mu \in \mathbb{Q}$. *Тогда*

$$L_3(q,m,n) \leqslant (1+\tau(n))\frac{mn}{\log_q(mn)}, \quad \tau(n) \asymp \frac{\log\log n}{\log n}.$$

Доказательство. (i) Пусть $m = n^{\alpha}$. Если

$$\frac{(3/2)\log\log n - O(1)}{\log n} \leqslant \alpha \leqslant \sqrt{\frac{\log(q\log n)}{\log n}},$$

то нужную оценку обеспечивают схемы глубины 2 из следствия 1. Поэтому далее полагаем, что $\alpha \geqslant \sqrt{\frac{\log(q\log n)}{\log n}}$.

Положим $\delta = \frac{1}{2}\sqrt{\frac{\log(q\log n)}{\log n}}$ и при помощи леммы 1 построим приближение

$$\frac{1}{1+\alpha} = P(r_t,\ldots,r_1) - \varepsilon, \quad 0 \leqslant \varepsilon \leqslant \delta, \quad R = r_1\cdots r_t \leqslant 2/\delta,$$

где $t = 2k$. Параметры r_i умышленно пронумерованы в обратном порядке по отношению к (6).

Можно считать, что $k > 0$, поскольку в силу выбора δ случай $k = 0$ соответствует малым значениям параметра α: $\alpha \lesssim \frac{1}{2}\sqrt{\frac{\log(q \log n)}{\log n}}$.

Далее опишем итерационную процедуру построения многослойной схемы, вычисляющей произвольную (m, n)-матрицу A. Предварительно дадим неформальное общее описание этой процедуры. На первом шаге методом из леммы 2 для матрицы строится схема глубины 2. При этом один слой схемы, условно говоря, простой (с относительно небольшим числом рёбер), а второй — сложный. Пусть A_0 — булева матрица, вычисляемая сложным слоем. На следующем шаге процедуры слой, реализующий матрицу A_0, заменяется двухслойной схемой, построенной методом из леммы 3. В этой схеме также один из слоёв простой, а другой — сложный. Булева матрица, реализуемая сложным слоем, обозначается через A_1. На очередном шаге вместо этого слоя подставляется двухслойная схема, построенная методом из леммы 3, и т. д. При реализации матриц A_i используются поочерёдно прямая и двойственная версия схем леммы 3. В итоге получим схему из $2k + 1$ слоёв, среди которых только один внутренний слой сложный — он будет определять сложность всей схемы.

Перейдём к формальному описанию процедуры, характеризующейся параметрами $\beta, a_1, \ldots, a_{2k}$, значения которых будут определены позднее.

1) Пусть β — некоторый параметр, при котором $\beta \log_q n$ — целое число. Для реализации матрицы A применим транспонированное преобразование из леммы 2 с разрезанием на горизонтальные полосы высоты $s = \beta \log_q n$. Второй слой схемы заполняется пучками из не более чем qs рёбер, а на первом слое реализуется некоторая булева матрица A_0 из $m_0 \leqslant mn^{\beta}/(\beta \log_q n)$ строк, $n_0 = n$ столбцов, веса $v_0 \leqslant n(m/(\beta \log_q n)+1)$.

Размеры возникающих далее в процессе трансформации схемы булевых матриц A_i обозначим через $m_i \times n_i$, а вес — через v_i.

2) Нечётная итерация. Разбиваем матрицу A_{2i} на вертикальные полосы ширины $n^{a_{2i+1}}$. При помощи леммы 3 с параметром $r = r_{2i+1}$ реализуем матрицу из каждой полосы. Первый слой схемы составляют пучки из не более чем r_{2i+1} вентилей. На втором слое реализуется некоторая матрица A_{2i+1} с размерными параметрами $m_{2i+1} = m_{2i}$, $n_{2i+1} \leqslant n_{2i} n^{a_{2i+1}(r_{2i+1}-1)}$ и весом $v_{2i+1} \leqslant v_{2i}/r_{2i+1} + m_{2i}(n_{2i} n^{-a_{2i+1}} + 1)$. Подставим эту двухслойную реализацию в схему для матрицы A.

3) Чётная итерация. Разбиваем матрицу A_{2i-1} на горизонтальные полосы высоты $n^{a_{2i}}$. При помощи леммы 3 с параметром $r = r_{2i}$ реализуем транспонированную матрицу из каждой полосы. Второй слой схемы составляют пучки из не более чем r_{2i} вентилей. На первом слое

реализуется некоторая матрица A_{2i} с размерными параметрами $m_{2i} \leqslant m_{2i-1}n^{a_{2i}(r_{2i}-1)}$, $n_{2i} = n_{2i-1}$ и весом $v_{2i} \leqslant v_{2i-1}/r_{2i} + n_{2i-1}(m_{2i-1}n^{-a_{2i}}+1)$. Подставим описанную конструкцию в схему для матрицы A.

