

**САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ**

Математико-механический факультет

Кафедра системного программирования

Метод отжига в задачах оптимизации

Дипломная работа студента 59 группы
Лопатина Андрея Сергеевича

Научный руководитель,/О. Н. Граничин/
/подпись/

Рецензент,/Н. Н. Вояковская/
/подпись/

“Допустить к защите”
Заведующий кафедрой,
профессор/А. Н. Терехов/
/подпись/

Санкт-Петербург
2004

Содержание

1	Введение	4
2	Алгоритм метода отжига	4
3	Общие схемы метода отжига	6
3.1	Больцмановский отжиг	6
3.2	Отжиг Коши (быстрый отжиг)	6
3.3	Сверхбыстрый отжиг	7
3.4	Алгоритм Ксин Яо	8
3.5	Методы “тушения”	9
3.6	Масштабирование в ходе отжига	9
4	Моделирование схем отжига	10
4.1	Больцмановский отжиг	10
4.2	Отжиг Коши	10
4.3	Сверхбыстрый отжиг	11
4.4	Методы “тушения”	11
5	Сравнение схем отжига на задачах нахождения минимума	12
5.1	Обозначения к графикам	12
5.2	Многочлены от одной переменной	12
5.2.1	Генерация случайных многочленов	12
5.2.2	Выбор тестов	13
5.2.3	Выбор начальной точки	13
5.2.4	Выбор параметров сверхбыстрого отжига	13
5.2.5	Выбор стартовой температуры	13
5.2.6	Результаты тестирования	13
5.2.7	Анализ результатов	17
5.3	Функция с разрывом производной в точке минимума	17
5.3.1	Выбор параметров и начальной точки	17
5.3.2	Результаты тестирования	18
5.3.3	Анализ результатов	19
5.4	Функция с большим количеством локальных минимумов	19
5.4.1	Выбор параметров и начальной точки	20
5.4.2	Результаты тестирования	20
5.4.3	Анализ результатов	22
5.5	Функция со случайной помехой	22
5.5.1	Выбор параметров и начальной точки	22
5.5.2	Результаты тестирования	22
5.5.3	Анализ результатов	24
5.6	Почти аналитическая функция с плохой сходимостью метода отжига	24
5.6.1	Выбор параметров и начальной точки	25
5.6.2	Результаты тестирования	25
5.6.3	Анализ результатов	27

6	Нахождение корней функций с помощью метода отжига	27
6.1	Многочлены от одной переменной	27
6.1.1	Результаты тестирования	27
6.1.2	Анализ результатов	32
7	О выборе параметров метода отжига	32
7.1	Пример	33
8	Заключение	34
9	Список литературы	34

1 Введение

В настоящей дипломной работе рассмотрены реализации и свойства различных вариантов метода отжига.

Метод отжига (синонимы: метод обжига, метод симуляции отжига, метод модельной закалки, *simulated annealing*) — это техника оптимизации, использующая упорядоченный случайный поиск на основе аналогии с процессом образования веществом кристаллической структуры с минимальной энергией при охлаждении.

В настоящее время метод отжига применяется для решения многих оптимизационных задач — финансовых (см. [13], [16]), компьютерной графики (см. [14–15]), комбинаторных (см. [17]), в телекоммуникационных сетях (см. [18]), и многих других. Зачастую метод отжига используют для обучения нейронных сетей (см. [10–11]).

Несмотря на такую широкую область применения, скорость сходимости метода отжига всё ещё мало изучена.

История метода отжига начинается с 1953 года. В этом году Н. Метрополисом был разработан алгоритм симуляции установления равновесия в системе с множеством степеней свободы при заданной температуре [1]. В начале 80-х у С. Киркпатрика впервые появилась идея использовать этот алгоритм не только для моделирования физических систем, но и для решения некоторых задач оптимизации [2].

Огромным преимуществом метода отжига является свойство избежать “ловушки” в локальных минимумах оптимизируемой функции, и продолжить поиск глобального минимума. Это достигается за счет принятия не только изменений параметров, приводящих к уменьшению значения функции, но и некоторых изменений, увеличивающих ее значение, в зависимости от т.н. *температуры* T — характеристики моделируемого процесса. Чем выше температура, тем большие “ухудшающие” изменения (аналогичные случайным флуктуациям в нагретом веществе) допустимы, и больше их вероятность.

Еще одним преимуществом является то, что даже в условиях нехватки вычислительных ресурсов для нахождения глобального минимума, метод отжига, как правило, выдает весьма неплохое решение (один из локальных минимумов).

В [3] Л. Ингбером показано, что метод отжига и его модификации являются одним из наиболее эффективных методов случайного поиска оптимального решения для большого класса задач. В [4] Л. Ингбером проведен сравнительный анализ адаптивного метода отжига (*Adaptive Simulated Annealing* — *ASA*) и генетических алгоритмов, из которого следует, что на большинстве задач метод отжига не проигрывает генетическим алгоритмам, а на многих и выигрывает.

К настоящему времени разработано множество различных вариантов метода отжига, как общих, так и их специализаций для решения конкретных задач.

В данной работе исследованы известные на настоящий момент общие схемы отжига, а также изучено их поведение на различных задачах оптимизации. Предложен способ решения задачи нахождения какого-либо корня функции с помощью метода отжига. Также предложен новый подход к выбору параметров метода отжига.

2 Алгоритм метода отжига

Метод отжига служит для поиска глобального минимума некоторой функции $f(x)$, заданной для x из некоторого пространства S , дискретного или непрерывного. Элементы множества S представляют собой состояния воображаемой физической системы (“энергетические уровни”), а значение функции f в этих точках используется как

энергия системы $E = f(x)$. В каждый момент предполагается заданной *температура* системы T , как правило, уменьшающаяся с течением времени. Находясь в состоянии x при температуре T , следующее состояние системы выбирается в соответствии с заданным *порождающим семейством вероятностных распределений* $\mathcal{G}(x, T)$, которое при фиксированных x и T задаёт случайный элемент $G(x, T)$ со значениями в пространстве S . После генерации нового состояния $x' = G(x, T)$, система с вероятностью $h(\Delta E, T)$ переходит к следующему шагу в состояние x' , в противном случае процесс генерации x' повторяется. Здесь ΔE обозначает приращение функции энергии $f(x') - f(x)$. Величина $h(\Delta E, T)$ называется *вероятностью принятия* нового состояния.

Как правило, в качестве функции $h(\Delta E, T)$ выбирается либо точное значение соответствующей физической величины

$$h(\Delta E, T) = \frac{1}{1 + \exp(\Delta E/T)},$$

либо приближенное значение

$$h(\Delta E, T) = \exp(-\Delta E/T).$$

Вторая формула используется наиболее часто. При её использовании $h(\Delta E, T)$ оказывается больше единицы в случае $\Delta E < 0$, и тогда соответствующая вероятность считается равной 1. Таким образом, если новое состояние даёт лучшее значение оптимизируемой функции, то переход в это состояние произойдет в любом случае.

Итак, конкретная схема метода отжига задается следующими параметрами:

- Выбором закона изменения температуры $T(k)$, где k — номер шага.
- Выбором порождающего семейства распределений $\mathcal{G}(x, T)$.
- Выбором функции вероятности принятия $h(\Delta E, T)$.

Алгоритм:

1. Случайным образом выбирается начальная точка $x = x_0, x_0 \in S$. Текущее значение энергии E устанавливается в значение $f(x_0)$.
2. k -я итерация основного цикла состоит из следующих шагов:
 - (а) Сравнить энергию системы E в состоянии x с найденным на текущий момент глобальным минимумом. Если $E = f(x)$ меньше, то изменить значение глобального минимума.
 - (б) Сгенерировать новую точку $x' = G(x, T(k))$.
 - (с) Вычислить значение функции в ней $E' = f(x')$.
 - (д) Сгенерировать случайное число α из интервала $[0; 1]$
 - (е) Если $\alpha < h(E' - E, T(k))$, то установить $x \leftarrow x', E \leftarrow E'$ и перейти к следующей итерации. Иначе повторить шаг (б), пока не будет найдена подходящая точка x' .

Известны следующие модификации этого алгоритма:

- Модификация А. На шаге 2.(е) переход к следующей итерации происходит и в том случае, если точка x' не являлась подходящей. При этом следующая итерация начинается с точки x , но уже с новым значением температуры.

- Модификация Б. В качестве оценки точки минимума возвращается последнее значение x . Это может незначительно ускорить алгоритм в случае большой размерности S , но с небольшой вероятностью может привести к тому, что будет получено худшее решение (особенно, если температура к моменту завершения алгоритма остаётся значительно больше нуля).
- Модификация В. На шаге 2.(b) x' вычисляется рекуррентно с использованием формулы $x' = G(x', T(k))$. Изначально на шаге 1 устанавливается $x' \leftarrow x_0$. Это позволяет избежать “застревания” алгоритма, однако такая реализация теряет множество преимуществ метода отжига, т.к. не очень сильно отличается от обычного случайного поиска (особенно, если это комбинируется с вариантом 1 — в этом случае проверку $h(\Delta E, T)$ вообще можно исключить).

В следующем разделе будут рассмотрены основные схемы выбора параметров метода отжига.

3 Общие схемы метода отжига

3.1 Больцмановский отжиг

Исторически первой схемой метода отжига является т.н. схема *Больцмановского отжига*. Именно эта схема использовалась Н. Метрополисом и др. в [1] для вычисления многомерных интегралов пути в задачах статистической физики, а также С. Киркпатриком и др. в [2] для решения задачи нахождения оптимальной разводки микросхем. В Больцмановском отжиге изменение температуры задаётся формулой

$$T(k) = \frac{T_0}{\ln(1 + k)}, k > 0.$$

Семейство распределений $\mathcal{G}(x, T)$ выбирается как семейство нормальных распределений с математическим ожиданием x и дисперсией T , т.е. задаётся плотностью

$$g(x'; x, T) = (2\pi T)^{-D/2} \exp(-|x' - x|^2/(2T)),$$

где D — размерность пространства состояний. Пространство состояний предполагается метрическим.

Для Больцмановской схемы доказано (см. [5]), что при достаточно больших T_0 и общем количестве шагов K , выбор такого семейства распределений гарантирует нахождение глобального минимума.

3.2 Отжиг Коши (быстрый отжиг)

Основным недостатком Больцмановского отжига является очень медленное убывание температуры. Например, для того, чтобы понизить исходную температуру в 40 раз, требуется $e^{40} \approx 2.35 \cdot 10^{17}$ итераций, что уже вряд ли приемлемо при решении каких-либо задач. Ввиду этого Цу и Хартли в [6] предложили алгоритм, который позволяет использовать для изменения температуры схему

$$T(k) = \frac{T_0}{k} \tag{1}$$

без потери гарантии нахождения глобального минимума. Это достигается за счет использования в качестве \mathcal{G} распределений Коши с плотностью

$$g(x'; x, T) = \frac{T}{(|x' - x|^2 + T^2)^{(D+1)/2}},$$

соответствующим образом нормированных. Например, в случае $D = 1$ приходим к плотности

$$g(x'; x, T) = \frac{1}{\pi} \frac{T}{|x' - x|^2 + T^2}.$$

К сожалению, это распределение не очень удобно моделировать в пространстве размерности больше 1. Этого можно избежать, например, с помощью перемножения D одномерных распределений Коши:

$$g(x'; x, T) = \frac{1}{\pi^D} \prod_{i=1}^D \frac{T}{|x'_i - x_i|^2 + T^2},$$

но в этом случае нахождение глобального минимума гарантируется только при законе изменения температуры не быстрее чем:

$$T(k) = \frac{T_0}{k^{1/D}},$$

что гораздо медленнее схемы (1).

3.3 Сверхбыстрый отжиг

Недостатки двух предыдущих методов привели к тому, что в 1989 году американским исследователем Л. Ингбером был разработан метод сверхбыстрого отжига (Very Fast Annealing, [7]). В нём пространство S считается состоящим из D -мерных векторов $\begin{pmatrix} x_1 \\ \vdots \\ x_D \end{pmatrix}$, где $x_i \in [A_i, B_i]$. Кроме этого, температура по каждой из координат может различаться, таким образом, T также является вектором размерности D .

Семейство распределений строится следующим образом. Вводится функция

$$g_T(y) = \prod_{i=1}^D \frac{1}{2(|y_i| + T_i) \ln(1 + 1/T_i)} \equiv \prod_{i=1}^D g_{(i;T)}(y_i), \quad y_i \in [-1, 1]$$

В качестве y для получения плотности распределений \mathcal{G} используется $\frac{\Delta x}{B_i - A_i}$, таким образом, новое значение x'_i вычисляется по формуле $x'_i = x_i + z_i(B_i - A_i)$, где z_i — случайная величина с плотностью $g_{(i;T)}$ на $[-1, 1]$. При этом выходящие за границы интервала значения параметра генерируются заново (пока не произойдет попадание в интервал) или приравниваются соответствующим границам.

Такую случайную величину легко промоделировать:

$$z_i = \text{sgn}(\alpha_i - \frac{1}{2}) T_i ((1 + 1/T_i)^{|2\alpha_i - 1|} - 1), \quad (2)$$

где α_i — независимые случайные величины, распределенные равномерно на $[0, 1]$.

Доказано (см. [7]), что закон изменения температуры

$$T_i(k) = T_{(i;0)} \exp(-c_i k^{1/D}), \quad c_i > 0$$

даёт статистическую гарантию нахождения глобального минимума. Для вероятности принятия также используется отдельная шкала температуры, изменяющаяся по такому же закону. Как правило, при реализации этого метода c_i управляется двумя параметрами:

$$c_i = m_i \exp(-n_i/D).$$

Преимущества такого метода очевидны. Во-первых, экспоненциальное убывание температуры гораздо быстрее достижимого в предыдущих методах. Во-вторых, разделение размерностей может дать большой выигрыш, как и благодаря отдельным температурам (т.к. специфика конкретной задачи может сильно различать параметры), так и благодаря ускорению процесса, в случае, если не нужно менять все координаты одновременно.

