

Федеральное государственное бюджетное учреждение науки  
Научно-Исследовательский Институт Системных Исследований РАН

На правах рукописи

УДК

Карандашев Яков Михайлович

**МЕТОД ТРАНСФОРМАЦИИ МАТРИЦЫ В ЗАДАЧЕ  
КВАДРАТИЧНОЙ БИНАРНОЙ ОПТИМИЗАЦИИ**

05.13.01 – «Системный анализ, управление и обработка информации»  
(по математическим отраслям и информатике)

АВТОРЕФЕРАТ

диссертации на соискание ученой степени  
кандидата физико-математических наук

Москва

2013

Работа выполнена в Федеральном государственном бюджетном учреждении науки Научно-исследовательском институте системных исследований РАН

Научный руководитель: член-корреспондент РАН  
доктор физико-математических наук  
Крыжановский Борис Владимирович

Официальные оппоненты: звание, уч. степень ФИО  
звание, уч. степень ФИО

Ведущая организация:

Защита состоится ... (число, месяц, год) в 16 часов на заседании диссертационного совета Д 002.265.01 при НИИСИ РАН по адресу 117218, Москва, Нахимовский проспект, д. 36, корп. 1, конференц-зал.

С диссертацией можно ознакомиться в библиотеке НИИСИ РАН (комн. 13-21а)

Автореферат разослан ... (число, месяц, год)

Ученый секретарь  
диссертационного совета,  
кандидат физико-математических  
наук, доцент

В.Б. Демидович

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

### Обзор тематики

Настоящая работа направлена на повышение эффективности процедуры случайного поиска, применяемой при решении задач квадратичной бинарной минимизации. Квадратичный функционал определяет энергию системы взаимодействующих спинов в магнитной решётке. В случае случайных положительных и отрицательных взаимодействий такая система в физике называется спиновым стеклом и характеризуется беспорядком и фрустрацией, приводящей к очень сложной энергетической поверхности. Поэтому разнообразные алгоритмы адаптировались на протяжении многих лет специально для поиска основных состояний и для исследования спиновых стёкол при конечных температурах и при переходе к пределу нулевой температуры. Ввиду большой сложности задач разрабатывались специализированные компьютеры для их решения.

Задача минимизации в пространстве бинарных переменных кардинально отличается от той же задачи в непрерывном пространстве с вещественными переменными, когда она может быть решена с применением известных вычислительных процедур. В случае бинарных переменных квадратичный функционал является многоэкстремальным, даже если матрица является положительно-определенной.

В случае когда система является ферромагнетиком, т.е. взаимодействия имеют все одинаковый положительный знак, задача отыскания основного состояния является тривиальной, поскольку решением является состояние, в котором все спины сонаправлены (все  $+1$  или все  $-1$ ). Даже в случае добавления к функционалу линейного члена (внешнее случайное поле), задача для ферромагнетика остаётся полиномиально разрешимой при помощи алгоритмов, основанных на методах максимального потока в сети.

В случае двумерного спинового стекла без внешнего поля и с периодическим граничным условием не более чем в одном измерении существуют эффективные полиномиальные алгоритмы отыскания основного состояния, многие из которых основаны на паросочетаниях в графах. Полиномиально разрешима также двумерная решётка с периодическими граничными условиями в обоих направлениях, если связи имеют вид  $\pm J$ . Для гауссовых связей вопрос до сих пор открыт. Как только мы добавляем в систему внешнее поле, т.е. линейный член в функционал энергии, задача становится NP-сложной для всех типов взаимодействий. Для трёхмерной кубической решётки задача является NP-сложной даже в отсутствии внешнего поля.

Известно, что задача квадратичной бинарной оптимизации напрямую связана с задачами на графах. Задачи о разрезах (MIN-CUT и MAX-CUT) имеют множество применений, например в разметке

электронных схем, в повышении надёжности сетей, распознавании зрительных сцен и др. Для неотрицательных весов рёбер, задача MIN-CUT может быть решена при помощи метода потоков в сети благодаря дуальности максимального потока и минимального разреза в сети (алгоритм Форда-Фалкерсона), либо при помощи алгоритма Штора-Вагнера. Задача MAX-CUT для некоторых классов графов, например, как сказано выше, для планарных графов, для графов не содержащих  $K_5$  (полный граф из пяти вершин), для слабо двудольных графов, в случае целочисленных весов когда род графа ограничен константой и значения весов рёбер ограничены по абсолютному значению полиномом от размера графа, а также для некоторых других случаев имеет полиномиальные алгоритмы решения. Для общего случая весов рёбер, задача MAX-CUT (как и задача MIN-CUT, получающаяся инверсией знаков весов рёбер) является NP-сложной.

*Строгие алгоритмы.* Для решения задачи в общем случае предложено несколько алгоритмов. Простейший метод состоит в перенумерации всех  $2^N$  возможных состояний и очевидно имеет экспоненциальное время работы. Но даже кубическая решётка состоящая из  $4^3$  элементов уже оказывается слишком большой.

Для решения NP-сложных задач предложены алгоритмы  $\varepsilon$ -аппроксимации, т.е. полиномиальные алгоритмы, строго дающие решение, отличающееся не более чем на  $\varepsilon$  от энергии глобального минимума. Например, в работе Goemans and Williamson (1994) был предложен алгоритм, который за полиномиальное время даёт решение задачи MAX-CUT со значением, составляющим не менее 0.878 от величины максимального разреза, т.е. отличающимся менее чем на 13% от оптимума. С другой стороны, как следует из работы Bellare et al (1995), в предположении  $P \neq NP$ , не существует полиномиального алгоритма, который доказуемо даёт решение со значением разреза, составляющим хотя бы 98% от оптимального. Исследование алгоритмов  $\varepsilon$ -аппроксимации привело к открытию методов сведения задачи бинарной оптимизации к линейному и полуопределённому программированию (так называемые LP- и SDP-релаксации), которые обычно применяются в рамках метода ветвей и границ.

*Эвристические методы, локальный поиск.* Для решения задач больших размерностей, а также с более плотными матрицами применяются эвристические методы. Широко распространённым в физике подходом является эволюция системы при некоторой температуре используя метод Монте-Карло (MC). На этом основаны методы симуляции отжига (SA), а также более изошрёпный алгоритм параллельной закалки (PT). Предлагались подходы основанные на нейронных сетях. С разным успехом применялись варианты генетических алгоритмов. Самыми простыми и потому наиболее популярными эвристиками являются локальные одношаговые

алгоритмы поиска. В настоящей работе в качестве основной процедуры локальной минимизации выбрана асинхронная динамика модели Хопфилда.

### **Актуальность темы**

Несмотря на многочисленные работы по данной тематике (см. выше), NP-полнота задачи квадратичной оптимизации с бинарными переменными оставляет эту задачу трудно решаемой. Линейное увеличение скорости вычислений компьютеров (например, при помощи графических ускорителей) не может превозмочь экспоненциальную от размера входных данных сложность задачи. Поэтому задача остаётся актуальной (и вероятно будет оставаться такой вплоть до создания квантовых вычислителей).

### **Цели и задачи диссертации**

Диссертационная работа направлена на исследование способов трансформации энергетической поверхности функционала с целью резкого увеличения эффективности алгоритмов минимизации.

Для достижения данной цели ставятся следующие задачи:

1) Исследовать энергетическую поверхность функционала, в частности форму локальных минимумов.

2) Исследовать зависимость вероятности отыскания минимума при случайном поиске от глубины минимума.

3) Исследовать влияние трансформации матрицы функционала на изменение глубины минимумов, сдвиг минимумов в пространстве и смещение спектра находимых минимумов.

4) На основе проведенных исследований построить эффективный алгоритм минимизации квадратичного функционала.

### **Основные положения выносимые на защиту**

1) Исследованы свойства энергетической поверхности, описываемой квадратичным функционалом в  $N$ -мерном пространстве состояний с бинарными переменными. Получено выражение для формы минимума, её зависимость от глубины минимума.