Окончательно получаем схему, состоящую из $2k+2$ слоёв, в которой внутренний слой (с номером $k+1$, считая от входов) вычисляет матрицу A_{2k}, размеры и вес которой удовлетворяют оценкам:

$$m_{2k} \leqslant \frac{m \cdot n^{\beta+a_2(r_2-1)+a_4(r_4-1)+\cdots+a_{2k}(r_{2k}-1)}}{\beta \log_q n}, \tag{10}$$

$$n_{2k} \leqslant n^{1+a_1(r_1-1)+a_3(r_3-1)+\cdots+a_{2k-1}(r_{2k-1}-1)}, \tag{11}$$

$$\begin{aligned} v_{2k} \leqslant{}& \frac{mn}{R \cdot \beta \log_q n} \\ &+ \frac{m \cdot n^{1+\beta+a_1(r_1-1)+a_2(r_2-1)+\cdots+a_{2k-1}(r_{2k-1}-1)-a_{2k}}}{\beta \log_q n} \\ &+ \frac{m \cdot n^{1+\beta+a_1(r_1-1)+a_2(r_2-1)+\cdots+a_{2k-2}(r_{2k-2}-1)-a_{2k-1}}}{r_{2k} \cdot \beta \log_q n} \\ &+ \cdots + \frac{m \cdot n^{1+\beta-a_1}}{r_2 \cdots r_{2k} \cdot \beta \log_q n} \\ &+ n_{2k-1} + \frac{m_{2k-2}}{r_{2k}} + \cdots + \frac{m_0}{r_2 \cdots r_{2k}} + \frac{n}{r_1 \cdots r_{2k}}. \end{aligned} \tag{12}$$

Положим $\beta = (1+\alpha)/R - \gamma_0$, $a_1 = \beta + \gamma$ при некоторых положительных параметрах $\gamma_0, \gamma \in o(\delta)$ с условием, что n^{a_1} — целое число. Пусть далее $a_i = a_{i-1}r_{i-1}$ для всех $i > 1$.

При таком выборе для любого $i \leqslant 2k-1$ имеем

$$\begin{aligned} &\beta + a_1(r_1-1) + a_2(r_2-1) + \cdots + a_i(r_i-1) \\ &\qquad = \beta + (a_2-a_1) + (a_3-a_2) + \cdots + (a_{i+1}-a_i) = a_{i+1} - \gamma, \end{aligned}$$

тем самым числители в средней группе слагаемых оценки (12) равны $mn^{1-\gamma}$, и сумма этих членов в таком случае не превосходит

$$\frac{tmn^{1-\gamma}}{\beta \log_q n}. \tag{13}$$

Сумму членов в последней строке (12) грубо оценим сверху через

$$t(n_{2k} + m_{2k}). \tag{14}$$

В результате получаем оценку

$$v_t \leqslant \frac{mn}{(1+\alpha-\gamma_0 R)\log_q n} + \frac{tmn^{1-\gamma}}{\beta \log_q n} + t(n_{2k}+m_{2k}). \tag{15}$$

Схема глубины 3 для исходной матрицы A получается следующим образом. На среднем слое реализуется матрица A_{2k}. Остальные слои построенной схемы вокруг слоя, реализующего матрицу A_{2k}, склеиваются. Первый слой полученной таким образом схемы составляют n_{2k} пучков из не более чем $r_1 r_3 \cdots r_{2k-1}$ вентилей, а третий слой — m_{2k} пучков из не более чем $r_2 r_4 \cdots r_{2k} \cdot q\beta \log_q n$ вентилей. Суммарное число вентилей в этих двух слоях можно оценить через

$$Rq\beta \log_q n \cdot (n_{2k}+m_{2k}). \tag{16}$$

Используя формулы $a_i = r_{i-1}\cdots r_1 a_1$ и $\beta = a_1 - \gamma$, показатель степени у n в (10) перепишем так:

$$\begin{aligned}\beta + a_2(r_2-1) + a_4(r_4-1) + \cdots + a_{2k}(r_{2k}-1)\\ = a_1 R \cdot P(r_{2k},\ldots,r_1) - \gamma = (1+\alpha-(\gamma_0-\gamma)R)\left(\frac{1}{1+\alpha}+\varepsilon\right) - \gamma\\ = 1 + \varepsilon(1+\alpha) - (\gamma_0-\gamma)R\left(\frac{1}{1+\alpha}+\varepsilon\right) - \gamma.\end{aligned}$$

Аналогично сумма в показателе степени в (11) преобразуется следующим образом:

$$\begin{aligned}a_1(r_1-1) + a_3(r_3-1) + \cdots + a_{2k-1}(r_{2k-1}-1)\\ = a_1 R(1 - P(r_{2k},\ldots,r_1)) = (1+\alpha-(\gamma_0-\gamma)R)\left(\frac{\alpha}{1+\alpha}-\varepsilon\right)\\ = \alpha - \varepsilon(1+\alpha) - (\gamma_0-\gamma)R\left(\frac{\alpha}{1+\alpha}-\varepsilon\right).\end{aligned}$$