Кроме того, в отличие от отжига Коши, сверхбыстрый отжиг, как было показано, допускает очень быстрое моделирование распределения \mathcal{G} независимо от размерности S .

Среди недостатков этого метода можно назвать то, что ввиду большого количества параметров иногда требуется по несколько месяцев, чтобы хорошо настроить его для решения конкретной задачи.

3.4 Алгоритм Ксин Яо

Алгоритм Ксин Яо [8] был получен повторным применением идеи предыдущего алгоритма. В качестве $g_T(y)$ выбирается

$$g_T(y) = \prod_{i=1}^D g_{(i;T)}(y_i) = \prod_{i=1}^D \frac{1}{2(|y_i| + \frac{1}{\ln(1/T_i)}) \ln(1 + \ln(1/T_i))}.$$

Утверждается, что при изменении температуры по закону

$$T_i(k) = T_{(i;0)} \exp(-\exp(b_i k^{1/D})), \quad b_i > 0$$

достигается статистическая гарантия нахождения глобального минимума.

Однако, как показано в [8], увеличение скорости убывания температуры вовсе не означает ускорения в решении задачи. Более того, “размазанность” распределения приводит к тому, что метод генерирует огромное число “длинных” переходов, которые отвергаются в силу низкой вероятности их принятия.

Таким образом, несмотря на то, что этот процесс можно итерировать до бесконечности, получая законы изменения температуры вроде

$$T_i(k) = T_{(i;0)} \exp(-\exp(\exp(\dots \exp(b_i k^{1/D}) \dots))),$$

ценность таких “улучшений” представляется сомнительной. Более того, легко видеть, что в пределе это приводит к тривиальному методу случайного поиска, которым является метод отжига при $T = 0$.

Это в небольшой степени применимо и к методу сверхбыстрого отжига, так что вопрос о скорости сходимости этих методов, а также о других методах, обеспечивающих не такое быстрое убывание температуры, но бóльшую скорость сходимости, остаётся открытым.

Следовательно, работу [8] можно рассматривать и как некую скрытую критику метода сверхбыстрого отжига. Однако, вполне возможны задачи, на которых вторая итерация вышеописанного процесса может давать неплохие результаты, поэтому этот метод и рассмотрен в настоящей работе.

3.5 Методы “тушения”

Далеко не всегда хватает вычислительных ресурсов на поиск глобального минимума. Кроме того, зачастую достаточно не глобально оптимального решения задачи, а достаточно близкого к нему. Методы “тушения” (simulated quenching) не гарантируют нахождения глобального минимума, но, как правило, быстро находят близкое решение, а на практике зачастую и сам оптимум.

Основная идея этих методов заключается в том, чтобы скомбинировать семейство распределений \mathcal{G} одного из предыдущих четырёх методов с более быстрым законом убывания температуры.

Например, можно рассматривать нормальные распределения \mathcal{G} из Больцмановского отжига, но при этом уменьшать температуру по закону

$$T_{k+1} = cT_k.$$

Как правило, в этом случае c выбирается между 0.7 и 0.99. Такой метод очень быстро сходится, и для конкретных задач может давать весьма неплохое решение, близкое к оптимальному, в условиях реального времени.

Некоторые из методов “тушения” рассмотрены в [3] и [9]. Зачастую они основаны либо на нормальном распределении, либо на распределении для сверхбыстрого отжига. Кроме того, встречаются специальные распределения, подобранные опытным путём для решения конкретных задач.

3.6 Масштабирование в ходе отжига

Зачастую при реализации сверхбыстрого отжига в задачах с большой размерностью используется технология *масштабирования отжига* (reannealing), иногда также применяемая и к другим вариантам отжига. При использовании этой технологии периодически во время отжига производится следующая операция. Обозначим за s_i значение некоторой оценки производной целевой функции по i -й координате в точке текущего минимума:

$$s_i \geq \left| \frac{\partial f}{\partial x_i}(x_i^{min}) \right|.$$

Кроме того, пусть $s_{max} = \max_{1 \leq i \leq D} s_i$. После этого номер шага (называемый в этом алгоритме *временем отжига* (annealing-time)) и температура для каждой размерности изменяются следующим образом:

$$T_i' = T_i(k_i)(s_{max}/s_i)$$

$$k_i' = (\ln(T_{(i;0)}/T_i')/c_i)^D.$$

Таким образом, значение $T_{(i;0)}$ сохраняется, а время отжига (которое теперь может принимать не только целые значения) масштабируется согласно температуре.

Температура, используемая для расчета вероятности принятия, также масштабируется. Способ масштабирования зависит от задачи, но, как правило, коэффициент масштабирования определяется значением целевой функции в точке текущего минимума (чем меньше значение функции, тем меньше температура). Время отжига для пересчета температуры принятия также изменяется аналогичным образом.

Кроме приведенного, возможны другие способы масштабирования, зависящие от конкретной задачи. Как правило, они сводятся к другому выбору *коэффициентов чувствительности* s_i . Приведенный выше способ выбора s_i используется для нелинейных физических задач большой размерности.

Алгоритм сверхбыстрого отжига с использованием этой технологии описан в [3], [7] и [9], и носит название *адаптивного метода отжига* (adaptive simulated annealing).

4 Моделирование схем отжига

Для сравнения различных схем метода отжига одной из задач дипломной работы являлось их моделирование. Вопрос моделирования различных схем отжига разбивается на две части:

- Промоделировать соответствующее распределение.
- Промоделировать схему изменения температуры.

Исходные коды модуля, реализующего каждую из схем, можно найти в Приложении.

4.1 Больцмановский отжиг

Нормальное распределение моделировалось с помощью центральной предельной теоремы. Обозначим α_i независимые реализации равномерного распределения на $[0; 1]$. Тогда $\mathbf{E}\alpha_i = 0$, $\mathbf{D}\alpha_i = 1/12$ и

$$\frac{\sum_{i=1}^n \alpha_i}{\sqrt{n/12}} \xrightarrow{n \rightarrow \infty} \mathbf{N}(0; 1).$$

Таким образом, можно сложить n независимых реализаций α_i , и разделив их сумму на $\sqrt{n/12}$, мы будем иметь достаточно хорошее приближение к нормальному распределению. Как показывает практика, достаточно взять $n = 24$. Для моделирования требуемого $\mathbf{N}(0; T)$ достаточно домножить получившееся число на \sqrt{T} . В случае размерности, большей единицы, по каждой из координат вводились независимые возмущения с таким распределением.

Для моделирования убывания температуры использовалась непосредственно формула $T_0 / \ln k$, $k = 2, 3, \dots$

4.2 Отжиг Коши

Распределение Коши моделировалось методом обратных функций. Пусть \mathbf{X} — случайная величина, имеющая распределение Коши с плотностью

$$g(x) = \frac{1}{\pi} \frac{T}{x^2 + T^2}.$$

Вычислим $\mathbf{P}\{\mathbf{X} < y\}$:

$$\mathbf{P}\{\mathbf{X} < y\} = \int_{-\infty}^y \frac{1}{\pi} \frac{T}{x^2 + T^2} dx = \frac{1}{\pi} \left(\operatorname{arctg} \left(\frac{y}{T} \right) + \frac{\pi}{2} \right).$$

Следовательно, если α — реализация равномерно распределенной на $[0; 1]$ случайной величины, то величина

$$T \operatorname{tg} \left(\pi \alpha - \frac{\pi}{2} \right) \quad (3)$$

имеет требуемое распределение.

В случае задач с размерностью $D > 1$ использовалось произведение D одномерных распределений Коши, т.е. вносилось возмущение по каждой из координат, вычисленное по формуле (3) для независимых α_i .

Для температуры в этом случае используется формула

$$T(k) = \frac{T_0}{k^{1/D}},$$

причем в случае $D = 1$ она вычислялась как T_0/k , в случае $D = 2$ — как T_0/\sqrt{k} для увеличения скорости работы программ.

4.3 Сверхбыстрый отжиг

Для сверхбыстрого отжига использовалась также полученная методом обратных функций формула (2). Закон изменения температуры также имел отличия при $D = 1$ и $D = 2$. Если при $D = 2$ отличие, аналогично предыдущему случаю, сводилось к вычислению $k^{1/2}$ как \sqrt{k} , то при $D = 1$ возможна более существенная оптимизация, т.к.

$$\frac{T(k+1)}{T(k)} = \frac{e^{-c_i/(k+1)}}{e^{-c_i/k}} = e^{-c_i} = \text{const.}$$

Значение этой константы вычислялось заранее, и при каждом уменьшении температуры производилось умножение на нее вместо вычисления экспоненты и деления, что примерно в 200 раз быстрее.

Для предотвращения деления на 0 в формулах температура была ограничена снизу значением 10^{-4000} .

4.4 Методы “тушения”

Кроме вышеописанных методов, были смоделированы четыре метода тушения:

- Больцмановский отжиг с уменьшением температуры как в методе Коши.
- Больцмановский отжиг с экспоненциальным уменьшением температуры ($T_{k+1} = cT_k$, $c = 0.99$).
- Метод Коши с экспоненциальным уменьшением температуры ($T_{k+1} = cT_k$, $c = 0.99$).
- Сверхбыстрое тушение.

Сверхбыстрое тушение получается из метода сверхбыстрого отжига с использованием следующих формул:

$$T_i(k) = T_{(i;0)} \exp(-c_i k^{Q/D}), \quad c_i > 0, \\ c_i = m_i \exp(-n_i/D),$$

где Q — так называемый *множитель тушения*. При тестировании множитель тушения брался равным 2.

5 Сравнение схем отжига на задачах нахождения минимума

В этом разделе приводятся результаты сравнения реализованных мной схем отжига на различных задачах оптимизации. Каждая схема запускалась в двух вариантах — классическом, а также в варианте с уменьшением температуры и после неудач (вариант 2.А). Для получившихся 14 схем приводятся графики невязки, полученной в результате усреднения по 100 запускам метода.

5.1 Обозначения к графикам

Для каждого теста приводится по два графика — на первом сравниваются методы в классическом варианте, на втором — в варианте 2.А. На графике отображено изменение невязки по мере вычисления значений функции в новых точках.

Ниже приводится таблица обозначений для графиков, приведенных для тестов:

- Больцмановский отжиг
- - - - - Отжиг Коши
- ⋯ Сверхбыстрый отжиг
- - - - - Больцмановское тушение (убывание T по Коши)
- · - · - - - - Больцмановское тушение (экспоненциальное убывание T)
- · - - - - Тушение по Коши-1
- ⋯ Сверхбыстрое тушение

5.2 Многочлены от одной переменной

5.2.1 Генерация случайных многочленов

Для генерации псевдослучайного многочлена степени N одной переменной использовалась следующая технология. Генерируется число c в пределах от 0 до $\lfloor \frac{N-1}{2} \rfloor$ — количество комплексных корней производной. Далее генерируются c комплексно-сопряжённых пар вида $x_k \pm iy_k$, $-1 \leq x_k \leq 1$, $-1 \leq y_k \leq 1$ и $N-1-2c$ вещественных чисел $-1 \leq r_j \leq 1$. Производная многочлена имеет вид

$$P'(t) = \prod_{k=1}^c (t^2 - 2x_k t + x_k^2 + y_k^2) \cdot \prod_{j=1}^{N-2c-1} (t - r_j) \quad (4)$$

Многочлен получается тривиальным интегрированием, и достигает своего минимума в одной из точек r_j . Свободный член после интегрирования генерируется случайным образом.

Исходные коды процедуры генерации можно найти в приложении.

5.2.2 Выбор тестов

Тестирование производилось на случайных многочленах степеней 2, 10 и 100. Как правило, работа методов на различных многочленах одной степени не очень сильно отличается, поэтому для каждой степени приводятся результаты только для одного специально подобранного многочлена, иллюстрирующего особенности работы методов на многочленах.

5.2.3 Выбор начальной точки

В каждом из методов начальная точка выбиралась как случайное число от -5 до 5 , несмотря на то, что из процедуры генерации очевидно, что минимум лежит между -1 и 1 . Это позволяет протестировать каждый метод на работу в условиях сильного промаха при генерации случайной точки.

5.2.4 Выбор параметров сверхбыстрого отжига

Как показали испытания, сверхбыстрый отжиг достаточно эффективно работает на многочленах при $M = 1$, $N = 5$. При существенном изменении этих параметров метод иногда не находит глобальный минимум за время тестирования.