2) Радиус области притяжения минимума  $x_0$  полностью определяется глубиной минимума  $\gamma = |E|\sqrt{N}$  из уравнения:

$$\frac{x_0}{1-x_0} = \Phi[\gamma(1-2x_0)],$$

где  $\Phi(\cdot)$  - интеграл вероятностей. Поскольку спектр энергий минимумов имеет узкое распределение и глубина минимумов отличается от среднего значения глубины  $\bar{\gamma}$  не более чем на 10-20%, то можно считать зависимость радиуса от глубины линейной:  $x_0 \approx \bar{x} + k\bar{x}(\gamma - \bar{\gamma})$ , где  $\bar{x}$  и  $k$  - константы (порядка  $\bar{x} \approx 0.384$  и

$k \approx 0.076$ ). Поэтому вероятность отыскания минимума при случайном поиске, равная отношению объёму области притяжения минимума во всём пространстве конфигураций, экспоненциально зависит от глубины минимума  $P(\gamma) = x^N / 2^N \approx \bar{P} e^{kN(\gamma-\bar{\gamma})}$ .

3) Показано, что возведение матрицы  $T$  функционала в степень  $k$  ( $k = 2, 3, 4, \dots$ ) меняет энергетическую поверхность функционала, однако наиболее глубокие минимумы исходного функционала  $E$  остаются минимумами в новом функционале  $E_k$ , либо присутствуют минимумы в их малой  $d$ -окрестности (при  $k=3$  сдвиг  $d \approx 0.026N$ ) и при этом оказываются глубже (для матриц 2D EA при  $k=3$  происходит углубление минимума, соответствующего глобальному на исходном функционале, примерно на 15%).

4) При сложении матрицы  $T$  и матрицы  $T^2$  с весами  $(1-z)$  и  $z$ , соответственно, получается функционал  $E_z$ , который обладает теми же свойствами, что и  $E_k$ , только в этом случае сдвиг и углубление минимумов контролируются параметром  $z$  ( $z$  от 0 до 1). Получены выражения для величины углубления минимума и его сдвига в пространстве. Показано, что существует оптимальное значение  $z_{opt} \approx 0.51$ , при котором происходит максимальное углубление (порядка 45%), при этом сдвиг оказывается около  $\bar{d}_{opt} = 0.01N$ .

5) На основании двух предложенных выше трансформаций, в работе предложены два алгоритма минимизации (DDK и MM), состоящие из двух шагов. На первом шаге происходит минимизация трансформированного функционала (спуск по поверхности  $E_k$  в алгоритме DDK и спуск по поверхности  $E_z$  в алгоритме MM). Далее из найденной точки процедура минимизации продолжается на исходном функционале (продолжение спуска по поверхности  $E$ ). Такая двухэтапная процедура алгоритмов позволяет находить минимумы исходного функционала, нивелируя сдвиги, введённые трансформацией поверхности.

6) Показано, что использование предложенных алгоритмов (DDK и MM) приводит к существенному сдвигу спектра находимых минимумов в сторону глобального минимума и к экспоненциальному по  $N$  росту эффективности алгоритма случайного поиска, при этом количество операций, затрачиваемых на один старт, увеличивается вдвое. Дано строгое обоснование эффективности предложенных алгоритмов для полностью связанных матриц модели Шеррингтона-Киркпатрика; применение алгоритмов для другого типа матриц является эвристическим. Эффективность предложенных двухэтапных алгоритмов проверена на нескольких типах случайных матриц и продемонстрирована сравнением с известными алгоритмами на задачах MaxCut и Graph Bipartitioning.

## Методы исследования

Для целей исследования были выбраны матрицы SK-модели – модели Шеррингтона-Киркпатрика, в которой матричные элементы можно рассматривать как независимые случайные величины с нулевым средним. Для таких типов матриц удалось сделать некоторые теоретические оценки, пользуясь центральной предельной теоремой. Также разработанные алгоритмы проверялись на разреженных матрицах моделей Эдвардса-Андерсона (2D EA и 3D EA модели, описывающие двумерные и трехмерные спиновые решетки с периодическими граничными условиями, в которых взаимодействие отлично от нуля только между ближайшими соседями в решётке и описывается нормально распределенными матричными элементами).

Анализ энергетической поверхности и формы минимумов производился при помощи разложения матрицы  $T$  на две части: той, которая полагается ответственной за существование минимума ( $T_0$ ), и оставшейся части ( $T_1$ ). При этом матрицы  $T_0$  и  $T_1$  оказываются некоррелированными (подробнее см. главу 2). Исходя из некоррелированности этих матриц *делалось допущение об их независимости*, и на основании этого матрица  $T_1$  полагалась чисто случайной по отношению к минимуму. Данное допущение позволило сделать важные оценки, которые во многом соответствовали эксперименту. В том числе вероятность сдвига конфигурации вычислялась путём поиска отношения «сигнал/шум», где сигнальной части соответствовала матрица  $T_0$ , а шумовой – матрица  $T_1$ .

При решении поставленных задач в данной работе были использованы методы системного анализа, вычислительной математики, комбинаторики, математической статистики и теории вероятностей. Вычислительные эксперименты проводились на персональном компьютере (Intel Core i3 CPU 3.08ГГц, 2 Гб ОЗУ), и обычно занимали от нескольких минут до нескольких дней счёта (в зависимости от размерности задачи, числа стартов при случайном поиске и количества матриц, на которых набиралась статистика).

## Научная новизна

Обычно эффективность процедуры случайного поиска пытаются повысить видоизменяя динамику спуска по поверхности, описываемой функционалом  $E(S)$ . В отличие от этого, в настоящей работе предлагается изменять не динамику спуска по поверхности, а трансформировать саму поверхность так, чтобы увеличился радиус области притяжения глобального минимума (равно как и других минимумов, сравнимых с глобальным по глубине).

## **Практическая ценность**

Благодаря предложенным в работе методам трансформации, вероятность отыскания глобального минимума возрастает как  $\exp(\alpha N)$ , где  $N$  – размерность задачи; спектр энергий находимых минимумов сильно сдвигается в глубокую сторону так, что среднее по спектру значение отличается от энергии глобального минимума всего лишь на 4% - 6% (независимо от размерности задачи). Разработанные методы оптимизации найдут своё применение в прикладных задачах.

## **Апробация работы и публикации**

По материалам диссертации опубликовано **22** работы, из них **10** – в российских и зарубежных журналах (в том числе **5** из перечня ВАК), **12** – в трудах конференций. Основные результаты работы докладывались на конференциях:

1. XI Всероссийская научно-техническая конференция **Нейроинформатика 2009**. Москва, МИФИ (2009)
2. **ICONS 2009**: The 2nd IFAC International Conference on Intelligent Control Systems and Signal Processing. Turkey, Istanbul (2009)
3. **ИКТМР-2009**: Международная конференция Актуальные проблемы информационно-компьютерных технологий мехатроники и робототехники. Дивноморск (2009)
4. **МСО – 2009**: Всероссийская конференция Методы и средства обработки информации. Москва, МГУ (2009)
5. **ММРО-14**: 14-ая Всероссийская конференция Математические Методы Распознавания Образов. Суздаль (2009)
6. 52-я научная конференция **МФТИ**. Долгопрудный, МФТИ (2009)
7. X Всероссийская научно-техническая конференция **Нейроинформатика-2010**. Москва, МИФИ (2010)
8. **IJCNN 2010**: The International Joint Conference on Neural Networks. Spain, Barcelona (2010)
9. **ICANN 2010**: The 20 International Conference on Artificial Neural Networks, Greece, Thessaloniki (2010)
10. **ICANN 2011**: The 21st International Conference on Artificial Neural Networks. Finland, Espoo (2011).
11. **ICCGI 2012**: The Seventh International Multi-Conference on Computing in the Global Information Technology. Italy, Venice (2012)
12. **ICANN 2012**: The 22 International Conference on Artificial Neural Networks. Switzerland, Lausanne (2012)

## **Объём и структура диссертации.**

Диссертация состоит из **5** глав (включая **введение**), **заклучения** и **списка литературы**. Список использованной литературы содержит **85** наименований. Текст диссертации содержит **122** страницы машинописного текста, включая **31** рисунок, **9** таблиц.



## КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

В **первой главе** даётся введение в задачу минимизации квадратичного бинарного функционала  $E(S)$ :

$$E(S) = -\frac{1}{N^2 \sigma_T} \sum_{i=1}^N \sum_{j=1}^N T_{ij} s_i s_j, \quad (1)$$

построенного на заданной  $N \times N$ -матрице  $T$  в  $N$ -мерном пространстве состояний  $S = (s_1, s_2, \dots, s_N)$  с бинарными переменными  $s_i = \pm 1$ ;  $\sigma_T$  - стандартное отклонение матричных элементов  $T_{ij}$ .

В качестве основной процедуры локальной минимизации выбрана асинхронная динамика модели Хопфилда. Стандартная (асинхронная) динамика минимизации состоит в следующем: вычисляется величина локального поля  $h_i \sim -\partial E(S) / \partial s_i$ , действующего на произвольно выбранный  $i$ -й спин:

$$h_i = \sum_{j \neq i}^N T_{ij} s_j \quad (2)$$

Если спин направлен вдоль поля ( $h_i s_i \geq 0$ ), то его состояние не изменяется. Если спин находится в неустойчивом состоянии ( $h_i s_i < 0$ ), то он разворачивается вдоль поля, принимая состояние  $s_i = \text{sign } h_i$ . Эта процедура последовательно применяется ко всем спинам до тех пор, пока система не конвергирует в устойчивое состояние. Уменьшение энергии при каждом перевороте неустойчивого спина гарантирует нам, что процесс, описываемый асинхронной динамикой, за конечное число шагов приводит систему в устойчивое положение с минимумом энергии. Конечно, найденный минимум скорее всего окажется локальным, в то время как нам бы хотелось найти самый глубокий (глобальный) минимум функционала. Стандартным приёмом здесь является случайный поиск (Standard Random Search (SRS)), суть которого состоит в многократном повторении описанных выше спусков из различных случайных стартовых конфигураций.

В **главе 2** показано, что с ростом глубины минимума растёт радиус бассейна притяжения, т.е. возрастает вероятность его отыскания при случайном поиске.

Пусть  $S_0 = (s_1^{(0)}, s_2^{(0)}, \dots, s_N^{(0)})$  - конфигурация, соответствующая некоторому минимуму  $E_0 = E(S_0)$ . Совокупность точек, находящихся на расстоянии  $n$  (по Хеммингу) от  $S_0$ , назовём  $n$ -окрестностью точки  $S_0$ . Тогда, объём области притяжения минимума (число всех точек пространства, стартуя с которых спускаемся в  $S_0$ ) равен:

$$V = 2 \sum_{n=1}^{N/2} \binom{N}{n} P_n \quad (3)$$

где  $P_n$  - вероятность попадания из  $n$ -окрестности в минимум  $S_0$ , которую мы оценим ниже.

Вычислим вероятность перехода из  $n$ -окрестности в  $(n-1)$ -окрестность, которая совпадает с вероятностью события  $h_i(S_n)s_i^{(0)} > 0$ . Как показано в работе, величина  $h_i(S_n)s_i^{(0)}$  представляет собой нормально распределенную величину со средним  $(N-2n)|E_0|$  и дисперсией  $N\sigma_T^2$ . Соответственно, для искомой вероятности перехода из  $n$ - в  $(n-1)$ -окрестность получим  $\Pr\{h_i(S_n)s_i^{(0)} > 0\} = \Phi(\gamma_n)$ , где:

$$\Phi(\gamma_n) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\gamma_n} e^{-\frac{1}{2}z^2} dz, \quad \gamma_n = \gamma(1 - 2n/N), \quad (4)$$

где  $\gamma = |E_0|\sqrt{N}$  - глубина минимума  $S_0$ . Тогда выражение для вероятности конвергенции из  $S_n$  в минимум  $S_0$  примет вид:  $P_n = \Phi(\gamma_n) \cdot \Phi(\gamma_{n-1}) \cdot \dots \cdot \Phi(\gamma_1)$ . Подставляя это выражение в (3), используя соотношение Стирлинга и переходя от суммирования к интегрированию по переменной  $x = n/N$  для объема области притяжения получим:

$$V = \sqrt{2N} \int_0^{0.5} \frac{e^{NF(x)}}{\sqrt{\pi x(1-x)}} dx, \quad (5)$$

где

$$F(x) = -x \ln x - (1-x) \ln(1-x) + \int_0^x \ln[\Phi(\gamma_z)] dz, \quad \gamma_z = \gamma(1 - 2z). \quad (6)$$

Интеграл (6) оценим методом перевала. Приравняв производную  $F(x)$  к нулю, получим уравнение для седловой точки  $x_0$ :

$$\frac{x_0}{1-x_0} = \Phi[\gamma(1 - 2x_0)] \quad (7)$$

и выражение (5) преобразуется к виду:

$$V = 2e^{NF(x_0)} \quad (8)$$

Величина  $x_0$  фактически представляет собой радиус области притяжения минимума, поскольку объем области притяжения с точностью до несущественного коэффициента совпадает с величиной  $x_0^N$ . Анализ выражений (7-8) показывает, что с ростом глубины минимума  $\gamma$  радиус области притяжения возрастает. Одновременно экспоненциально возрастает и объем области притяжения. Чтобы убедиться в этом, получим из (7-8) более простые оценочные выражения. Эксперимент показывает, что глубина среднестатистического минимума  $\bar{\gamma}$  отличается от глубины

глобального минимума  $\gamma_0$  всего лишь на 10-15%. Это позволяет для относительно глубоких минимумов ( $\bar{\gamma} < \gamma \leq \gamma_0$ ) представить функцию  $x_0 = x_0(\gamma)$  в виде разложения в окрестности значения  $\bar{\gamma}$ , которое наиболее легко определяется в эксперименте. Ограничиваясь линейным членом разложения, получим оценочное выражение:

$$x_0 \approx \bar{x} + k\bar{x}(\gamma - \bar{\gamma}) \quad (9)$$

где  $\bar{x}$  - радиус притяжения для минимума среднестатистической глубины  $\bar{\gamma}$ , а  $k$  - коэффициент, слабо зависящий от типа матрицы. Аналогично, проводя в (8) разложение в окрестности величины  $\bar{\gamma}$ , для вероятности отыскания минимума  $P(\gamma) = V / 2^N$  получим:

$$P(\gamma) = \bar{P} e^{kN(\gamma - \bar{\gamma})} \quad (10)$$

где  $\bar{P} = P(\bar{\gamma})$  есть вероятность отыскания минимума среднестатистической глубины  $\bar{\gamma}$ .

Для некоторого среднестатистического минимума, характерного для матриц модели Шеррингтона-Киркпатрика (СК-модель), глубина равна  $\bar{\gamma} = 1.36$ . Из (7) и (9) получаем для него  $\bar{x} \approx 0.384$  и  $k \approx 0.076$ . Это означает, что при случайном поиске большая часть точек, из которых система конвергирует в данный минимум, расположена в его  $n$ -окрестности, где  $n \approx 0.384N$ . Соответственно, вероятность отыскания минимума глубины  $\gamma$  равна  $P \approx \bar{P} e^{0.076N(\gamma - \bar{\gamma})}$ . Как видим, чем глубже минимум, тем больше вероятность его отыскания, причём зависимость экспоненциальная. Данная зависимость выполняется не строго, а лишь в статистическом смысле. При небольших значениях  $N$  её можно подтвердить экспериментально.