Теперь неравенства (10) и (11) можно переписать так:

$$m_{2k} \leqslant \frac{mn^{1+\varepsilon(1+\alpha)-(\gamma_0-\gamma)R\left(\frac{1}{1+\alpha}+\varepsilon\right)-\gamma}}{\beta \log_q n}, \tag{17}$$

$$n_{2k} \leqslant mn^{1-\varepsilon(1+\alpha)-(\gamma_0-\gamma)R\left(\frac{\alpha}{1+\alpha}-\varepsilon\right)}. \tag{18}$$

Осталось указать подходящие параметры. Выберем $\gamma \asymp \frac{1}{R}\sqrt{\frac{\log(q\log n)}{\log n}}$ и $\gamma_0 \sim 2\gamma$ так, что показатели степеней у n в оценках (13), (17), (18)

не превосходят $1 - \log\bigl(q\log^3 n\bigr)/\log n$, при этом $\beta\log_q n$ и n^{a_1} — целые числа. (Благодаря выбору множителя $\frac{1}{2}$ в определении величины δ разность $\frac{\alpha}{1+\alpha} - \varepsilon$ в показателе степени в (18) неотрицательна и по порядку равна α.) Требование $\beta\log_q n \in \mathbb{Z}$ выполнимо, поскольку $\gamma\log_q n = \Omega(1)$.

Тогда вклад слагаемых (13), (14) и (16) в оценку сложности является величиной $O(mn/\log^2 n)$. Теперь сложность схемы оценивается как сумма (15) и (16):

$$\frac{mn}{(1+\alpha-\gamma_0 R)\log_q n} + O\biggl(\frac{mn}{\log^2 n}\biggr) = \frac{mn}{(1+\alpha)\log_q n}\biggl(1 + O\biggl(\sqrt{\frac{\log(q\log n)}{\log n}}\biggr)\biggr).$$

Докажем (ii). Если $\mu = 0$, то достаточно воспользоваться следствием 1. Иначе пусть $m = n^\alpha$. Как следует из п. (ii) леммы 1, существует приближение

$$\frac{1}{1+\alpha} = P(r_1,\ldots,r_k) - \varepsilon, \quad |\varepsilon| = O\biggl(\frac{\log\log n}{\log n}\biggr), \quad r_1\cdots r_k = O(1).$$

Если $\varepsilon > 0$, то просто применяем конструкцию из доказательства п. (i). При этом можно положить $\delta, \gamma \asymp \frac{\log\log n}{\log n}$.

Случай $\varepsilon < 0$ сводится к случаю $\varepsilon > 0$. Для этого матрица разбивается на вертикальные полосы ширины $\widehat{n} = \frac{n}{\log^c n}$, где константа c выбирается из условия $m \geqslant \widehat{n}^\mu$. Подматрицы в полосах реализуются независимо. Затем схемы для подматриц объединяются путём отождествления выходов. Теорема 1 доказана.

4. Схемы глубины 4

Ценой роста глубины можно ослабить зависимость сложности схемы в теореме 1 от q. В этой схеме увеличению q препятствуют два обстоятельства: множитель q в оценке (16) сложности граничного слоя и требование целочисленности параметра ширины полосы $s = \beta\log_q n$. Оба препятствия обусловлены конструкцией из леммы 2.

Первое препятствие (лишний множитель) можно устранить, добавив в схему ещё один слой. Второе препятствие устраняется, если разрешить разбиение на полосы дробной ширины. Оба приёма совмещены в следующей лемме, идея которой просматривается в методе из [10], и которую можно использовать в качестве альтернативы лемме 2.

Лемма 4. *Пусть $\sigma > 0$. Произвольная (m, n)-матрица над E_q может быть реализована вентильной схемой глубины 3 с $2nq/\log q$ вентилями на первом слое, $n(\log q + 1)$ вершинами на глубине 1, $12q^\sigma(n/\sigma + 1) \times (\sigma + 3)\log q$ вентилями на среднем слое, $4q^\sigma(n/\sigma + 1)$ вершинами на глубине 2 и $m(n/\sigma + 1)$ вентилями на третьем слое.*

Доказательство. Положим *ёмкость* ячейки матрицы над E_q — число возможных значений, которые она может заключать, — равной q. Допустим возможность разбиения некоторых ячеек на части, которые будем называть *неполными ячейками*. Неполная ячейка может заключать числа из $aE_b = \{0, a, 2a, \ldots, (b-1)a\}$ и имеет по определению ёмкость b. Ячейка для чисел из aE_b (в случае $a = 1$ и $b = q$ это целая ячейка, иначе — неполная) может быть разбита на две части по следующему правилу. Пусть $b_1 b_2 \geqslant b$ и $b_1, b_2 < b$. Разрешим одной части ячейки заключать числа из множества aE_{b_1}, а другой — числа из $ab_1E_{b_2}$. Можно представить произвольное число из aE_b суммой двух чисел из aE_{b_1} и $ab_1E_{b_2}$ и разместить слагаемые в подходящих частях ячейки.

Ёмкость множества ячеек, включая, возможно, неполные, естественно определить как произведение ёмкостей составляющих (ёмкость целой ячейки может быть меньше ёмкости множества её частей).