5.2.5 Выбор стартовой температуры

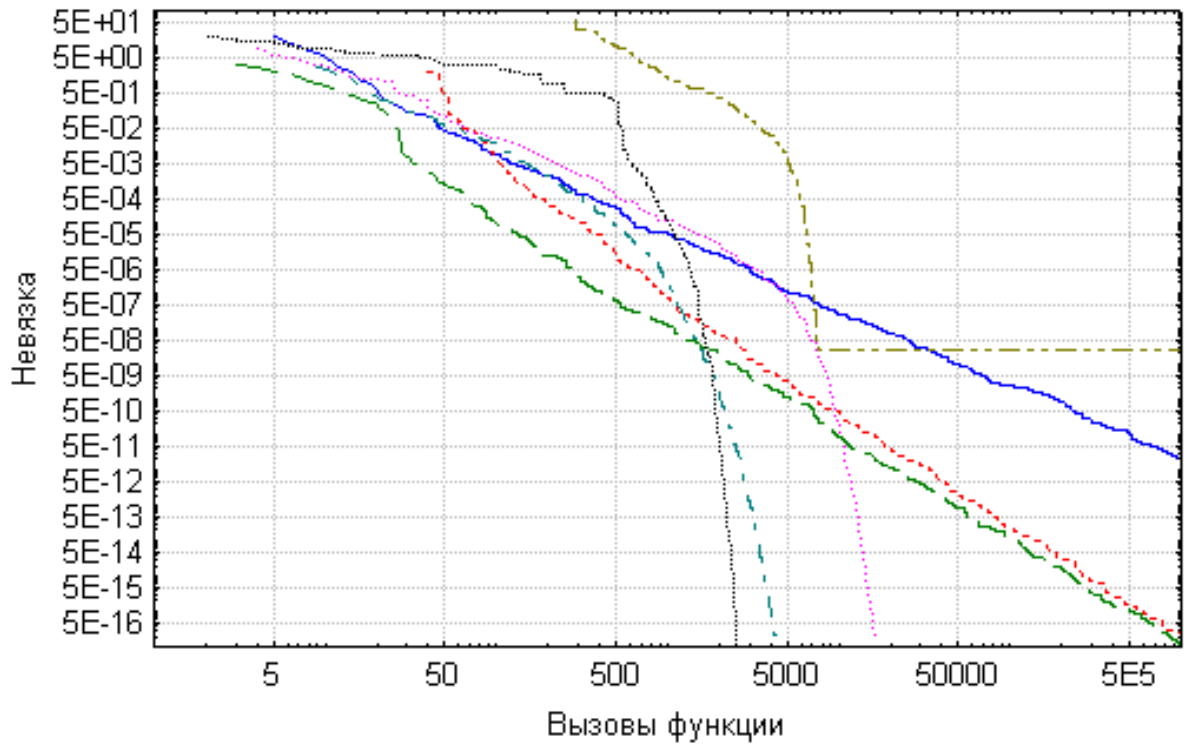
Температура для каждого из методов подбиралась опытным путём. Найденное значение использовалось для всех многочленов, независимо от степени. Используемые значения приведены в таблице:

Метод	T_0
Больцмановский отжиг без уменьшения T после неудач	80
Больцмановский отжиг с уменьшением T после неудач	80
Отжиг Коши без уменьшения T после неудач	60
Отжиг Коши с уменьшением T после неудач	60
Сверхбыстрый отжиг без уменьшения T после неудач	239
Сверхбыстрый отжиг с уменьшением T после неудач	239
Больцмановское тушение (по Коши) без уменьшения T после неудач	10
Больцмановское тушение (по Коши) с уменьшением T после неудач	5
Больцмановское тушение ($T_{k+1} = cT_k$) без уменьшения T после неудач	50
Больцмановское тушение ($T_{k+1} = cT_k$) с уменьшением T после неудач	500
Тушение Коши ($T_{k+1} = cT_k$) без уменьшения T после неудач	1000
Тушение Коши ($T_{k+1} = cT_k$) с уменьшением T после неудач	60
Сверхбыстрое тушение без уменьшения T после неудач	10^7
Сверхбыстрое тушение с уменьшением T после неудач	600

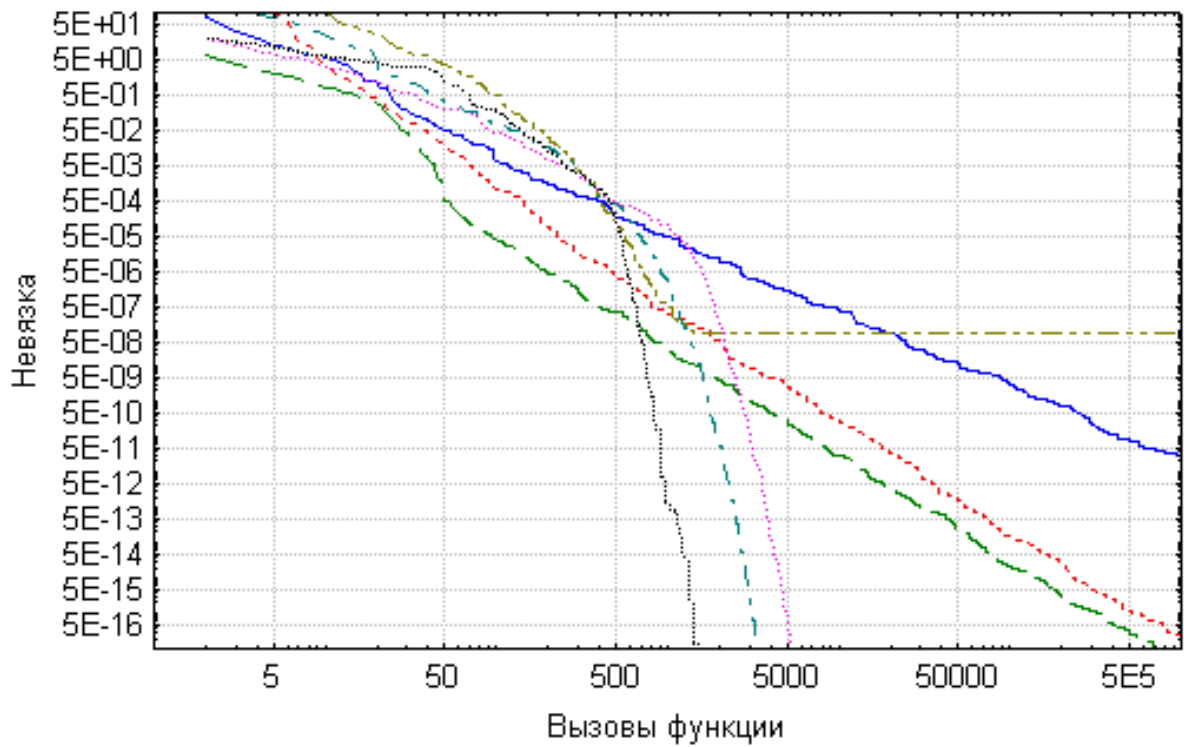
5.2.6 Результаты тестирования

Для многочлена степени 2, 10 и 100 приводится по два графика — результаты работы классической реализации методов и результаты работы методов в варианте 2.А.

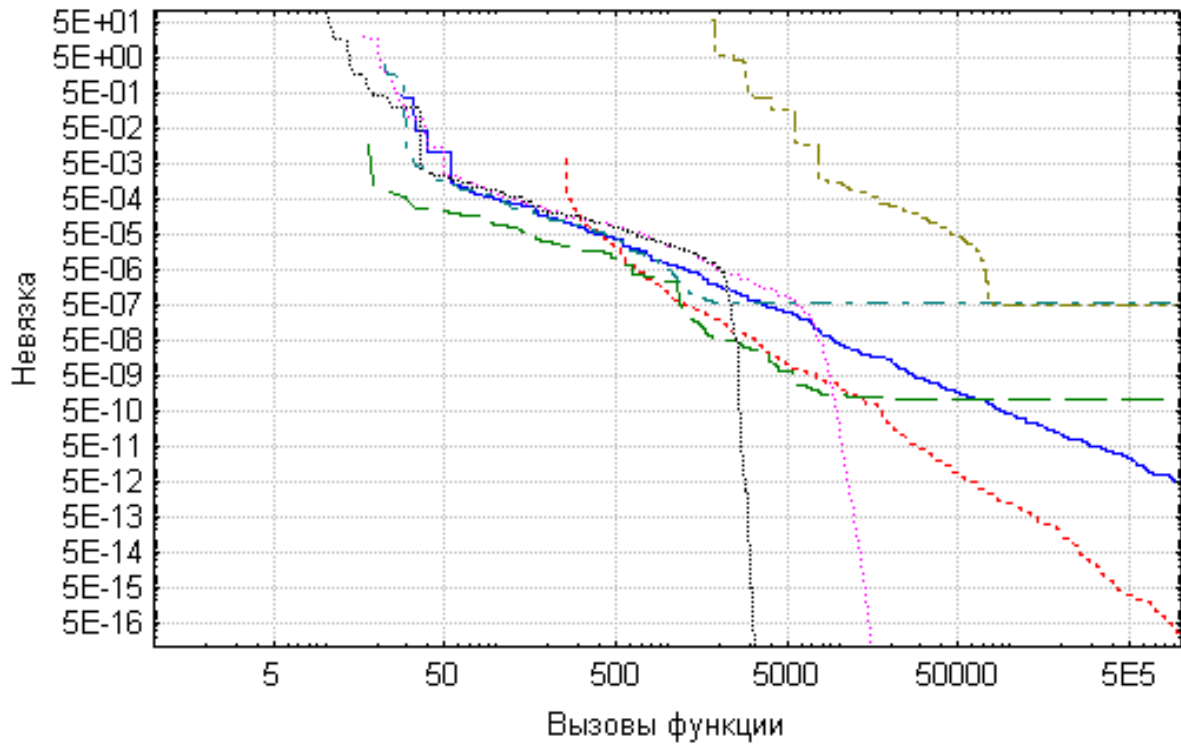
Случайный квадратный трёхчлен (без уменьшения T после неудач)



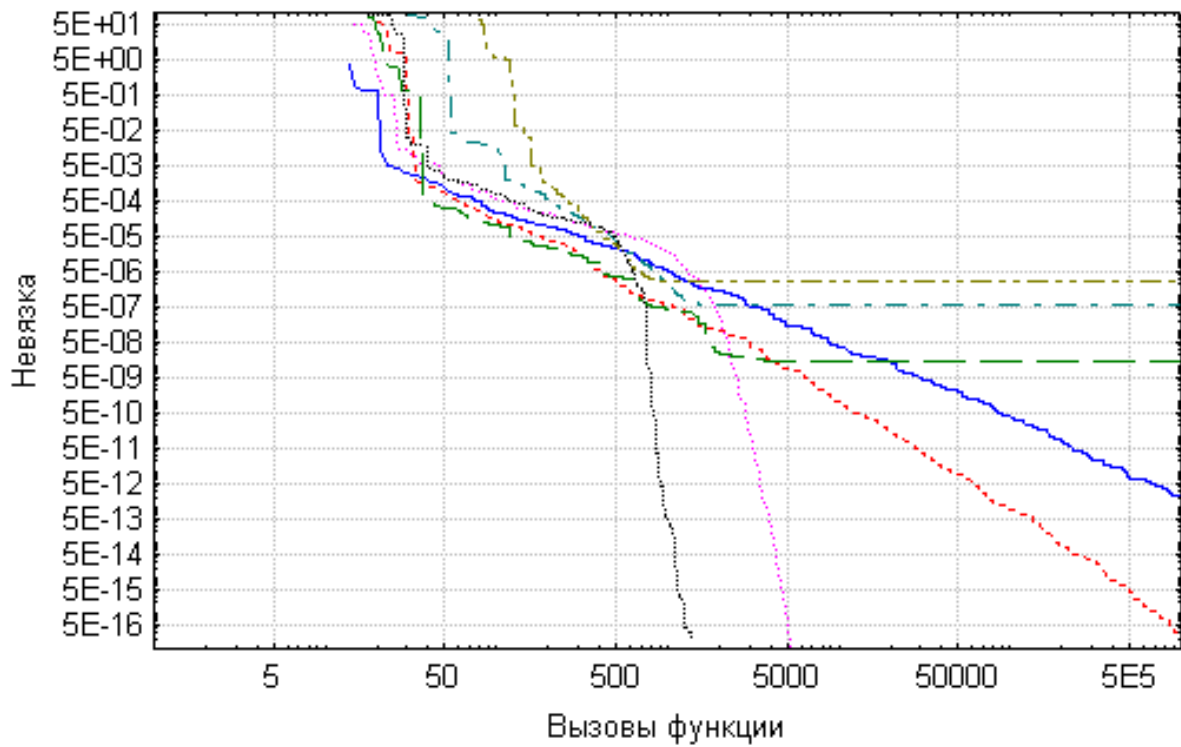
Случайный квадратный трёхчлен (с уменьшением T после неудач)



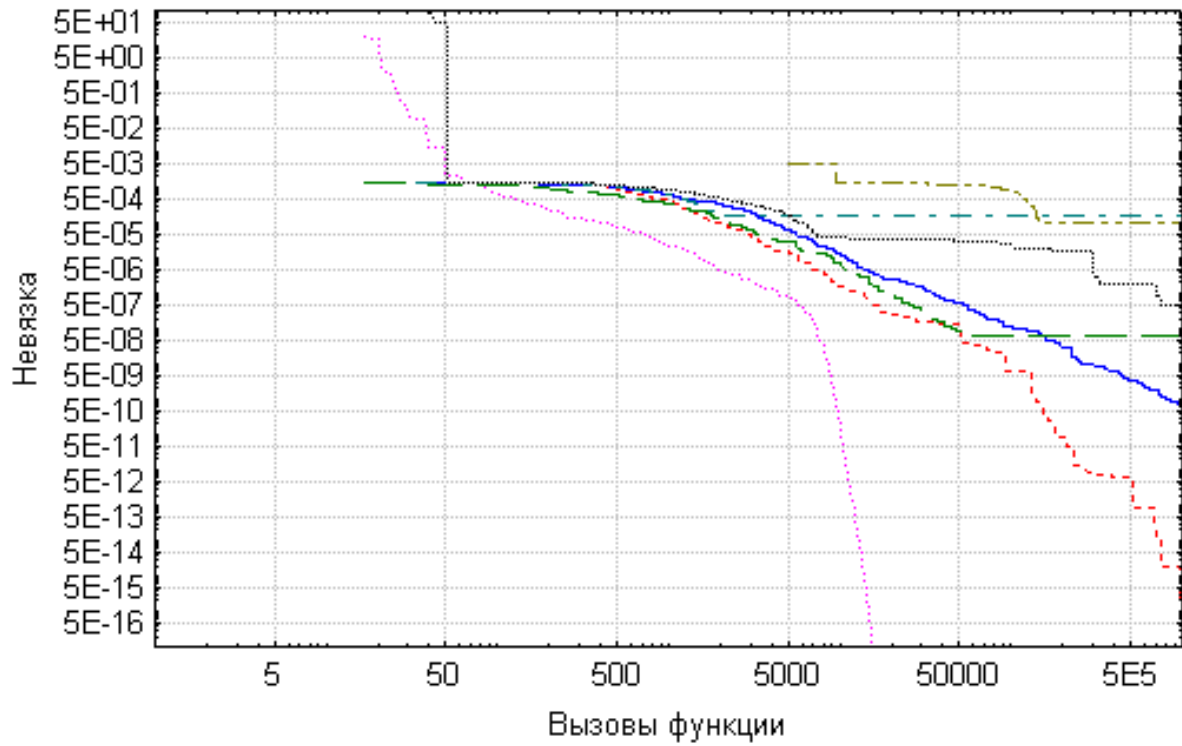
Случайный многочлен степени 10 (без уменьшения T после неудач)