Такую жесткую связь между глубиной минимума и вероятностью его отыскания несложно понять из следующих соображений. Форма энергетической поверхности вокруг минимума описывается выражениями:

$$\bar{E}_0(n) \approx E_0 \left(1 - \frac{2n}{N}\right)^2, \quad \sigma_0(n) \approx \frac{2}{N} \sqrt{\frac{2n}{N} \left(1 - \frac{n}{N}\right)} \quad (11)$$

где  $\bar{E}_0(n)$  и  $\sigma_0(n)$  - соответственно среднее значение и стандартное отклонение энергии в  $n$ -окрестности минимума. Энергетическая поверхность (11) очень гладкая, поскольку дисперсия энергии в  $n$ -окрестности с ростом размерности стремится к нулю как  $\sigma \sim 1/N$ . Более того, все минимумы имеют одинаковую форму, задаваемую выражением (11), причём ширина минимума (на полувысоте) прямо пропорциональна его глубине  $E_0$ . Это и означает, что число точек внутри области притяжения экспоненциально возрастает с ростом глубины минимума.

Как показано в работе, любую матрицу можно представить в виде взвешенного разложения по внешним произведениям случайных конфигурационных векторов:

$$T = \sigma_T \sum_{m=0}^{\infty} r_m S_m^+ S_m, \quad \sum r_m^2 = 1 \quad (12)$$

Для такого типа матриц справедливы все выводы, полученные в [главе 5](#), на основании статистической физики. В частности, любой из векторов  $S_m$ , присутствующий в разложении матрицы  $T$ , будет минимумом функционала только в том случае, если его статвес  $r_m$  больше критического значения:

$$r_c = \frac{1}{2\sqrt{\alpha_c N}} \quad (13)$$

где  $\alpha_c \approx 0.138$  - критическое значение параметра загрузки в модели Хопфилда. Это утверждение прежде всего относится к точке  $S_0$ , которая по определению является минимумом функционала (1), и для нее справедливы соотношения:

$$1 \geq r_0 \geq r_c, \quad -1 \leq E_0 \leq E_c, \quad E_c = -r_c \quad (14)$$

Ограничение на  $r_0$  сверху очевидно и неинтересно. В подавляющем большинстве минимизационных приложений реализуется ситуация  $r_0 \sim r_c$  и  $E_0 \sim E_c$ , когда вероятность отыскания глобального минимума чрезвычайно мала.

В [третьей главе](#) рассмотрен простейший вид трансформации энергетической поверхности функционала. Подставим в выражение (1) матрицу  $M = T^k$ , получаемую возведением матрицы  $T$  в степень  $k$  с последующим обнулением диагональных элементов. Тогда поверхность, описываемая функционалом  $E(S)$ , трансформируется в поверхность, описываемую функционалом:

$$E_k(S) = -c_k \sum_{i=1}^N \sum_{j=1}^N M_{ij} s_i s_j, \quad c_k = \frac{1}{\sigma_M N(N-1)} \quad (15)$$

где  $\sigma_M$  - стандартное отклонение элементов матрицы  $M$ . Все рассуждения проведем на примере трансформации глобального минимума. Очевидно, что при трансформации поверхности глобальный минимум сдвигается в пространстве, одновременно изменяются его глубина и ширина области притяжения. Ниже мы покажем, что при относительно небольших значениях параметра  $k$  ( $2 \leq k \leq 5$ ) указанная трансформация приводит к значительному увеличению глубины минимума при относительно небольшом сдвиге.

Чтобы нивелировать сдвиг, предлагается следующий двухэтапный алгоритм минимизации. На первом этапе осуществляется спуск по поверхности  $E_k(S)$  и находится

конфигурация  $S_m^{(k)}$ , являющаяся минимумом функционала  $E_k(S)$ . На втором этапе проводится коррекция: из точки  $S_m^{(k)}$  по поверхности  $E(S)$  спускаемся в ближайший минимум  $S_m$  функционала  $E(S)$ . Алгоритм спуска по поверхности  $E_k(S)$  обычный: рассчитывается локальное поле  $h_i^{(k)} = -\partial E_k(S) / \partial s_i$  на  $i$ -м спине:

$$h_i^{(k)} = \sum_{j \neq i}^N M_{ij} s_j \quad (16)$$

и при  $h_i^{(k)} s_i < 0$  состояние спина обновляется по правилу  $s_i = \text{sgn } h_i^{(k)}$ .

На основе предлагаемого двухэтапного спуска мы реализовали алгоритм случайного поиска. В дальнейшем мы будем обозначать его аббревиатурой **DDK** (Double Descent Algorithm с параметром  $k = 2, 3, \dots$ ) и сравнивать его эффективность со стандартным алгоритмом случайного поиска **SRS**. Во избежание недоразумений отметим, что при  $k = 1$  трансформированный функционал  $E_k(S)$  совпадает с исходным  $E(S)$ , а предлагаемый алгоритм DDK в этом случае есть не что иное, как стандартный алгоритм случайного поиска SRS ( $DD1 \equiv SRS$ ).

Обоснование алгоритма DDK проведем на примере случая  $k = 2$ . Покажем, что в результате трансформации поверхности глобальный минимум становится заметно глубже, а его сдвиг в пространстве невелик.

*Углубление минимумов.* Рассмотрим энергию  $E_{20} = E_2(S_0)$  в точке  $S_0$ . Как показано в работе, в пределе  $N \gg 1$  энергию  $E_{20}$  можно рассматривать как нормально распределенную величину со средним  $\bar{E}_{20}$  и относительно малым стандартным отклонением  $\sigma_E$ :

$$\bar{E}_{20} = -\sqrt{N} E_0^2, \quad \sigma_E = (1 - r_0^2) / N \quad (17)$$

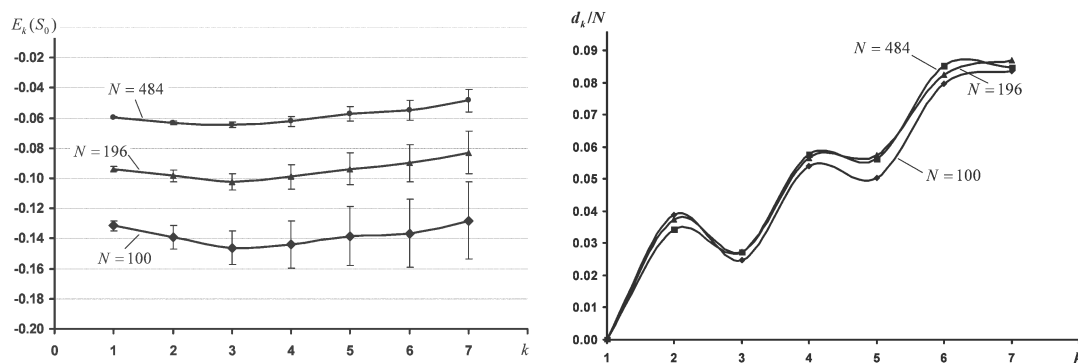
Поскольку  $\bar{E}_{20} / E_0 = r_0 \sqrt{N} \geq r_c \sqrt{N} \approx 1.35$ , то следует ожидать, что с ростом  $N$  вероятность углубления  $\Pr\{E_{20} < E_0\}$  асимптотически стремится к единице. Иными словами, трансформация поверхности с подавляющей вероятностью приводит к существенному увеличению глубины минимума (примерно в 1.35 раз), что в соответствии со сказанным в [главе 2](#) должно приводить к экспоненциальному по  $N$  возрастанию вероятности отыскания глобального минимума.

*Сдвиг минимума.* Среднюю величину сдвига можно представить в виде  $d_k = NP$ , где  $P = \Pr\{s_i^{(0)} h_i^{(k)} < 0 \mid h_i s_i^{(0)} > 0\}$  - вероятность того, что направления спина  $s_i^{(0)}$  и локального поля  $h_i^{(k)}$  в конфигурации минимума  $S_0$  не совпадают. Для рассматриваемого здесь случая  $k = 2$  эта вероятность выражается через функцию ошибок:

$$P = \frac{1}{P_0 \sqrt{\pi}} \int_0^{+\infty} e^{-(x-\gamma)^2} \operatorname{erfc}(\sqrt{2}\gamma x) dx \quad (18)$$

где  $P_0 = 1 + \operatorname{erf} \gamma$  и  $\gamma = \sqrt{N} r_0$ .