Разобьём множество ячеек строки матрицы на подмножества с ёмкостью каждого подмножества в интервале $[2q^\sigma, 4q^\sigma]$ (включая, возможно, одно подмножество меньшей ёмкости). Для этого совершаем проход вдоль строки, добавляя в текущее подмножество ёмкости C очередную ячейку ёмкости b, если $Cb \leqslant 2q^\sigma$, иначе — часть этой ячейки с ёмкостью $b_1 = \lceil 2q^\sigma/C \rceil$. Вторая часть ячейки приобретает ёмкость $\lceil b/b_1 \rceil$ и становится очередной для следующего подмножества.

Указанное разбиение делит матрицу на вертикальные полосы. Строка в полосе может содержать не более $\sigma + 1$ целых ячеек и не более двух неполных. Число полос можно оценить, исходя из того, что выделение одного подмножества ёмкости не менее $2q^\sigma$ уменьшает ёмкость оставшейся части множества минимум в q^σ раз. Отталкиваясь от ёмкости исходного множества q^n, получаем, что матрица содержит не более $n/\sigma + 1$ полос.

Теперь опишем схему. На первом слое схемы для каждого входа x вычисляется вектор кратностей: $(1, 2, 4, \ldots, 2^{\lfloor \log(q/\log q) \rfloor})x$. Один такой вектор реализуется при помощи не более чем $2q/\log q$ вентилей. На втором слое вычисляются всевозможные суммы в полосах: число сумм в одной полосе не превосходит $4q^\sigma$ по построению, для вычисления одной суммы используется не более $\sigma + 3$ ячеек, для вычисления значения в ячейке

из кратностей, вычисленных на первом уровне схемы, достаточно $3\log q$ вентилей. Последняя оценка опирается на возможность разложения произвольного числа $a \in E_q$ в сумму $a = a_1 l + a_2$, $l = 2^{\lfloor \log(q/\log q)\rfloor}$, при этом $a_1 \leqslant 2\log q$, $a_2 < l$. На заключительном слое, как и в лемме 2, суммы в строках складываются из сумм в полосах. Лемма 4 доказана.

Доказанная лемма позволяет расширить результат следствия 1 и дополнить результат теоремы 1 о возможностях асимптотически оптимального синтеза в глубине 3.

Следствие 2. *Пусть* $q = o(n/\log^2 n)$, $m = n^{o(1)}$. *Тогда*

$$L_3(q,m,n) \leqslant (1+\tau(m,n))\frac{mn}{\log_q(mn)} + n, \quad \tau(m,n) \asymp \frac{\log(m\log n)}{\log n}.$$

Доказательство. Оценка достигается на схеме из леммы 4 (транспонированная версия) с параметром $\sigma = \log_q(n/\log^2 n)$. Следствие 2 доказано.

Теорема 2. (i) *Пусть* $m \leqslant n$, $m = \Omega(\log^{3/2} n)$ *и* $q = o(n/\log^2 n)$. *Тогда*

$$L_4(q,m,n) \leqslant (1+\tau(n))\frac{mn}{\log_q(mn)}, \quad \tau(n) \asymp \sqrt{\frac{\log\log n}{\log n}}.$$

(ii) *Пусть дополнительно* $m = \Omega\left(\log^2 n\right)$ *и* $m \in n^{\mu}\log^{\pm O(1)} n$ *для некоторой постоянной* $\mu \in \mathbb{Q}$. *Тогда*

$$L_4(q,m,n) \leqslant (1+\tau(n))\frac{mn}{\log_q(mn)}, \quad \tau(n) \asymp \frac{\log\log n}{\log n}.$$

Доказательство. (i) Положим $\delta = \frac{1}{2}\sqrt{\frac{\log\log n}{\log n}}$. Величины $\alpha, \varepsilon, k, t, R$ определим как в теореме 1.

При $\alpha \leqslant 2\delta$ схемы глубины 3 из следствия 2 обеспечивают нужный результат. При бóльших α повторяем построение теоремы 1, только используем в качестве отправной точки схему из леммы 4 вместо схемы из леммы 2 (т. е. модифицируется шаг п. 1 в доказательстве теоремы 1).

1) Пусть β — некоторый параметр. Для реализации (m,n)-матрицы A применим транспонированное преобразование из леммы 4 с параметром $\sigma = \beta\log_q n$. В полученной схеме третий слой содержит не более чем $2mq$ вентилей. Второй слой схемы образуют пучки из не более чем $(\beta\log_q n + 3)(\log q + 1)$ рёбер, а на первом слое реализуется некоторая булева матрица A_0 из $m_0 \leqslant 4n^{\beta}(m + \beta\log_q n)/(\beta\log_q n)$ строк, $n_0 = n$ столбцов, веса $v_0 \leqslant n(m/(\beta\log_q n) + 1)$.