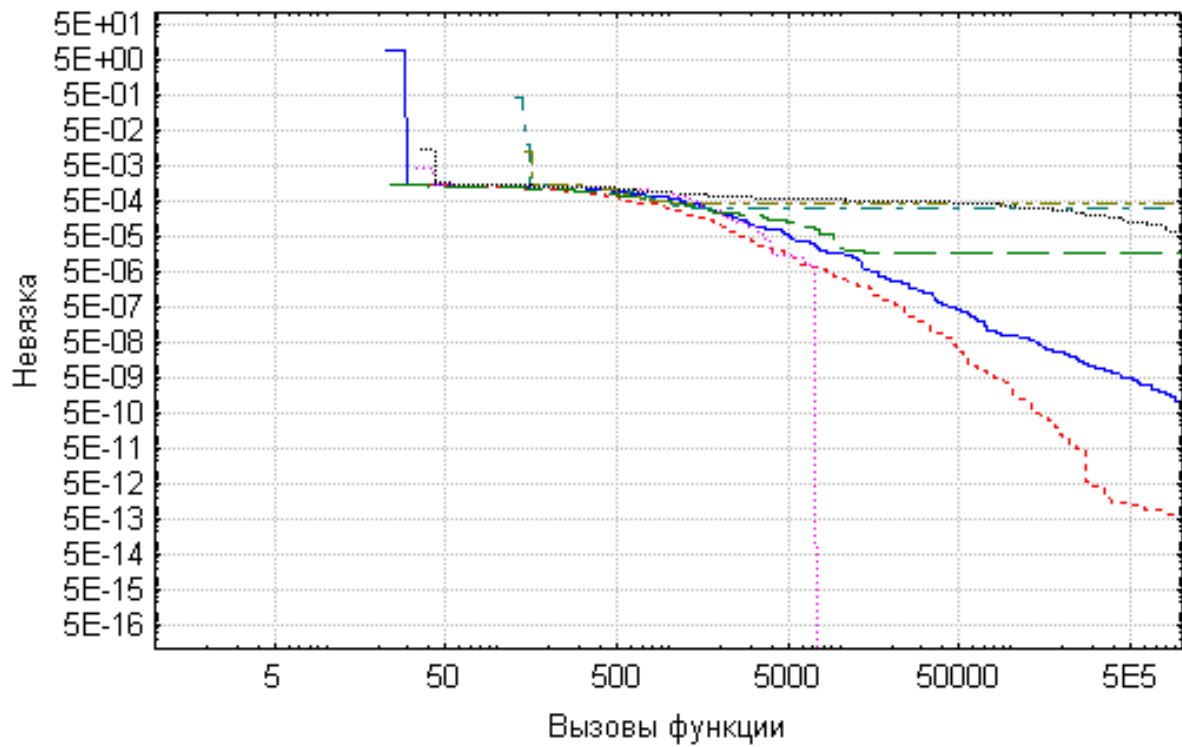
Случайный многочлен степени 10 (с уменьшением T после неудач)



Случайный многочлен степени 100 (без уменьшения T после неудач)



Случайный многочлен степени 100 (без уменьшения T после неудач)



5.2.7 Анализ результатов

Из полученных результатов очевидно, что наиболее эффективным из методов нахождения минимума для многочленов является метод сверхбыстрого отжига в варианте реализации с уменьшением температуры после неудач. На всех трёх многочленах этот метод нашёл точное значение минимума не более чем за 7300 итераций. Для многочлена сотой степени этот результат является весьма эффективным даже в сравнении с методами, использующими сам многочлен, а не только его значения в каких-то точках.

Методы тушения, чуть более эффективные, чем сверхбыстрый отжиг, на многочленах маленькой степени, с повышением степени иногда начинают попадать не в глобальный, а в один из локальных минимумов и там оставаться. Этим обусловлена горизонтальная прямая линия на графике. Такой эффект, видимо, наблюдается по причине того, что методы не имеют статистической гарантии глобального нахождения минимума.

Следует отметить, что метод сверхбыстрого тушения, в отличие от прочих методов тушения, может на итерациях с большим номером “выправлять” свое положение, отказываясь от локального минимума. Это хорошо видно по графику работы этого метода на многочлене степени 100.

Кроме этого, из графиков видно, что методы со статистической гарантией нахождения глобального минимума, как правило, продолжают сходиться и на больших номерах итераций, правда, гораздо медленнее, чем метод сверхбыстрого отжига.

Также можно сделать вывод, что использование тушения Коши с экспоненциальным убыванием температуры является крайне неэффективным для такого рода задач.

При сравнении методов Больцмана и Коши, реализованных по классическому алгоритму и по алгоритму с модификацией 2.A, видно, что разница между этими вариантами на многочленах незначительна. Для всех остальных методов модификация 2.A даёт тот же результат в несколько раз быстрее.

5.3 Функция с разрывом производной в точке минимума

В качестве функции с разрывом производной в точке минимума использовалась 10-мерная сумма модулей координат

$$\sum_{i=1}^D |x_i|.$$

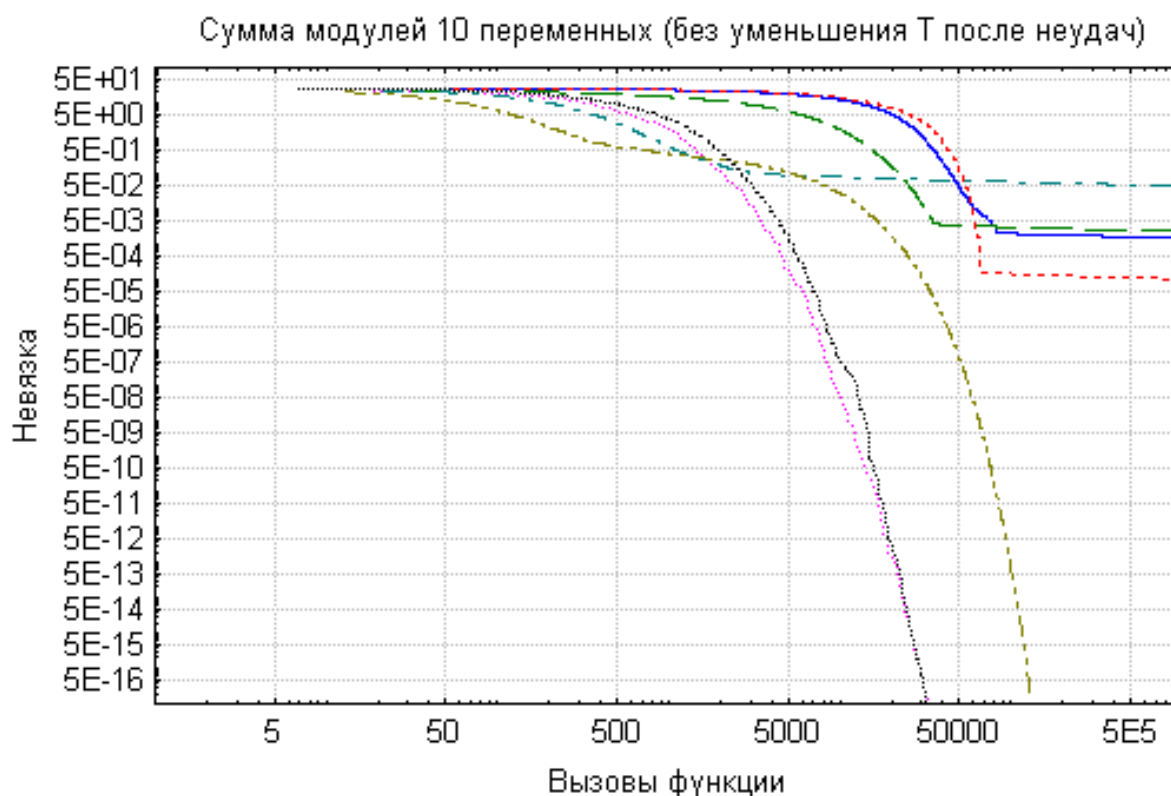
5.3.1 Выбор параметров и начальной точки

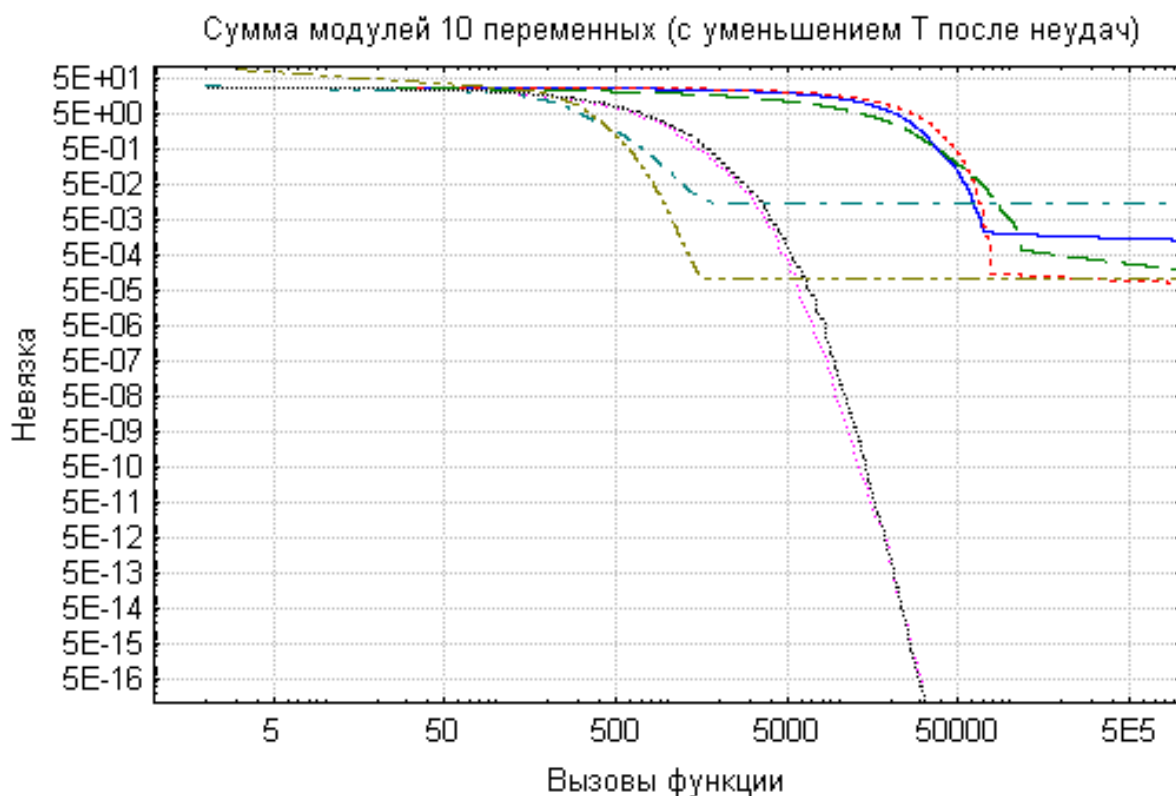
Каждая из координат инициализировалась реализацией случайной величины, распределенной равномерно на $[-5; 5]$. Параметры сверхбыстрого отжига были заданы как $M_i = 1$, $N_i = 5$, $1 \leq i \leq 10$. Температура по всем координатам использовалась одна и та же. Стартовая температура, как и в предыдущем тесте, подбиралась опытным путём; использованные значения приведены в таблице:

Метод	T_0
Больцмановский отжиг без уменьшения T после неудач	10^{-5}
Больцмановский отжиг с уменьшением T после неудач	10^{-5}
Отжиг Коши без уменьшения T после неудач	10^{-4}
Отжиг Коши с уменьшением T после неудач	10^{-4}
Сверхбыстрый отжиг без уменьшения T после неудач	10^{-100}
Сверхбыстрый отжиг с уменьшением T после неудач	10^{-100}
Больцмановское тушение (по Коши) без уменьшения T после неудач	0.01
Больцмановское тушение (по Коши) с уменьшением T после неудач	0.01
Больцмановское тушение ($T_{k+1} = cT_k$) без уменьшения T после неудач	0.01
Больцмановское тушение ($T_{k+1} = cT_k$) с уменьшением T после неудач	5
Тушение Коши ($T_{k+1} = cT_k$) без уменьшения T после неудач	0.1
Тушение Коши ($T_{k+1} = cT_k$) с уменьшением T после неудач	5
Сверхбыстрое тушение без уменьшения T после неудач	10^{-130}
Сверхбыстрое тушение с уменьшением T после неудач	10^{-120}

5.3.2 Результаты тестирования

Как и для предыдущей задачи, приведены два графика для различных вариантов основного алгоритма:





5.3.3 Анализ результатов

Из графиков видно, что большинство методов не очень эффективно работают на кусочно-линейных функциях. Единственным исключением является супербыстрый отжиг, а также тушение Коши без уменьшения T после неудач. Все остальные методы хоть и сходятся к глобальному минимуму, но крайне медленно — именно этим обусловлены почти прямые линии на графиках.