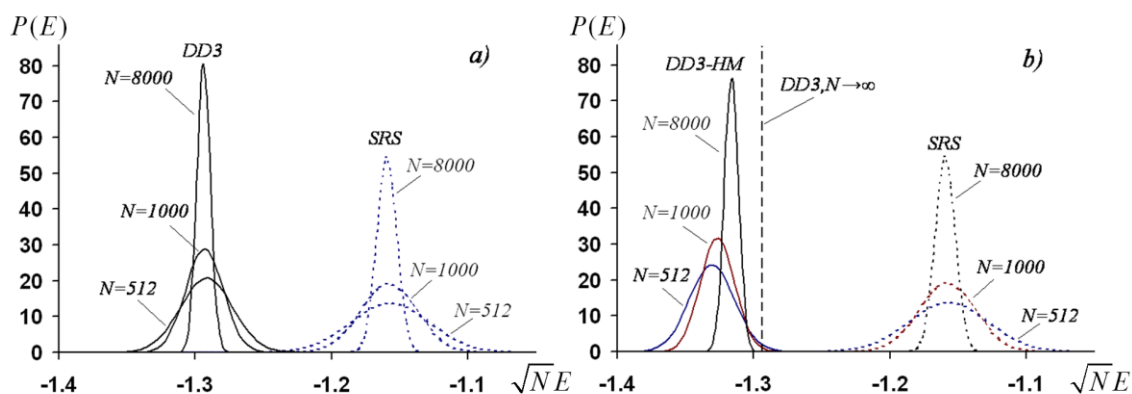
**Рис. 1 слева** показывает наблюдаемое углубление глобальных минимумов. Как видим, глобальные минимумы становятся глубже для  $k = 2, 3, 4$  и даже для  $k = 5$ , однако с увеличивающейся дисперсией. Наибольшее среднее углубление наблюдается для  $k = 3$ . Вызываемый трансформацией сдвиг минимума в пространстве относительно невелик (**рис. 1 справа**). Видно, что наименьший сдвиг также наблюдается при  $k = 3$ .



**Рис. 1** Углубление и сдвиг глобальных минимумов при трансформации поверхности при возведении матрицы в степень  $k = 2, 3, \dots, 7$ . Матрицы 2D EA размера  $N = 100, 196, 484$ . Планки слева показывают ст. отклонение.

**Эффективность алгоритма DDK.** Для иллюстрации работоспособности предложенных идей мы выбрали матрицы модели 3-мерной модели Эдвардса-Андерсона размера  $N = L^3$  при  $L = 8, 10, 20$ . Эксперименты включали в себя  $10^6$  стартов для динамики Хопфилда и  $10^4$  стартов для динамики Хоудайера-Мартина (НМ) (см. дальше). Результаты были усреднены по 100 матрицам каждого размера. На **рис. 2а** показана типичная картина распределения плотности вероятности для алгоритмов DD3 и SRS. Как видно из рисунка, кривые, полученные для достаточно больших значений  $N$  (больше 500), хорошо аппроксимируются гауссовым распределением с параметрами: 1) для DD3 среднее значение  $\bar{E}_3 \approx -1.29 / \sqrt{N}$  и стандартное отклонение  $\sigma_3 \approx 0.41 / N$ ; 2) для SRS среднее  $\bar{E}_1 \approx -1.16 / \sqrt{N}$  и стандартное отклонение  $\sigma_1 \approx 0.66 / N$ .

Алгоритм DD3 имеет существенное преимущество перед SRS. При  $N \geq 1000$ , ни в одном из экспериментов алгоритм SRS не нашёл минимум с энергией  $E < \bar{E}_3$ . Этот результат вполне ожидаем поскольку вероятность такого события оказывается  $P < 1.3 \cdot 10^{-10}$  для уже при  $N = 1000$ , т.е. выигрыш составляет 10 порядков!



**Рис. 2.** а) Плотность вероятности  $P(E)$  нахождения минимумов с помощью алгоритма SRS (точечная кривая) и с помощью алгоритма DD3 (сплошная кривая); б) Сплошные кривые показывают плотность вероятности  $P(E)$ , получаемую модифицированным алгоритмом случайного поиска DD3-НМ; для сравнения, точечная кривая показывает плотность вероятности, получаемая стандартным алгоритмом SRS. Данные для 3D EA модели с  $L = 8, 10, 20$ . По оси абсцисс отложена энергия помноженная на  $\sqrt{N}$ .

Также в работе мы попробовали вставить динамику спуска Хоудайера-Мартина (НМ) вместо динамики Хопфилда, с целью улучшить алгоритм за счёт выхода из мелких каверн.

Алгоритм DD3-НМ позволяет найти глобальный минимум для 2D EA матриц размера  $N = 100$  с вероятностью примерно 6%, которая в 30 раз выше чем вероятность (порядка 0.2%) даваемая алгоритмом DD3 и в 25000 раз больше чем вероятность ( $\approx 2.4 \cdot 10^{-6}$ ) даваемая алгоритмом SRS. Результаты для DD3-НМ для больших размерностей показаны на [рис. 2б](#). Пики описываются гауссовым распределением со средним и стандартным отклонением:

$$\bar{E} \approx -\frac{\varepsilon}{\sqrt{N}}, \quad \sigma \approx \frac{t}{N},$$

где  $\varepsilon = 1.331, 1.327, 1.315$  и  $t = 0.37, 0.40, 0.46$  для  $N = 512, 1000, 8000$  соответственно. Ясно, что алгоритм DD3-НМ показывает гораздо лучшие результаты чем DD3. Однако с ростом размерности это преимущество исчезает: пик кривой плотности вероятности сдвигается в сторону пунктирной линии, которая соответствует центру распределения DD3 на [рис. 2а](#) ( $\varepsilon \rightarrow 1.29$ ).

**Применения алгоритма к задаче разбиения графа.** Суть задачи о разбиении графа в следующем: пусть имеется некоторый граф  $G$ , у которого  $N$  вершин  $V = (v_1, v_2, \dots, v_N)$ , а веса рёбер описываются матрицей  $T$  размера  $N \times N$ . Требуется разбить вершины графа на две *равные* непересекающиеся части  $V_1$  и  $V_2$  так, чтобы сумма весов рёбер, связывающих разные части, была минимальна. Легко показать,

что эта задача сводится к минимизации квадратичного бинарного функционала (1) при ограничении  $\sum_{i=1}^N s_i = 0$ .

Алгоритм DDK для этой задачи на первом шаге ищет разбиение графа с матрицей связи  $T^k$ , где  $T$  - матрица связи исходного графа,  $k$  - некоторая степень ( $k = 2, 3, \dots, 5$ ). Полученное на первом шаге разбиение используется как начальное на втором шаге. Как показывают эксперименты, такой двухступенчатый алгоритм существенно повышает эффективность алгоритма разбиения графа. А именно, такая трансформация связей графа понижает среднюю энергию находимых минимумов и разница между энергией глобального минимума и средней энергией находимых минимумов уменьшается почти в 3 раза. Как следствие, вероятность отыскания близких к глобальному минимумов повышается на порядки (для матриц Изинга уже при  $N = 100$ , а также для матриц с равномерным распределением при больших размерностях начиная с  $N = 200$ ).

В **четвёртой главе** мы описали ещё один способ трансформации энергетической поверхности. Матрица исходного функционала (1) заменяется на новую матрицу, которую мы назовём *микс-матрицей*:

$$M = \frac{1}{\sigma_T} (1 - z)T + \frac{1}{\sigma_{2T}} zT^2, \quad (19)$$

где  $\sigma_{2T}$  - среднеквадратичное отклонение элементов матриц  $T^2$ . При изменении параметра  $z$  от 0 до 1 происходит переход от матрицы  $M = T$  к матрице  $M = T^2$ . Соответственно, поверхность, описываемая функционалом  $E(S)$ , трансформируется в поверхность  $E_z(S)$ :

$$E_z(S) = -\frac{1}{\sigma_M N^2} \sum_{i=1}^N \sum_{j=1}^N M_{ij} s_i s_j \quad (20)$$

где  $\sigma_M = \sigma_T \sqrt{z^2 + (1 - z)^2}$  - стандартное отклонение элементов матрицы  $M$ .