Далее действуем так, как описано в пп. 2 и 3 доказательства теоремы 1. В итоге приходим к схеме из $2k+3$ слоёв, в которой средний слой (с номером $k+1$) вычисляет матрицу A_{2k}, для размеров и веса которой справедливы следующие оценки:

$$m_{2k} \leqslant \frac{4(m+\beta\log_q n)n^{\beta+a_2(r_2-1)+a_4(r_4-1)+\cdots+a_{2k}(r_{2k}-1)}}{\beta\log_q n}, \tag{19}$$

$$n_{2k} \leqslant n^{1+a_1(r_1-1)+a_3(r_3-1)+\cdots+a_{2k-1}(r_{2k-1}-1)}, \tag{20}$$

$$\begin{aligned} v_{2k} \leqslant{}& \frac{mn}{R\cdot\beta\log_q n} \\ &+ \frac{4(m+\beta\log_q n)n^{1+\beta+a_1(r_1-1)+a_2(r_2-1)+\cdots+a_{2k-1}(r_{2k-1}-1)-a_{2k}}}{\beta\log_q n} \\ &+ \frac{4(m+\beta\log_q n)n^{1+\beta+a_1(r_1-1)+a_2(r_2-1)+\cdots+a_{2k-2}(r_{2k-2}-1)-a_{2k-1}}}{r_{2k}\cdot\beta\log_q n} \\ &+\cdots+\frac{4(m+\beta\log_q n)n^{1+\beta-a_1}}{r_2\cdots r_{2k}\cdot\beta\log_q n} \\ &+ n_{2k-1}+\frac{m_{2k-2}}{r_{2k}}+\cdots+\frac{m_0}{r_2\cdots r_{2k}}+\frac{n}{r_1\cdots r_{2k}}. \end{aligned}$$

Положим $\beta=(1+\alpha)/R-\gamma_0$, $a_1=\beta+\gamma$ и $a_i=a_{i-1}r_{i-1}$ для всех $i>1$. В результате (см. доказательство теоремы 1) для веса матрицы A_{2k} выводится оценка

$$v_t \leqslant \frac{mn}{(1+\alpha-\gamma_0 R)\log_q n}+\frac{4t(m+\beta\log_q n)n^{1-\gamma}}{\beta\log_q n}+t(n_{2k}+m_{2k}). \tag{21}$$

Схема глубины 4 для исходной матрицы A получается следующим образом. Оставляем нетронутыми внешний слой из рёбер, соединённых с выходами, построенный на первом шаге, а также средний слой, где реализуется матрица A_{2k}. Слои в каждой из оставшихся двух связных частей схемы совмещаются. На первом слое полученной таким образом схемы расположены n_{2k} пучков из не более чем $r_1r_3\cdots r_{2k-1}$ вентилей, а на третьем слое — m_{2k} пучков из не более чем $r_2r_4\cdots r_{2k}\times 3(\beta\log_q n+3)\log q$ вентилей. Тогда суммарное число вентилей на первом, третьем и четвёртом слоях оценивается величиной

$$3R(\beta\log_q n+3)\log q\cdot(n_{2k}+m_{2k})+2mq/\log q. \tag{22}$$

Точно так же, как в теореме 1, оценки (19), (20) переписываются в виде

$$m_{2k} \leqslant \frac{4(m+\beta\log_q n)n^{1+\varepsilon(1+\alpha)-(\gamma_0-\gamma)R(\frac{1}{1+\alpha}+\varepsilon)-\gamma}}{\beta\log_q n}, \tag{23}$$

$$n_{2k} \leqslant mn^{1-\varepsilon(1+\alpha)-(\gamma_0-\gamma)R(\frac{\alpha}{1+\alpha}-\varepsilon)}. \tag{24}$$

Выберем $\gamma \sim \frac{c}{R}\sqrt{\frac{\log\log n}{\log n}}$ и $\gamma_0 \sim 2\gamma$ так, что показатели степеней у n в среднем слагаемом оценки (13) и оценках (23), (24) не превосходят $1 - 3\log\log n/\log n$ и при этом выполняется условие $n^{a_1} \in \mathbb{Z}$.

Теперь сложность схемы оценивается как сумма (21) и (22):

$$\frac{mn}{(1+\alpha-\gamma_0 R)\log_q n} + O\left(\frac{mn}{\log^2 n}\right) = \frac{mn}{(1+\alpha)\log_q n}\left(1 + O\left(\sqrt{\frac{\log\log n}{\log n}}\right)\right).$$

Доказательство п. (ii) не отличается от доказательства п. (ii) теоремы 1. Теорема 2 доказана.

5. Вычисление матриц с быстро растущими коэффициентами

В этом разделе рассматривается класс матриц с произвольным ограничением q на размер коэффициентов. Следующая лемма описывает строительный блок для оптимальных схем.