Стоит отметить высокую эффективность метода супербыстрого отжига на очень низких температурах (порядка 10^{-100}). Ни один из других методов на таких температурах практически не работает в силу того, что распределение оказывается очень близким к сосредоточенному в одной точке.

Разница в работе классического варианта и варианта 2.A на таких функциях практически отсутствует. Супербыстрое тушение по скорости сходимости не отличается от супербыстрого отжига.

5.4 Функция с большим количеством локальных минимумов

В качестве функции с большим количеством локальных минимумов была выбрана функция, использованная Л. Ингбером в [4] для сравнения супербыстрого отжига и генетических алгоритмов. Эта функция выглядит следующим образом:

$$f(x_1, \dots, x_4) = \sum_{i=1}^4 \begin{cases} (t_i \operatorname{sgn}(z_i) + z_i)^2 c d_i & \text{если } |x_i - z_i| < |t_i| \\ d_i x_i^2 & \text{иначе,} \end{cases}$$

$$z_i = \left\lfloor \left| \frac{x_i}{s_i} \right| + 0.49999 \right\rfloor \operatorname{sgn}(x_i) s_i,$$

$$s_i = 0.2, t_i = 0.05, \quad i = 1 \dots 4,$$

$$d_i = \{1, 1000, 10, 100\},$$

$$c = 0.15.$$

Здесь коэффициенты подобраны таким образом, что f определяет параболоид с небольшими “дырами”, которые увеличивают свою глубину по мере приближения к началу координат.

Минимизация такой функции проводилась при условиях $-1000 \leq x_i \leq 1000$. На таком множестве она имеет около 10^{16} локальных минимумов.

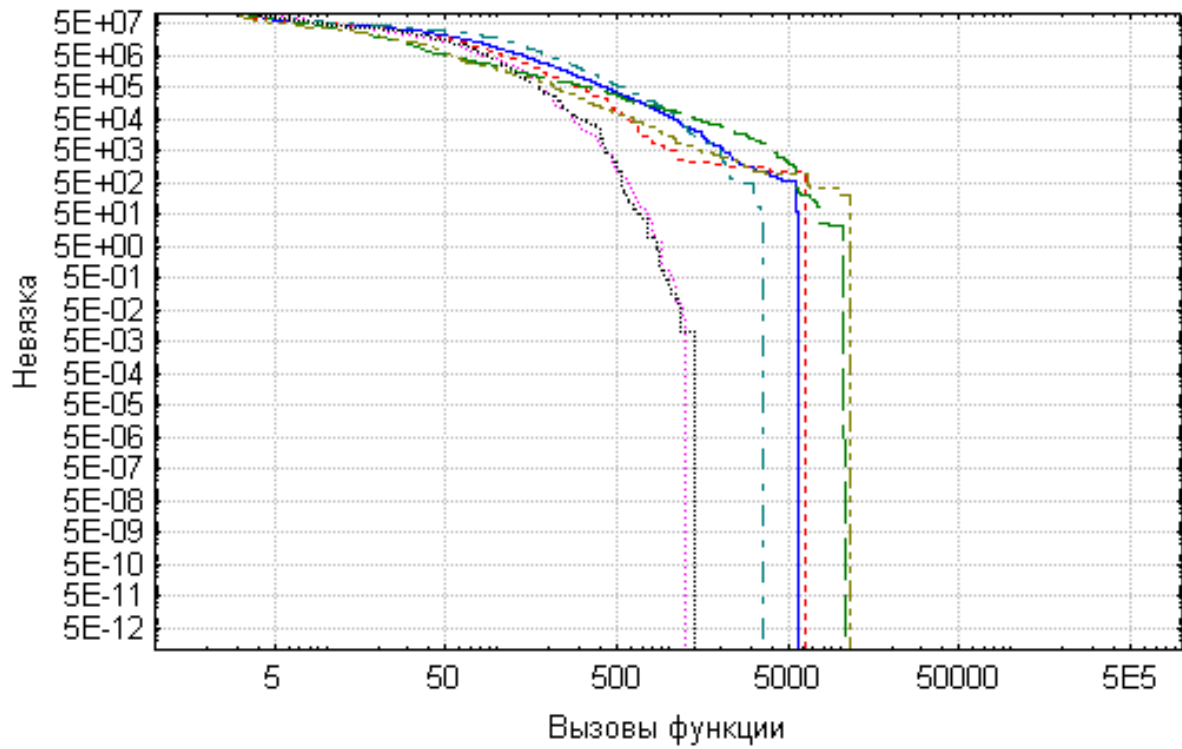
5.4.1 Выбор параметров и начальной точки

Начальная точка выбиралась равномерно распределенной на множестве $-1000 \leq x_i \leq 1000$. Для сверхбыстрого отжига использовались $m_i = 1$, $n_i = 3$, для сверхбыстрого тушения — $m_i = 1$, $n_i = 0$. Температура по всем координатам совпадала, стартовые значения подбирались опытным путем. Таблица стартовых температур для этой задачи приводится ниже:

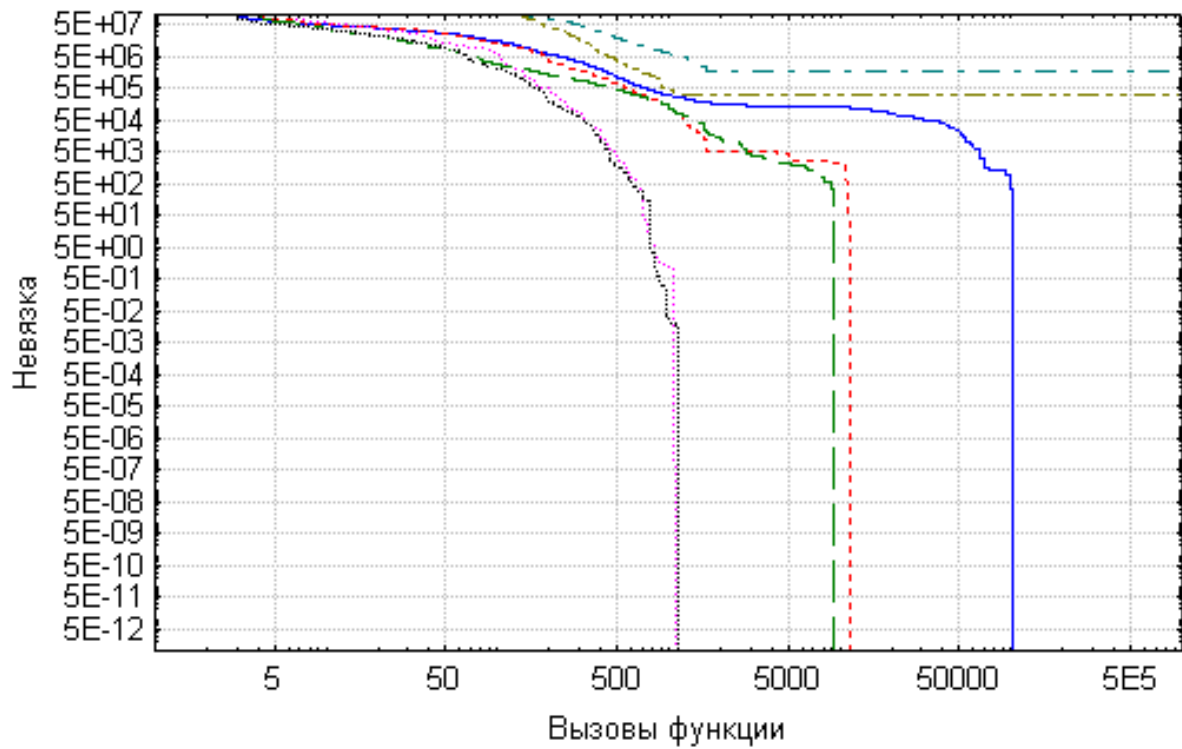
Метод	T_0
Больцмановский отжиг без уменьшения T после неудач	6 000
Больцмановский отжиг с уменьшением T после неудач	5 000
Отжиг Коши без уменьшения T после неудач	30
Отжиг Коши с уменьшением T после неудач	30
Сверхбыстрый отжиг без уменьшения T после неудач	10^{-8}
Сверхбыстрый отжиг с уменьшением T после неудач	10^{-8}
Больцмановское тушение (по Коши) без уменьшения T после неудач	200 000
Больцмановское тушение (по Коши) с уменьшением T после неудач	2 000 000
Больцмановское тушение ($T_{k+1} = cT_k$) без уменьшения T после неудач	2 000
Больцмановское тушение ($T_{k+1} = cT_k$) с уменьшением T после неудач	10^8
Тушение Коши ($T_{k+1} = cT_k$) без уменьшения T после неудач	50
Тушение Коши ($T_{k+1} = cT_k$) с уменьшением T после неудач	10 000
Сверхбыстрое тушение без уменьшения T после неудач	10^{-6}
Сверхбыстрое тушение с уменьшением T после неудач	10^{-4}

5.4.2 Результаты тестирования

Функция Ингбера (без уменьшения T после неудач)



Функция Ингбера (с уменьшением T после неудач)



5.4.3 Анализ результатов

Эта функция хорошо иллюстрирует эффективность метода отжига для поиска глобального минимума при наличии большого количества локальных. Практически все методы, кроме некоторых методов тушения с уменьшением T после неудач сошлись к глобальному минимуму.

На графике хорошо видно, что для методов Больцмана и Коши более эффективной является реализация без уменьшения температуры после неудач, а для сверхбыстрого отжига — реализация с уменьшением температуры.

5.5 Функция со случайной помехой

В качестве такой функции выбрана функция

$$f(x_1, \dots, x_{30}) = \sum_{i=1}^{30} x_i^4 + \alpha,$$

где α — случайная помеха, распределенная равномерно на $[0; 1]$.

Множество минимизации задавалось условиями $-1000 \leq x_i \leq 1000$.

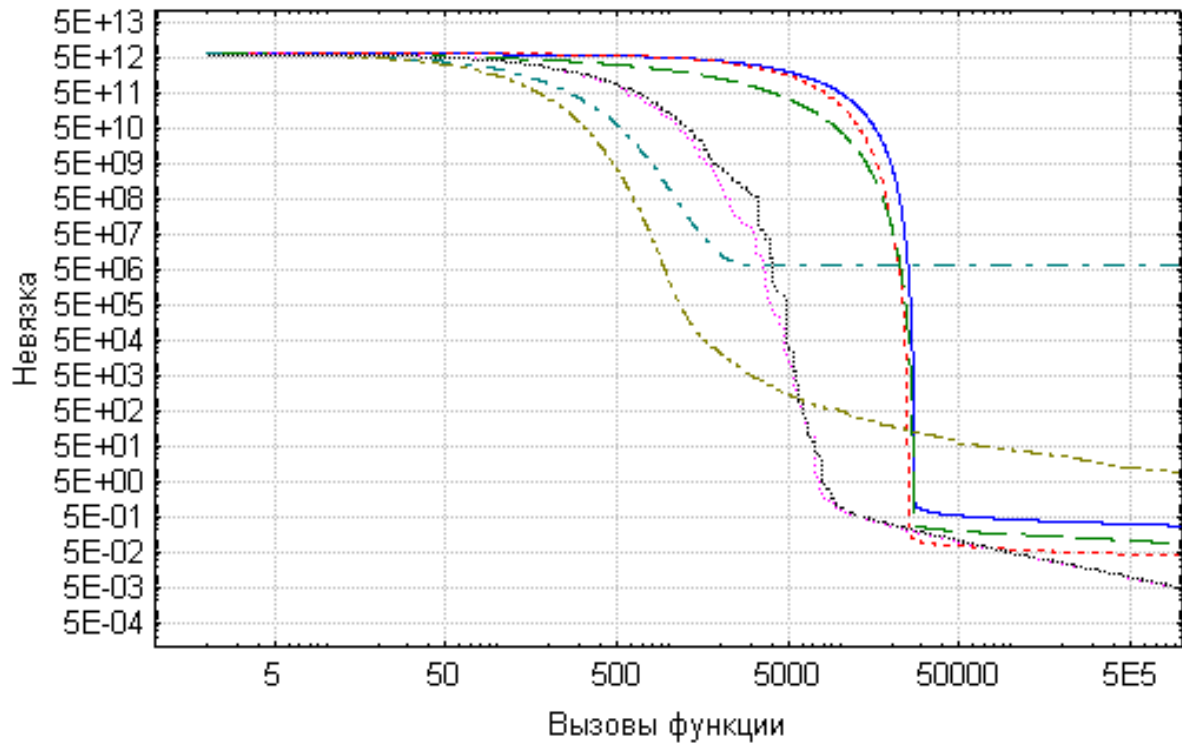
5.5.1 Выбор параметров и начальной точки

Начальная точка выбиралась равномерно распределенной на множестве минимизации. Для сверхбыстрого отжига использовались $m_i = 1$, $n_i = -160$, для сверхбыстрого тушения — $m_i = 1$, $n_i = -80$. Температура по всем координатам совпадала, стартовые значения подбирались опытным путем. Ниже приводится таблица стартовых температур:

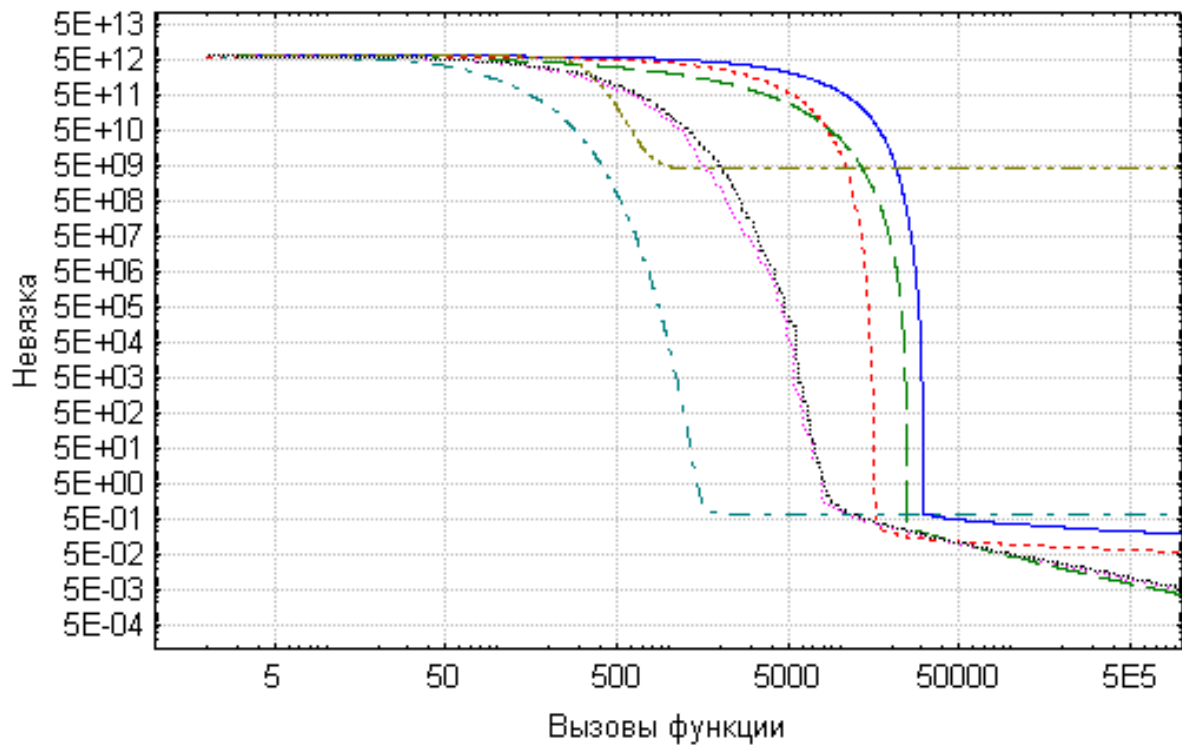
Метод	T_0
Больцмановский отжиг без уменьшения T после неудач	1
Больцмановский отжиг с уменьшением T после неудач	1
Отжиг Коши без уменьшения T после неудач	0.03
Отжиг Коши с уменьшением T после неудач	0.05
Сверхбыстрый отжиг без уменьшения T после неудач	0.001
Сверхбыстрый отжиг с уменьшением T после неудач	10^{-4}
Больцмановское тушение (по Коши) без уменьшения T после неудач	450
Больцмановское тушение (по Коши) с уменьшением T после неудач	1 000
Больцмановское тушение ($T_{k+1} = cT_k$) без уменьшения T после неудач	500
Больцмановское тушение ($T_{k+1} = cT_k$) с уменьшением T после неудач	40 000
Тушение Коши ($T_{k+1} = cT_k$) без уменьшения T после неудач	10
Тушение Коши ($T_{k+1} = cT_k$) с уменьшением T после неудач	800
Сверхбыстрое тушение без уменьшения T после неудач	0.5
Сверхбыстрое тушение с уменьшением T после неудач	10^{-4}

5.5.2 Результаты тестирования

Функция со случайной помехой (без уменьшения T после неудач)



Функция со случайной помехой (с уменьшением T после неудач)



5.5.3 Анализ результатов

Видно, что различные методы отжига хорошо работают с функциями с помехой, за исключением некоторых методов тушения. Известно, что скорость сходимости никакого метода, основанного только на значениях функции с невырожденными неубывающими помехами, не может быть быстрее, чем порядка корня числа итераций. Ниже приводится таблица невязки при использовании метода сверхбыстрого отжига в зависимости от числа вызовов функций:

Число итераций	Невязка
10000	0.763129571893273058
100000	0.043839954318724353
1000000	0.004662248394661734
10000000	0.000547701824698814
100000000	0.000056790943476683
1000000000	0.000007776614600166

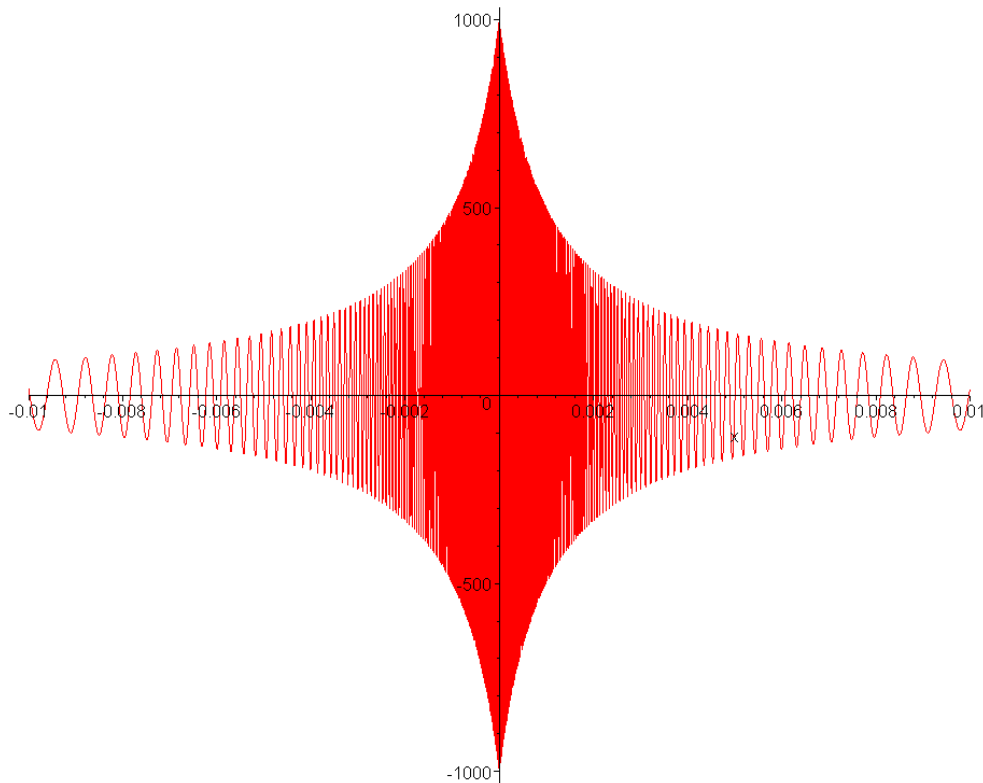
При тестировании использовалось усреднение по десяти запускам.

Непосредственно из приведённой таблицы видно, что скорость убывания невязки достаточно высока, и, скорее всего, совпадает с оптимальной теоретической оценкой. К сожалению, доказательство этого факта для метода отжига до сих пор не найдено.

5.6 Почти аналитическая функция с плохой сходимостью метода отжига

В целях иллюстрации того, что на некоторых почти аналитических функциях сходимость может быть не очень быстрой, была придумана функция

$$f(x) = \frac{\sin\left(\frac{1}{|x|+0.001}\right)}{|x|+0.001}.$$



Такая функция имеет несколько сотен локальных минимумов, причём их количество и глубина резко возрастает при приближении к началу координат, а длина “провала” — резко уменьшается. Кроме того, при выходе на бесконечность функция убывает к нулю. Во всех точках отличных от нуля, в частности, в точках глобального минимума, эта функция является аналитической.

Для поиска минимума этой функции методом отжига на $[-1000; 1000]$ пришлось ограничить выход за пределы этого промежутка даже у тех методов, которые не предусматривают такого ограничения, иначе x зачастую уходит на бесконечность при оптимизации.

5.6.1 Выбор параметров и начальной точки

Начальная точка выбиралась равномерно распределенной на отрезке минимизации. Для сверхбыстрого отжига использовались значения, приведенные в таблице:

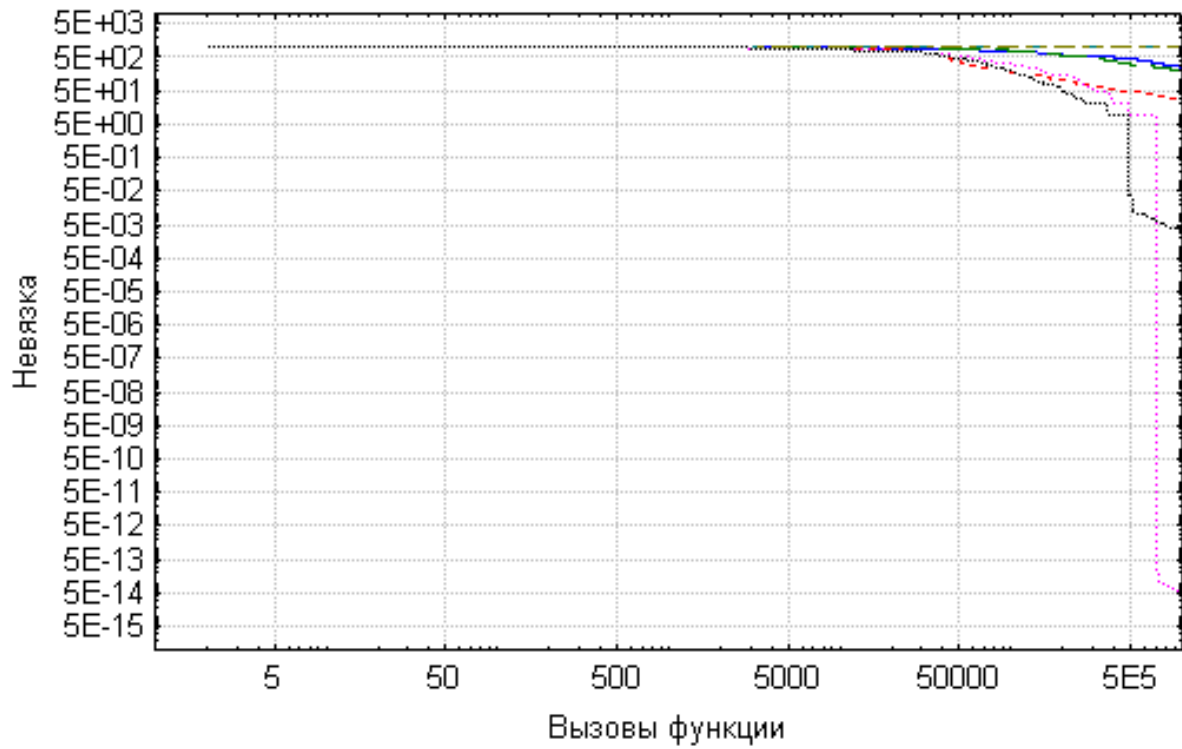
Метод	M	N
Сверхбыстрый отжиг без уменьшения T после неудач	1	16
Сверхбыстрый отжиг с уменьшением T после неудач	1	8
Сверхбыстрое тушение без уменьшения T после неудач	1	64
Сверхбыстрое тушение с уменьшением T после неудач	1	32

Ниже приведена таблица начальных значений температуры, подобранных опытным путём:

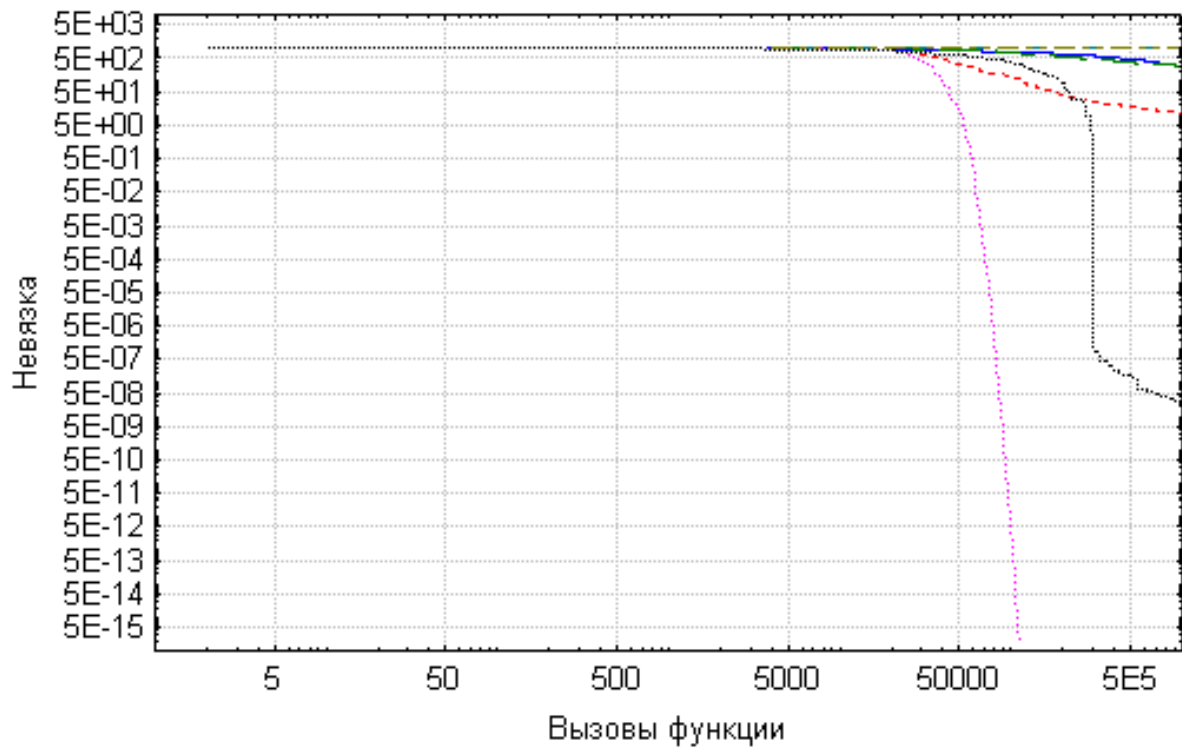
Метод	T_0
Больцмановский отжиг без уменьшения T после неудач	20
Больцмановский отжиг с уменьшением T после неудач	10
Отжиг Коши без уменьшения T после неудач	30000
Отжиг Коши с уменьшением T после неудач	30000
Сверхбыстрый отжиг без уменьшения T после неудач	10^{-13}
Сверхбыстрый отжиг с уменьшением T после неудач	1
Больцмановское тушение (по Коши) без уменьшения T после неудач	500 000
Больцмановское тушение (по Коши) с уменьшением T после неудач	400 000
Больцмановское тушение ($T_{k+1} = cT_k$) без уменьшения T после неудач	10^6
Больцмановское тушение ($T_{k+1} = cT_k$) с уменьшением T после неудач	10^7
Тушение Коши ($T_{k+1} = cT_k$) без уменьшения T после неудач	100 000
Тушение Коши ($T_{k+1} = cT_k$) с уменьшением T после неудач	10 000
Сверхбыстрое тушение без уменьшения T после неудач	10^{-7}
Сверхбыстрое тушение с уменьшением T после неудач	10^{-10}