Как и в **главе 3**, на основе предложенной трансформации (19), был разработан аналогичный DDK двухэтапный алгоритм, который мы назвали ММ (Mix-Matrix algorithm). Минимизация функционала (20) на первом шаге производится как обычно. Чтобы обосновать работоспособность алгоритма ММ, мы показали, что трансформация приводит к углублению минимумов при небольшом сдвиге.

*Углубление минимума.* Рассмотрим энергию  $E_{z_0} = E_{z_0}(S_0)$  в точке  $S_0$ . После ряда преобразований получаем:

$$E_{z_0} = E_0 \frac{1 - z + zA}{\sqrt{(1 - z)^2 + z^2}}, \quad A = \frac{\sigma_\lambda^2 + \gamma^2 - 1}{\gamma} \sqrt{q / (q - 1)}, \quad (21)$$



где  $q$  – число отличных от нуля элементов в строке матрицы  $T$  ( $q = N - 1$  в SK-модели и  $q = 4$  в модели 2D EA),  $\sigma_\lambda^2$  – дисперсия величины  $\lambda_i = h_i s_i / \sqrt{N}$  в минимуме.

Из (21) следует, что имеется оптимальное значение параметра  $z$

$$z_{opt} = \frac{A}{1 + A}, \quad (22)$$

при котором углубление достигает максимального значения:

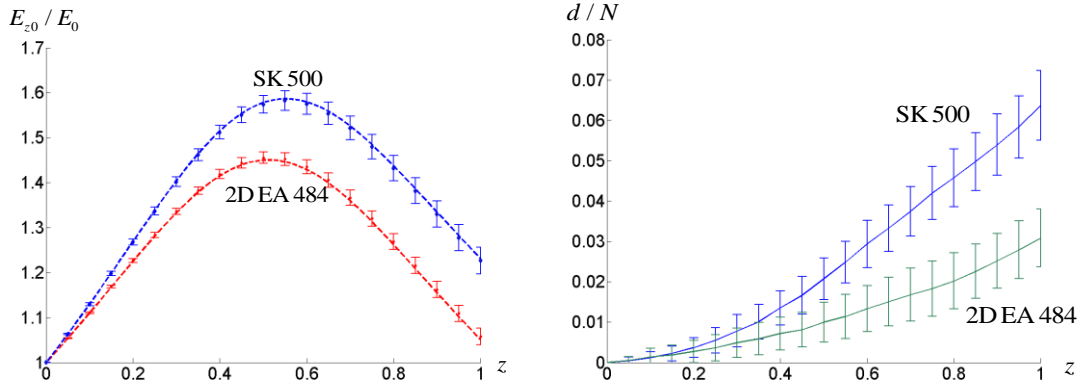
$$(E_{z_0} / E_0)_{opt} = \sqrt{1 + A^2} \quad (23)$$

Для матриц 2D EA имеем  $\gamma \approx 1.31$ ,  $\sigma_\lambda \approx 0.69$  и из (21)-(23) находим  $z_{opt} \approx 0.51$ ,  $(E_{z_0} / E_0)_{opt} \approx 1.45$ . Аналогично, для SK-модели  $\gamma \approx 1.51$ ,  $\sigma_\lambda \approx 0.76$ , значит  $z_{opt} \approx 0.55$ ,  $(E_{z_0} / E_0)_{opt} \approx 1.59$ . Видим, что оптимальное значение  $z$  находится где-то вблизи 0.5. Эксперимент подтверждает полученные соотношения (рис. 3 слева).

*Сдвиг минимума.* Аналогично (18) получаем вероятность сдвига:

$$P = \frac{1}{\sqrt{2\pi}\sigma_\lambda \Phi(\gamma / \sigma_\lambda)} \int_0^\infty dx e^{-\frac{(x-\gamma)^2}{2\sigma_\lambda^2}} \left[ 1 - \Phi\left(\frac{\alpha x + \sigma_\lambda^2 - 1}{\sigma_\lambda^2}\right) \right], \quad (24)$$

где  $\Phi(\cdot)$  - интеграл вероятностей,  $\alpha = \left[ (1 - z)\sqrt{(q-1)/q} + z\gamma \right] / z$ .



**Рис. 3** Углубление (слева) и сдвиг (справа) глобальных минимумов в зависимости от параметра  $z$ . Маркеры с планками - экспериментальные кривые: SK 500 - для  $N = 500$  модели Шеррингтона-Киркпатрика; 2D EA 484 - для матриц 2D модели Эдвардса-Андерсона размерности  $N = 22 \times 22$ . Штриховые линии (слева) – теоретические, построенные по формуле (21) с параметрами  $\gamma \approx 1.51$ ,  $\sigma_\lambda \approx 0.76$  для SK 500, и  $\gamma \approx 1.31$ ,  $\sigma_\lambda \approx 0.69$  для 2D EA 484.

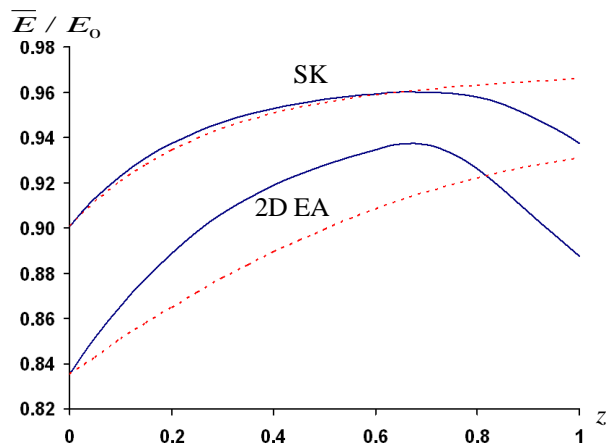
Как и следовало ожидать, формула (24) описывает монотонное увеличение величины  $\bar{d}$  с ростом  $z$ . Это подтверждается экспериментом (см. рис. 3 справа). Более того, с ростом  $z$  возрастают и флуктуации величины сдвига  $d$ . Это ожидаемый результат, поскольку аналогичные вычисления показывают, что дисперсия

сдвига описывается выражением  $\sigma^2(d) = \bar{d}$ . Значения сдвига, достигаемые при  $z = z_{opt}$ , приблизительно в 3 раза меньше чем при  $z = 1$ :  $\bar{d}_{opt} = 0.015N$  для SK-модели и  $\bar{d}_{opt} = 0.01N$  для модели 2D EA.

*Результаты для алгоритма ММ.* Для проверки эффективности предложенного алгоритма в процессе численных экспериментов строилась микс-матрица при значениях  $z$  от 0 до 1 с интервалом  $\Delta z = 0.05$ . Результаты экспериментов усреднялись по ансамблю из 100 матриц. Каждый эксперимент включал в себя по  $N_{runs} = 10^6$  стартов. В экспериментах помимо описанной выше микс-матрицы (19) мы также попробовали использовать микс-матрицу типа  $M = aT + bT^3$ .

На рис. 4 показано как меняется средняя энергия  $\bar{E}$  найденных минимумов при различных  $z$ . Результаты для 3D Изинга сливаются с результатами для 2D Изинга и потому не приведены на графике. Самым важным является то, что независимо от размерности задачи алгоритм ММ позволяет находить минимумы, всего лишь на несколько процентов отличающиеся от энергии глобального минимума. Так, для матриц двумерной и трёхмерной модели EA алгоритм SRS ( $z=0$ ) даёт минимумы с энергией на 17% хуже глобального, в то время как алгоритм ММ (при  $z=0.7$ ) даёт минимумы отличающиеся от энергии глобального минимума лишь на 6.5 - 7%. Аналогично, для матриц SK модели, алгоритм SRS не доходит до глобального минимума примерно на 10%, в то же время средняя глубина минимумов, получаемых алгоритмом ММ при  $z=0.7$ , оказывается лишь на 4% меньше глубины глобального минимума.