Лемма 5. *Пусть* $v = (3^{u_1}, \ldots, 3^{u_s})$, $u_i \in \mathbb{N}\cup\{0\}$, $u = \max\limits_{i=1,\ldots,s}\{u_i\}$. *Тогда при любом* d, $1 \leqslant d \leqslant u$, *справедливо*

$$L_d(v) \leqslant \min\{u,\, 2d-1\}\cdot 3^{u/d} + (s-1)3^{\lfloor (u-1)/d\rfloor}.$$

Доказательство. Удобнее рассмотреть реализацию транспонированного вектора v^T. На рис. 1 изображена базовая схема, вычисляющая вектор $w = (1, 3, \ldots, 3^u)^T$. Она имеет сложность $4u$ и глубину u. Нужная схема для v^T в случае $d = u$ получится, если из висячих рёбер оставить только те, которые ведут к выходам кратности 3^{u_i}.

При произвольном d запишем $u = pd + r$, $0 \leqslant r < d$. Обобщая схему на рис. 1, максимальную кратность 3^u вычислим в цепочке из d слоёв: из них r слоёв содержат группы из 3^{p+1} параллельно соединённых вентилей, остальные $d - r$ слоёв — группы из 3^p штук. Для вычисления

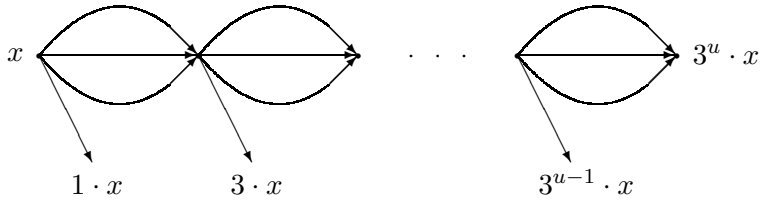

Рис. 1. Схема для вектора w

каждого из недостающих выходов достаточно добавить группу из не более чем 3^p вентилей, а если $r = 0$, то даже не более чем из 3^{p-1} вентилей.

Обозначим $x = r/d$. Число вентилей в цепочке, вычисляющей 3^u, оценивается так:

$$(d-r)3^p + r3^{p+1} = (d+2r)3^p$$
$$= d(1+2x)3^p \leqslant d(1+x)3^{p+x} = d(1+x)3^{u/d},$$

поскольку $1 + 2x \leqslant (1+x)(1+\ln 3 \cdot x) \leqslant (1+x)3^x$ для $x \geqslant 0$. Осталось заметить, что $r \leqslant \min\{u-d,\, d-1\}$. Лемма 5 доказана.

Замечание. Если к выходу схемы ведёт единственное ребро и он соединяется с входом другой схемы, то ребро можно удалить, отождествив его концы и при этом не нарушая функционирования объединённой схемы. Висячие рёбра, ведущие к выходам, как в схеме на рис. 1, нужны только затем, чтобы удовлетворить требованию определения схемы, запрещающей пути между входами или выходами. Соответственно когда выходы становятся внутренними вершинами, необходимость в этих рёбрах отпадает. По построению в схеме из доказательства леммы по меньшей мере $s - 1 - (u-d)$ выходов являются концами висячих рёбер.

Теорема 3. *Пусть $m \leqslant n$, $H = mn \log q \to \infty$.*

(i) *Пусть $1 \leqslant d \leqslant \lfloor \log_3(q-1) \rfloor$. Определим $s_d = \lfloor \log_3(q-1) \rfloor - d$ при $d \geqslant \log_3(q-1)/2$ и $s_d = d - 1$ в противном случае. Тогда*

$$L_{d+3}(q,m,n) \leqslant (1+\tau(H))\frac{H}{\log H} + n + mq^{1/d}(\lfloor \log_3(q-1) \rfloor + s_d), \quad (25)$$

где $\tau(H) \asymp \sqrt{\frac{\log\log H}{\log H}}$.

(ii) *Пусть дополнительно выполнено условие $n = \log^{O(1)} q$. Тогда*

$$L_d(q,m,n) \leqslant 3m \log_3(q-1) + (1+\tau(H))\frac{H}{\log H}, \quad (26)$$

где $d = \lfloor \log_3 q + \sqrt{\log_3 q} \rfloor + 3$ и $\tau(H) \asymp \frac{\log\log H}{\log H}$.

Доказательство. Случай $q \leqslant n/\log^3 n$ охватывают следствие 2, теоремы 1 и 2 (последнее слагаемое в оценке (25) в этом случае несущественно). Рассмотрим основной случай $q \geqslant n/\log^3 n$.

Базовый метод. Пусть $3^{k(t-1)} \leqslant q-1 < 3^{kt}$ при некоторых $k, t \in \mathbb{N}$. Используя запись в троичной системе счисления, каждый из элементов $a_{i,j}$ матрицы A представим в виде

$$a_{i,j} = bD_{i,j}c^T, \quad b = (1,\, 3^k,\, 3^{2k}, \ldots,\, 3^{(t-1)k}), \quad c = (1,\, 3,\, 3^2, \ldots,\, 3^{k-1}),$$

где $D_{i,j}$ — матрица размера $t \times k$ над E_3. Тогда

$$A = (a_{i,j}) = BDC, \quad B = \begin{pmatrix} b & 0 & \ldots & 0 \\ 0 & b & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & b \end{pmatrix},$$

$$D = \begin{pmatrix} D_{1,1} & \ldots & D_{1,n} \\ \ldots & \ldots & \ldots \\ D_{m,1} & \ldots & D_{m,n} \end{pmatrix}, \quad C = \begin{pmatrix} c^T & 0 & \ldots & 0 \\ 0 & c^T & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & c^T \end{pmatrix}.$$

Схему для матрицы A можно построить, соединяя последовательно схемы для матриц C, D, B.