5.6.2 Результаты тестирования

Функция для плохой сходимости (без уменьшения T после неудач)



Функция для плохой сходимости (с уменьшением T после неудач)



5.6.3 Анализ результатов

Видно, что большинство методов отжига на этой функции сходятся крайне медленно. Например, средняя невязка после 10^6 итераций метода Больцмана составляет около 250 притом, что разница между максимумом и минимумом этой функции меньше 2000. Отчасти это связано с убыванием функции при уходе на бесконечность.

Единственный метод, скорость сходимости которого приемлема в таких условиях — это метод сверхбыстрого отжига с уменьшением температуры после неудач.

6 Нахождение корней функций с помощью метода отжига

Применение метода отжига не ограничивается задачами нахождения минимума. С помощью этого метода можно также искать какой-либо корень заданной функции $f(x)$, так как эта задача эквивалентна нахождению минимума функции $|f(x)|$ или $f^2(x)$. В следующем разделе рассмотрена задача нахождения корня многочлена.

6.1 Многочлены от одной переменной

Как и при нахождении минимума, тестирование производилось на одном многочлене степени 2, 10 и 100. Генерация случайного многочлена производилась не по заданным минимумам и максимумам, а по заданным корням из интервала $[-1; 1]$, т.е. непосредственно по формуле (4).

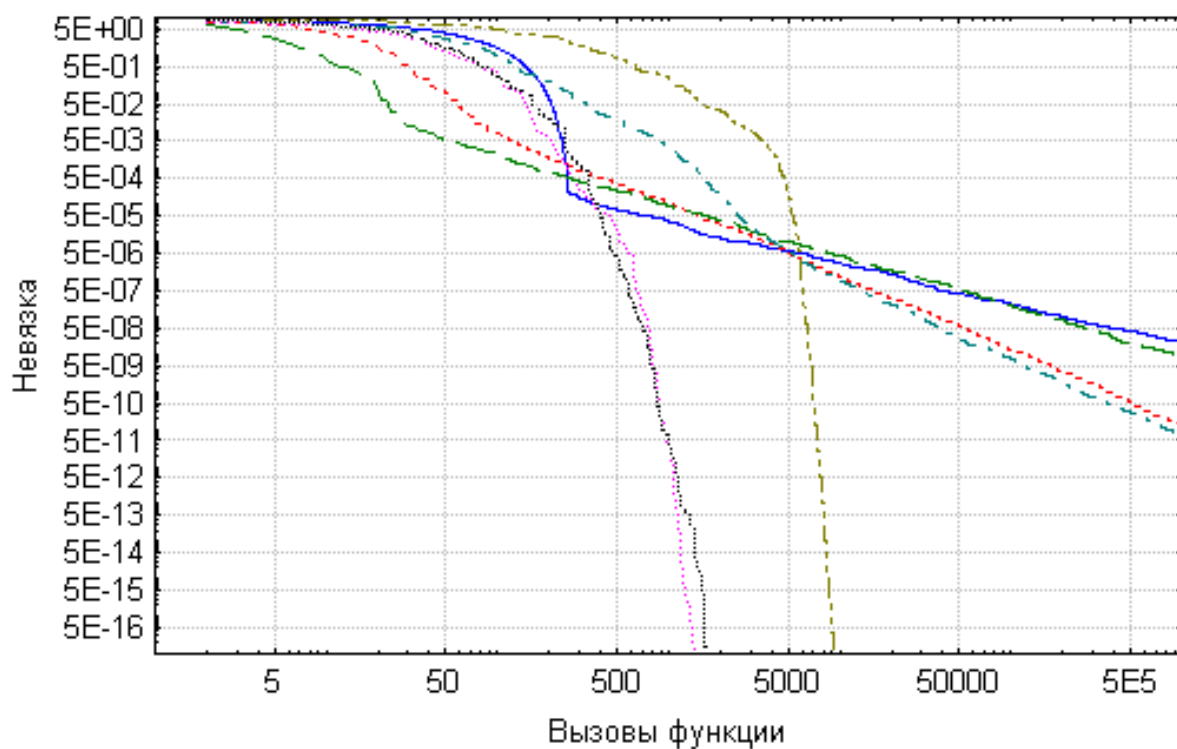
Функция для минимизации выбиралась как $|P(x)|$, отрезок, на котором искали корни — $[-5; 5]$.

Ниже приведены значения стартовой температуры, подобранные экспериментально:

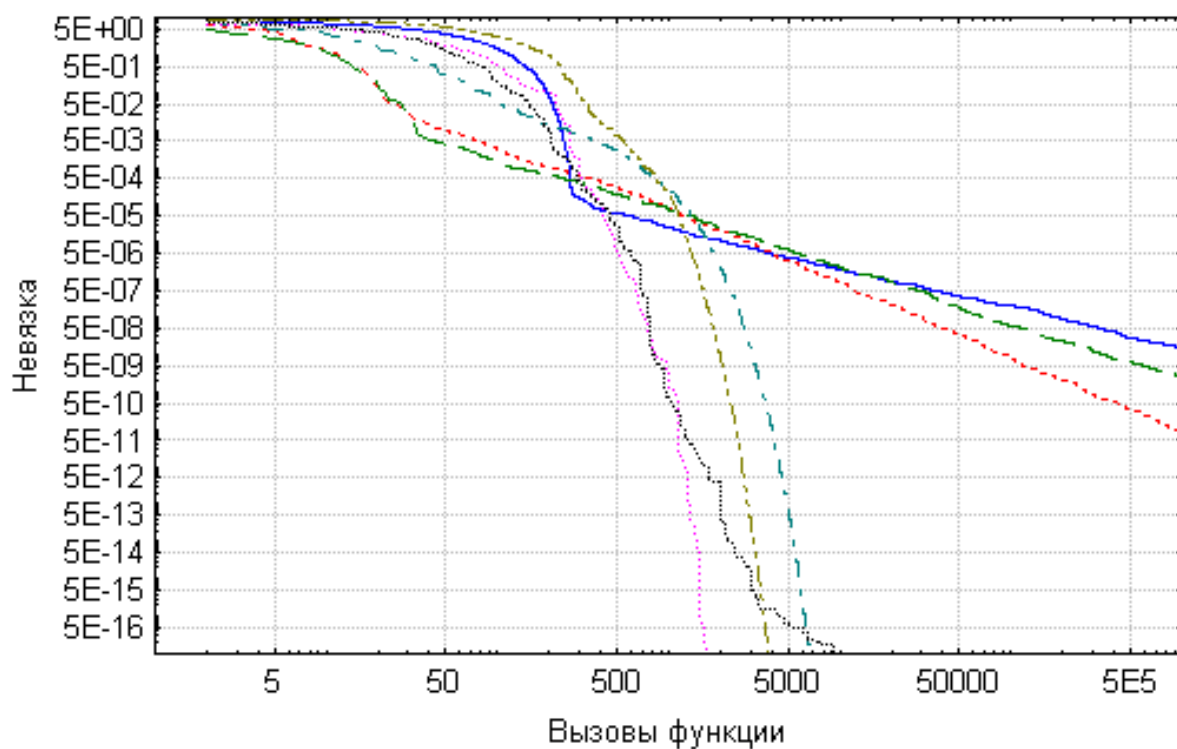
Метод	T_0
Больцмановский отжиг без уменьшения T после неудач	10^{-2}
Больцмановский отжиг с уменьшением T после неудач	10^{-2}
Отжиг Коши без уменьшения T после неудач	25
Отжиг Коши с уменьшением T после неудач	25
Сверхбыстрый отжиг без уменьшения T после неудач	10^{-15}
Сверхбыстрый отжиг с уменьшением T после неудач	10^{-15}
Больцмановское тушение (по Коши) без уменьшения T после неудач	8
Больцмановское тушение (по Коши) с уменьшением T после неудач	7
Больцмановское тушение ($T_{k+1} = cT_k$) без уменьшения T после неудач	4 000
Больцмановское тушение ($T_{k+1} = cT_k$) с уменьшением T после неудач	300
Тушение Коши ($T_{k+1} = cT_k$) без уменьшения T после неудач	260
Тушение Коши ($T_{k+1} = cT_k$) с уменьшением T после неудач	239
Сверхбыстрое тушение без уменьшения T после неудач	10^{-15}
Сверхбыстрое тушение с уменьшением T после неудач	10^{-13}

6.1.1 Результаты тестирования

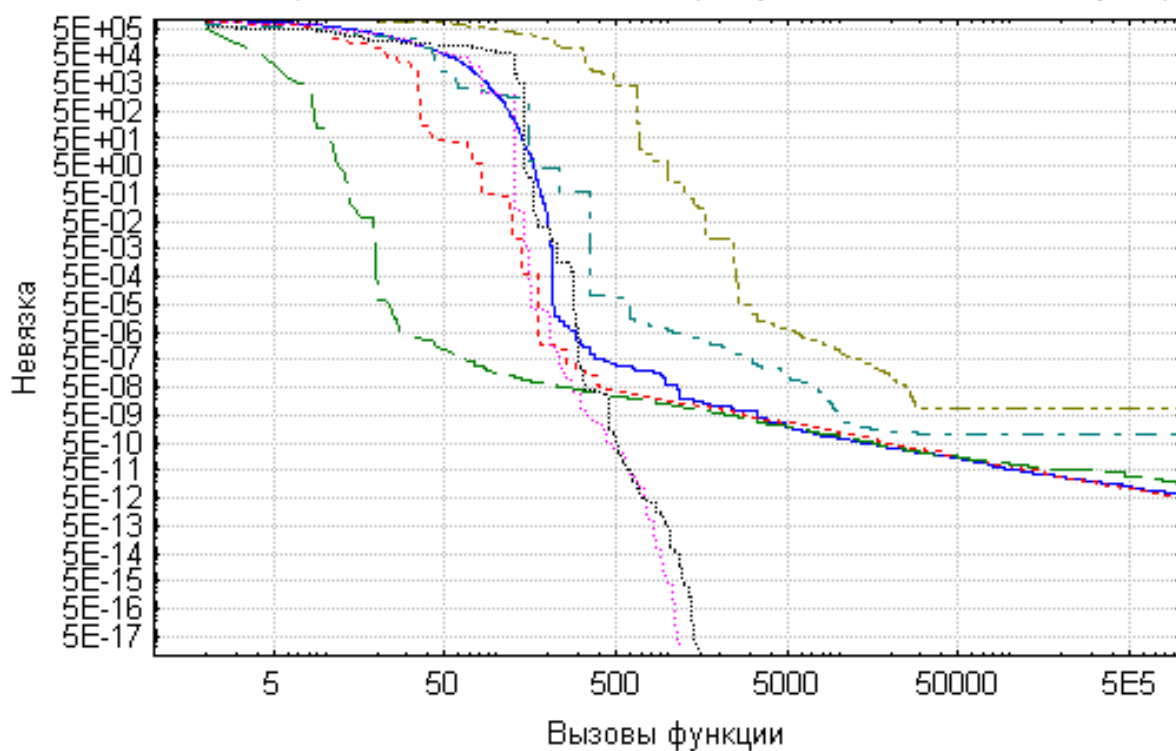
Поиск корня квадратного трёхчлена (без уменьшения T после неудач)