Пусть  $P_{SRS}$  и  $P_{MM}$  - вероятности отыскания минимумов с энергией  $E \in [-1; -0.99]$  с помощью алгоритмов SRS и ММ, соответственно. Разница между  $P_{MM}$  и  $P_{SRS}$  даже при относительно небольшой размерности задачи составляет примерно 3 порядка. При больших размерностях распределение плотности вероятности аналогично представленным на рис. 2.



**Рис. 4.** Среднее значение энергии локальных минимумов, найденных с помощью алгоритма ММ. Пунктирные линии – микс-матрица с  $T^3$ , сплошные - микс-матрица с  $T^2$ . Два типа матриц – SK и 2D EA. По оси Y величина средней энергии  $\bar{E}$  поделена на энергию глобального минимума  $E_0$  и примерно одинакова для разных  $N$

**Сравнение с алгоритмами для задачи MAXCUT.** Суть задачи о максимальном разрезе в следующем: для заданного взвешенного неориентированного графа  $G(V,W)$  требуется найти такое разбиение его вершин  $V$  на два непересекающихся подмножества  $V_1$  и  $V_2$ , чтобы сумма весов ребёр соединяющих вершины из разных подмножеств была максимальна, т.е.  $\max_{V_1, V_2} \sum_{u \in V_1} \sum_{v \in V_2} w_{uv}$ . Легко показать, что эта задача сводится к минимизации квадратичного функционала (1), построенного на матрице  $T_{ij} = -w_{ij} / 4$ .

Мы сравнили эффективность алгоритма ММ с алгоритмами, наиболее хорошо зарекомендовавшими себя для решения задачи MAX-CUT. В частности, это алгоритмы: Scatter Search [Marti et al \(2009\)](#) и CirCut [Burer et al \(2000\)](#). На сайте (<http://www.opticom.es>) выложены наборы задач MAX-CUT и результаты для этих алгоритмов, протестированных на данных наборах на компьютере примерно той же вычислительной мощности что и наш (процессор Intel 3.06 GHz).

Сравнение результатов показано в [Таблице 1](#). В первом столбце показаны названия матриц. Это матрицы того же сорта, что и матрицы 3D EA с периодическими (тороидальными) граничными условиями. Первые две задачи с гауссовыми связями  $T_{ij}$ , следующие две со связями  $T_{ij} = \pm 1$ . Размерность задач  $N = L \times L \times L$  при  $L = 15$  и  $8$ . Второй и третий столбцы показывают результаты (значение разреза и время в секундах), приведенные на сайте для алгоритмов Scatter Search и CirCut. В последнем столбце показаны усреднённые значения наилучших разрезов, найденные нашим алгоритмом за  $10^4$  стартов.

**Таблица 1.** Сравнительная таблица для 4-х примеров задачи MAXCUT.

Problem	Scatter Search		CirCut		Mix-Matrix $z=0.7, 10^4$ стартов	
	Obj Val	time	Obj Val	time	Obj Val	time
toursg3-15	281 029 888	1023	268 519 648	788	$(2.73 \pm 0.004) \cdot 10^8$	8.8
toursg3-8	40 314 704	65	41 684 814	53	$(4.10 \pm 0.010) \cdot 10^7$	2.3
tourspm3-15-50	2 964	333	2 895	427	$2 889.7 \pm 4.6$	8.9
tourspm3-8-50	442	48	454	38	$451.9 \pm 1.3$	2.2

Как видно из таблицы, наш алгоритм находит решение, сравнимое с одним из алгоритмов (CirCut или SS) и не более чем на 3% хуже другого, при том что CirCut и SS затрачивают в 20-100 раз больше времени. Наилучший результат оказался для 4-ой задачи, где наш алгоритм также нашёл в некоторых запусках решение 454, соответствующее лучшему из SS и CirCut. Более того, при увеличении числа стартов, удалось найти решение с величиной разреза 456, которое не было до сих пор получено другими алгоритмами.

В **пятой главе** подробно рассматривается проблема ограничения на глубину минимумов, вкратце рассмотренной в **главе 2**. В **главе 2** было показано, что любая матрица может быть представлена в виде квазихеббовского разложения по конфигурациям минимумов (12). При этом конфигурация  $S_m$  может быть минимумом, только если её вес  $r_m$  в разложении не меньше некоторого критического значения  $r_c$ . К сожалению, теоретически точно оценить значение  $r_c$  не удаётся. Мы пошли обратным путём и рассмотрели квазихеббовскую матрицу, *построенную* на случайных паттернах. В этом случае, пользуясь методами статистической физики, удаётся для самого общего вида распределения весов методом седловой точки получить уравнение, позволяющее исследовать свойства энергетической поверхности вблизи паттернов, в том числе получить оценки для ёмкости памяти модели Хопфилда и связанной с ней критическим весом  $r_c$ .

В качестве модели рассмотрена ситуация, когда только один вес  $r_1 = r$  отличается от всех остальных весов, тождественно равных 1:  $r_\mu \equiv 1$ ,  $\mu \geq 2$  и  $r \neq 1$ . Также исследованы три частных случая различных весов:

1) Для весов в виде гармонического ряда  $r_k = 1/k$  максимальное число запомненных паттернов невелико по сравнению с размерностью  $N$ , всего  $k_m \approx \sqrt{N / 2 \ln N}$ .

2) Для весов в виде геометрической прогрессии  $r_k = q^k$  было показано, что при некотором оптимальном значении  $q$ , достаточно близком к 1 ( $q_m \approx \sqrt{1 - \delta_0}$ , где  $\delta_0 \approx 1 / (0.165 \cdot N)$ ), число распознаваемых паттернов будет пропорционально  $N$ ,  $k_m \approx 0.05N$ .

3) Для весов в виде арифметической прогрессии, при оптимальном значении параметра шага прогрессии, получается максимальное значение загрузки  $k_m \approx 0.06N$ .

На основании числа  $k_m$ , можно сделать оценку критического значения веса  $r_c$ , ниже которого паттерны не распознаются, т.е. не являются минимумами. Так, для весов в виде гармонического ряда  $r_c$  равно  $1/k_m$ , что в точности совпадает с результатом, полученным методом «сигнал/шум»:  $r_c = \sqrt{2 \ln N / N}$ . Однако максимальное значение  $r_c$  будет в случае одинаковых весов паттернов (стандартная модель Хопфилда). Таким образом, если статвес паттерна  $r_m$  больше критического величины:  $r_c = 1 / \sqrt{\alpha_c N}$ , где  $\alpha_c \approx 0.14$  - критическое значение параметра загрузки стандартной модели Хопфилда, то он точно будет минимумом.

## ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

1) Исследованы свойства энергетической поверхности, описываемой квадратичным функционалом в  $N$ -мерном пространстве состояний с бинарными переменными. Получены выражения для формы минимумов. При случайной матрице связей средняя энергия в  $n$ -окрестности минимума описывается формулой  $\bar{E}(n) \approx E_0(1 - 2n/N)^2$ , где  $E_0$  - энергия минимума, при этом разброс относительно среднего стремится к нулю как  $1/N$ . Таким образом, форма минимума определяется лишь глубиной минимума.

2) Показано, что радиус области притяжения минимума  $x_0$  полностью определяется глубиной минимума  $\gamma = |E| \sqrt{N}$  из уравнения:

$$\frac{x_0}{1 - x_0} = \Phi[\gamma(1 - 2x_0)],$$

где  $\Phi(\cdot)$  - интеграл вероятностей. Поскольку спектр энергий минимумов имеет узкое распределение и глубина минимумов отличается от среднего значения глубины  $\bar{\gamma}$  не более чем на 10-20%, то можно считать зависимость радиуса от глубины линейной:

$$x_0 \approx \bar{x} + k \bar{x}(\gamma - \bar{\gamma}),$$

где  $\bar{x}$  и  $k$  - константы (порядка  $\bar{x} \approx 0.384$  и  $k \approx 0.076$ ). Из этого следует, что вероятность отыскания минимума при случайном поиске, равная отношению объёма области притяжения минимума во всём пространстве конфигураций, экспоненциально зависит от глубины минимума  $P(\gamma) = x^N / 2^N \approx \bar{P} e^{kN(\gamma - \bar{\gamma})}$ .