Если векторы-подматрицы матриц B, C реализуются схемами из леммы 5, то висячие рёбра обращены внутрь схемы, и их можно удалить согласно замечанию к лемме 5.

Для доказательства п. (i) выберем $k = 1$. Тогда

$$L_{d+3}(A) \leqslant L'_d(B) + L_3(D) \leqslant mL'_d(b) + L_3(3, mt, n),$$

где функционал L'_d отличается от L_d отсутствием учёта висячих рёбер, ведущих к выходам.

Согласно лемме 5 в силу замечания к ней имеем

$$L'_d(b) \leqslant (\min\{t-1, 2d-1\} + t - 1 - d)3^{(t-1)/d} \leqslant (\lfloor \log_3(q-1) \rfloor + s_d)q^{1/d}.$$

Величину $L_3(3, mt, n)$ оценим при помощи следствия 1 или теоремы 1 в зависимости от соотношения между mt и n следующим образом:

$$L_3(3, mt, n) \leqslant (1 + \tau(H))\frac{H}{\log H} + n.$$

Здесь существенно, что $t = \lceil \log_3 q \rceil = (1 + O(1/\log H)) \log_3 q$ в силу ограничения снизу на q.

Для доказательства п. (ii) достаточно выбрать $k \leqslant t$ так, что выполняется условие $mt \in (nk)^{\mu} \log^{\pm O(1)}(nk)$ для некоторой постоянной $\mu \in \mathbb{Q}$. Случай $m = n = 1$ тривиален. Пусть $n \geqslant 2$, $m = n^{\alpha}$ и $\log_3 q = n^{\beta}$. Выберем $\mu \in [\alpha,\, \min\{1,\, \alpha + \beta/2\}] \cap \mathbb{Q}$ и положим $k = \left\lfloor n^{\frac{\mu(1+\beta)-\alpha}{1+\mu}} \right\rfloor$, $t = \left\lceil \frac{\log_3 q}{k} \right\rceil$. Реализуем матрицы B и C схемами из леммы 5 (максимальной) суммарной глубины $kt - 1 \leqslant \log_3 q + k - 1$, а матрицу D реализуем схемой глубины 3, как в п. (i). По построению $kt \leqslant \log_3 q + k = (1 + O(1/\sqrt{H})) \log_3 q$. Теорема 3 доказана.

В частном случае $d = \lfloor \log_3(q-1) \rfloor$ оценка п. (i) теоремы является незначительным уточнением оценки (3).

Следствие 3. *При условиях $m \leqslant n$, $H = mn \log q \to \infty$ справедливо соотношение*

$$L(q, m, n) \leqslant 3m \log_3(q-1) + (1 + \tau(H)) \frac{H}{\log H} + n, \tag{27}$$

где $\tau(H) \asymp \sqrt{\frac{\log\log H}{\log H}}$, а при дополнительном ограничении $q = n^{O(1)}$ оценка (27) *достигается на схемах глубины $O(1)$.*

Доказательство. Вторая часть утверждения получается при надлежащем выборе $d \geqslant \log_q n + 1$. Следствие 3 доказано.

В п. (ii) мы ограничились указанием только достаточно общего случая (очень большое q), когда оценка (26) может быть доказана с улучшенным остаточным членом. Способ доказательства и результат теоремы 1 позволяют получить «правильный» порядок величины $\tau(H)$ в некоторых других ситуациях.

В целом оценка п. (ii) служит обобщением своего частного случая $m = n = 1$, для которого результат может быть получен адаптацией классического метода Брауэра [8] построения аддитивных цепочек к модели вентильных схем. По существу, применяемая в доказательстве теоремы 3 декомпозиция $A = BDC$ является центральным местом в методе Брауэра.

Автор благодарен рецензенту за внимательное изучение статьи и полезные замечания.