3) Показано, что возведение матрицы  $T$  в степень  $k$  ( $k = 2, 3, 4, \dots$ ) меняет энергетическую поверхность функционала, однако наиболее глубокие минимумы исходного функционала  $E$  остаются минимумами в новом функционале  $E_k$ , либо присутствуют минимумы в их малой  $d$ -окрестности (при  $k=3$  сдвиг  $d \approx 0.026N$ ) и при этом оказываются глубже (для матриц 2D EA при  $k=3$  происходит углубление минимума, соответствующего глобальному на исходном функционале, примерно на 15%).

4) При сложении матрицы  $T$  и матрицы  $T^2$  с весами  $(1 - z)$  и  $z$ , соответственно, получается новый функционал  $E_z$ , который обладает теми же свойствами, что и  $E_k$ , только в этом случае можно контролировать сдвиг и углубление минимумов, варьируя параметр  $z$  от 0 до 1. Получены выражения для величины углубления минимума и его сдвига в пространстве в зависимости от  $z$ . Показано, что существует некоторое оптимальное значение  $z_{opt} \approx 0.51$  (для матриц

2D EA), при котором происходит максимальное углубление (порядка 45%), при этом сдвиг оказывается около  $\bar{d}_{opt} = 0.01N$ .

5) На основании двух предложенных выше трансформаций функционала, в работе предложены два алгоритма минимизации (DDK и MM), состоящие из двух шагов. На первом шаге происходит минимизация трансформированного функционала (спуск по поверхности  $E_k$  в алгоритме DDK и спуск по поверхности  $E_z$  в алгоритме MM). Далее из найденной точки процедура минимизации продолжается на исходном функционале (продолжение спуска по поверхности  $E$ ). Такая двухэтапная процедура алгоритмов позволяет находить минимумы исходного функционала, нивелируя сдвиги, введённые трансформацией поверхности.

6) Показано, что использование предложенных алгоритмов (DDK и MM) приводит к существенному сдвигу спектра находимых минимумов в сторону глобального минимума и к экспоненциальному по  $N$  росту эффективности алгоритма случайного поиска, при этом количество операций, затрачиваемых на один старт, увеличивается вдвое. Дано строгое обоснование эффективности предложенных алгоритмов для полностью связанных матриц модели Шеррингтона-Киркпатрика; применение алгоритмов для другого типа матриц является эвристическим. Эффективность предложенных двухэтапных алгоритмов проверена на нескольких типах случайных матриц и продемонстрирована сравнением с известными алгоритмами на задачах MaxCut и Graph Bipartitioning.

## **РАБОТЫ АВТОРА ПО ТЕМЕ ДИССЕРТАЦИИ**

1. Я.М. Карандашев, Б.В. Крыжановский. Увеличение глубины глобального минимума квадратичного функционала путём возведения матрицы в степень. // XI Всероссийская научно-техническая конференция Нейроинформатика-2009. Москва, МИФИ (2009)
2. J.M. Karandashev, B.V. Kryzhanovsky. An Effective Increase of the Basin of a Quadratic Functional Global Minimum. // ICONS 2009: The 2nd IFAC International Conference on Intelligent Control Systems and Signal Processing. Turkey, Istanbul (2009)
3. Карандашев Я.М., Крыжановский Б.В. Эффективное увеличение области притяжения глобального минимума квадратичного бинарного функционала при нейросетевом поиске. // Международная конференция Актуальные проблемы информационно-компьютерных технологий мехатроники и робототехники (ИКТМР-2009). Дивноморск. (2009)
4. Я.М. Карандашев, Б.В. Крыжановский. Эффективное увеличение области притяжения глобального минимума квадратичного бинарного функционала. // Всероссийская конференция Методы и средства обработки информации (МСО - 2009). Москва, МГУ (2009)
5. Я.М. Карандашев, Б.В. Крыжановский. Эффективное увеличение области притяжения глобального минимума квадратичного бинарного функционала при нейросетевом поиске. //14-ая Всероссийская конференция Математические Методы Распознавания Образов (ММРО-14), Суздаль (2009)
6. Я.М. Карандашев, Б.В. Крыжановский. Улучшение нейросетевой оптимизации бинарного квадратичного функционала. // 52-я научная конференция МФТИ. Долгопрудный, МФТИ (2009)
7. Я.М. Карандашев, Б.В. Крыжановский. О трансформации энергетической поверхности в задаче бинарной оптимизации. //ДАН. Т. 429, №4, с. 465-469 (2009)
8. Ya. M. Karandashev and B. V. Kryzhanovsky. Transformation of Energy Landscape in the Problem of Binary Minimization. // Doklady Mathematics, 2009, Vol. 80, No. 3, pp. 927–931 (2009)
9. Я.М. Карандашев, Б.В. Крыжановский. Эффективное увеличение области притяжения глобального минимума квадратичного бинарного функционала при нейросетевом поиске. // Искусственный интеллект, № 4, с. 37-44 (2009)
10. Я.М. Карандашев, Б.В. Крыжановский. Повышение эффективности алгоритма Кернигана-Лина путём модификации

матрицы межсвязей. // Всероссийская конференция  
Нейроинформатика-2010. Москва, МИФИ (2010)

11. Ya. M. Karandashev, B. V. Kryzhanovsky. Binary Optimization: Efficient Increasing of Global Minimum Basin of Attraction. // *Optical Memory & Neural Networks*, Vol. 19, No. 2, pp. 110–125 (2010)
12. Ya. M. Karandashev and B. V. Kryzhanovsky. Efficient Energy Landscape Transformation in the Problem of Binary Minimization. // *The 2010 International Joint Conference on Neural Networks (IJCNN)*, Spain, Barcelona, pp. 1-6 (2010)
13. Yakov Karandashev, Boris Kryzhanovsky. Binary Minimization: Increasing the Attraction Area of the Global Minimum in the Binary Optimization Problem. // *ICANN (2) 2010, LNCS*, v. 6353/2010, pp. 525-530 (2010).
14. Iakov M. Karandashev, Boris V. Kryzhanovsky. Transformation of Edge Weights in a Graph Bipartitioning Problem. // *Proceedings of the 21st international conference on Artificial neural networks ICANN (Part 2)*, LNCS, v. 6792/2011, pp. 25-31 (2011).
15. Iakov Karandashev, Boris Kryzhanovsky. Increasing the attraction area of the global minimum in the binary optimization problem. // *arXiv:1109.0165v1*
16. Я.М. Карандашев, Б.В. Крыжановский, Л.Б. Литинский. Сильная неустойчивость спектра минимумов бинарного квадратичного функционала. *ДАН. - Т. 436, N 6. С. 733-737* (2011)
17. Ya. M. Karandashev, B. V. Kryzhanovsky and L. B. Litinskii. Strong instability of the minima spectrum of a quadratic binary functional. // *Doklady Mathematics* Vol. 83, 1, 116-120 (2011)
18. Iakov Karandashev, Boris Kryzhanovsky and Leonid Litinskii. Weighted Patterns as a Tool to Improve the Hopfield Model. // *Physical Review E* 85, 041925 (2012).
19. Я.М. Карандашев. Трансформация матрицы связей в задаче разбиения графа.// *Вестник компьютерных и информационных технологий. №1*, 33-37 (2012)
20. I. Karandashev and B. Kryzhanovsky . Mix-matrix Method in Problem of Discrete Optimization. // *ICCGI 2012 : The Seventh International Multi-Conference on Computing in the Global Information Technology*, Italy, Venice (2012)
21. Karandashev and B. Kryzhanovsky. The Mix-Matrix Method in the Problem of Binary Quadratic Optimization. // *Proc. of ICANN 2012, Part I*, LNCS 7552, pp. 41–48 (2012)
22. Karandashev and B. Kryzhanovsky. Increasing the attraction area of the global minimum in the binary optimization problem. // *J. Glob. Optim.*, V. 56, 3 , pp. 1167-1185 (2013), DOI: 10.1007/s10898-012-9947-7