ЛИТЕРАТУРА

1. **Гашков С. Б., Кочергин В. В.** Об аддитивных цепочках векторов, вентильных схемах и сложности вычисления степеней // Методы дискретного анализа в теории графов и сложности. Вып. 52. Новосибирск, Ин-т математики СО РАН, 1992. С. 22–40.
2. **Гашков С. Б., Сергеев И. С.** О применении метода аддитивных цепочек к инвертированию в конечных полях // Дискрет. математика. 2006. Т. 18, вып. 4. С. 56–72.
3. **Кочергин В. В.** О сложности вычисления систем одночленов с ограничениями на степени переменных // Дискрет. математика. 1998. Т. 10, вып. 3. С. 27–34.
4. **Кочергин В. В.** Теория вентильных схем (современное состояние) // Дискретная математика и ее приложения. Вып. 7. М.: Изд-во ИПМ РАН, 2013. С. 23–40.
5. **Лупанов О. Б.** О вентильных и контактно-вентильных схемах // Докл. АН СССР. 1956. Т. 111, № 6. С. 1171–1174.
6. **Нечипорук Э. И.** О вентильных схемах // Докл. АН СССР. 1963. Т. 148, № 1. С. 50–53.
7. **Нечипорук Э. И.** О топологических принципах самокорректирования // Проблемы кибернетики. Вып. 21. М.: Наука, 1969. С. 5–102.
8. **Brauer A.** On addition chains // Bull. Amer. Math. Soc. 1939. Vol. 45. P. 736–739.
9. **Jukna S., Sergeev I.** Complexity of linear Boolean operators // Found. Trends Theor. Comput. Sci. 2013. Vol. 9, No. 1. P. 1–123.
10. **Pippenger N.** The minimum number of edges in graphs with prescribed paths // Math. Syst. Theory. 1979. Vol. 12. P. 325–346.
11. **Pippenger N.** On the evaluation of powers and monomials // SIAM J. Comput. 1980. Vol. 9, No. 2. P. 230–250.

Сергеев Игорь Сергеевич

Статья поступила
19 апреля 2017 г.

Исправленный вариант —
4 сентября 2017 г.

DISKRETNYI ANALIZ I ISSLEDOVANIE OPERATSII
/DISCRETE ANALYSIS AND OPERATIONS RESEARCH/
January–March 2018. Volume 25, No. 1. P. 120–141

UDC 519.7 DOI: 10.17377/daio.2018.25.577

RECTIFIER CIRCUITS OF BOUNDED DEPTH

I. S. Sergeev
FSUE "RDI 'Kvant'",
15 Chetvyortyi Likhachyovskii Lane, 125438 Moscow, Russia
E-mail: isserg@gmail.com

Abstract. Asymptotically tight bounds are obtained for the complexity of computation of the classes of (m, n)-matrices with entries from the set $\{0, 1, \ldots, q-1\}$ by rectifier circuits of bounded depth d, under some relations between m, n, and q. In the most important case of $q = 2$, it is shown that the asymptotics of the complexity of Boolean (m, n)-matrices, $\log n = o(m)$, $\log m = o(n)$, is achieved for the circuits of depth 3. Illustr. 1, bibliogr. 11.

Keywords: rectifier circuit, complexity, depth.

REFERENCES

1. **S. B. Gashkov** and **V. V. Kochergin,** On addition chains of vectors, gate circuits, and the complexity of computations of powers, in *Metody diskretnogo analiza v teorii grafov i slozhnosti* (Methods of Discrete Analysis in Graph Theory and Complexity), Vol. 52, pp. 22–40, Izd. Inst. Mat., Novosibirsk, 1992. Translated in *Sib. Adv. Math.,* **4**, No. 4, 1–16, 1994.
2. **S. B. Gashkov** and **I. S. Sergeev,** An application of the method of additive chains to inversion in finite fields, *Diskretn. Mat.,* **18**, No. 4, 56–72, 2006. Translated in *Discrete Math. Appl.,* **16**, No. 6, 601–618, 2006.
3. **V. V. Kochergin,** On the complexity of calculating systems of monomials with restrictions on the powers of variables, *Diskretn. Mat.,* **10**, No. 3, 27–34, 1998. Translated in *Discrete Math. Appl.,* **8**, No. 4, 375–382, 1998.
4. **V. V. Kochergin,** The theory of rectifier circuits (the present state), in *Diskretnaya matematika i eyo prilozheniya* (Discrete Mathematics and Its Applications), Vol. 7, pp. 23–40, Izd. IPM RAN, Moscow, 2013.
5. **O. B. Lupanov,** On rectifier and switching-and-rectifier schemes, *Dokl. Akad. Nauk SSSR,* **111**, No. 6, 1171–1174, 1956.
6. **E. I. Nechiporuk,** Rectifier networks, *Dokl. Akad. Nauk SSSR,* **148**, No. 1, 50–53, 1963. Translated in *Sov. Phys., Dokl.,* **8**, 5–7, 1963.
7. **E. I. Nechiporuk,** On the topological principles of self-correction, in *Problemy kibernetiki* (Problems of Cybernetics), Vol. 21, pp. 5–102, Nauka, Moscow, 1969. Translated in *Systems Theory Res.,* **21**, 1–99, 1970.

8. A. Brauer, On addition chains, *Bull. AMS,* **45**, 736–739, 1939.

9. S. Jukna and **I. S. Sergeev,** Complexity of linear Boolean operators, *Found. Trends Theor. Comput. Sci.,* **9**, No. 1, 1–123, 2013.

10. N. Pippenger, The minimum number of edges in graphs with prescribed paths, *Math. Syst. Theory,* **12**, 325–346, 1979.

11. N. Pippenger, On the evaluation of powers and monomials, *SIAM J. Comput.,* **9**, No. 2, 230–250, 1980.

Igor S. Sergeev

Received
19 April 2017
Revised
4 September 2017