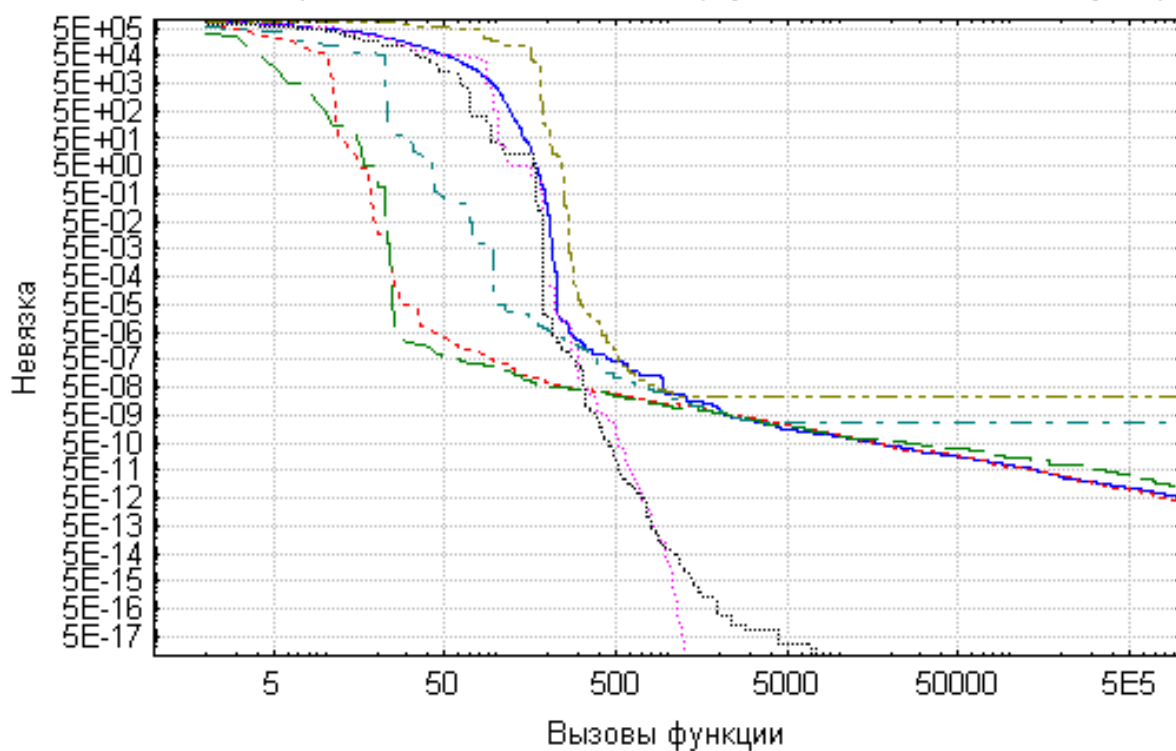
Поиск корня квадратного трёхчлена (с уменьшением T после неудач)



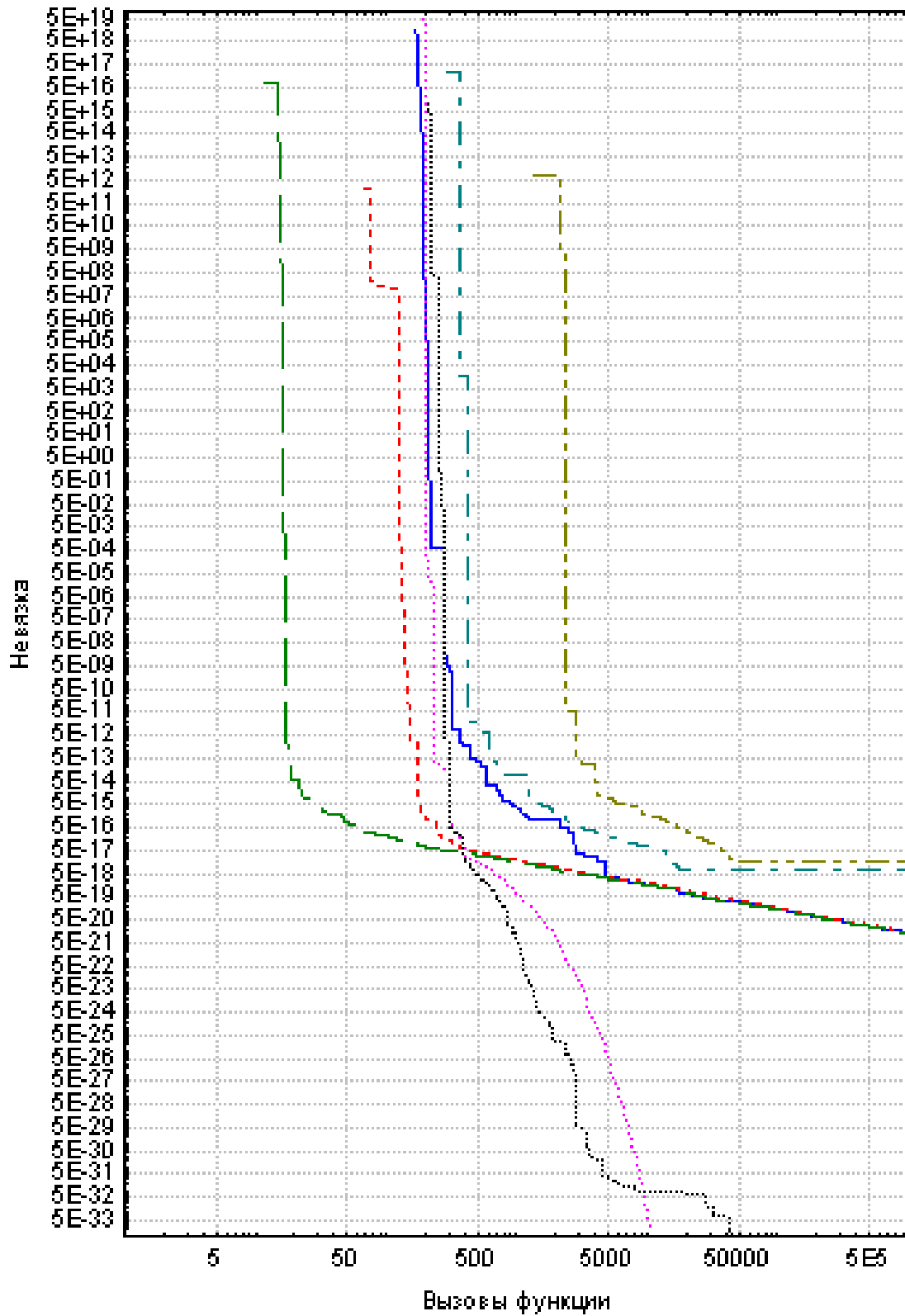
Поиск корня многочлена степени 10 (без уменьшения T после неудач)



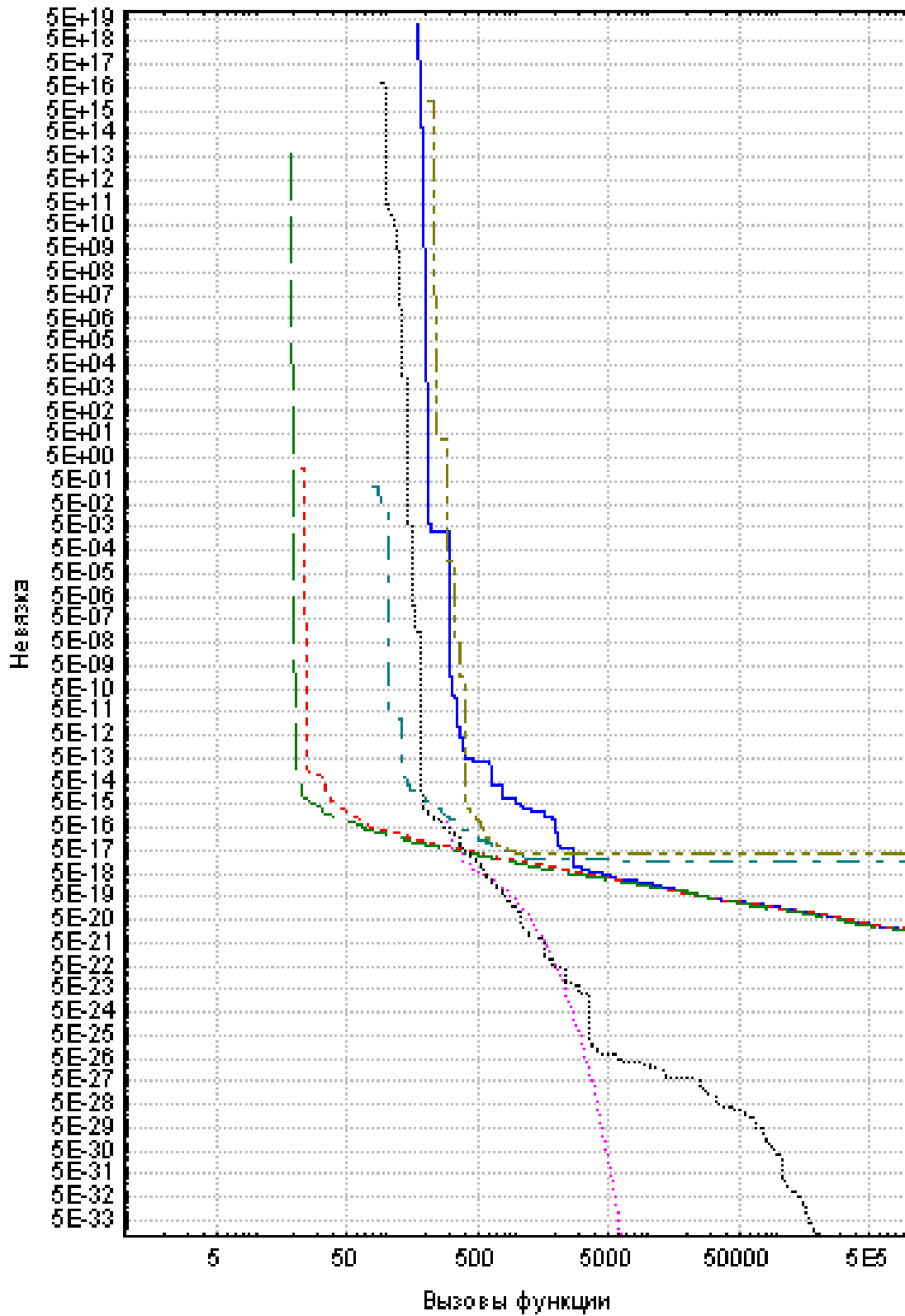
Поиск корня многочлена степени 10 (с уменьшением T после неудач)



Поиск корня многочлена степени 100 (без уменьшения T после неудач)



Поиск корня многочлена степени 100 (с уменьшением T после неудач)



6.1.2 Анализ результатов

Из графиков видна эффективная работа метода сверхбыстрого отжига, а в особенности его модификации с уменьшением температуры после неудач, для нахождения какого-либо корня многочлена. Остальные методы сходятся слишком медленно, и в практических целях вряд ли применимы.

7 О выборе параметров метода отжига

Задача выбора параметров имеет огромное значение для эффективности работы метода отжига. При неправильно выбранных параметрах скорость сходимости может оказаться в миллионы раз меньше, чем при оптимальных. В этом разделе рассмотрена задача выбора параметров, и предложен подход к её решению на основе “внешнего” метода отжига.

Обозначим пространство параметров метода отжига \mathcal{P} . Например, в случае Больцмановского отжига или отжига Коши \mathcal{P} представляет собой вещественную прямую значений T_0 , а в случае сверхбыстрого отжига \mathcal{P} — вещественное векторное пространство, содержащее значения T_0 и c для каждой из координат. Если же реализуется сверхбыстрый отжиг с одинаковыми T_0 и c_i для всех координат, то \mathcal{P} — двумерное вещественное пространство.

Результат работы заданного числа итераций метода отжига можно рассматривать как случайную величину, зависящую от точки в пространстве \mathcal{P} . Ниже приведен график зависимости невязки от T_0 при 10 000 итераций сверхбыстрого отжига на вырожденном случае задачи о пересечении двух кривых второго порядка на плоскости (две прямые $x + 2y + 3 = 0$ и $(x + y)^2 + 2x + 2y + 1 = 0$) при $c = e^{-12} \approx 3 \cdot 10^{-7}$:

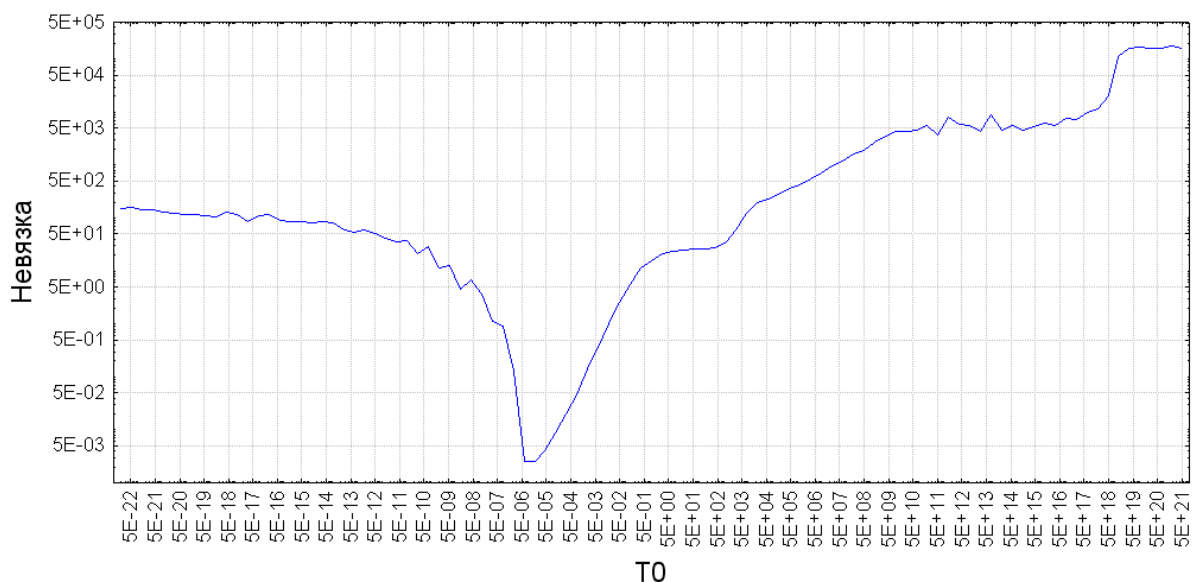
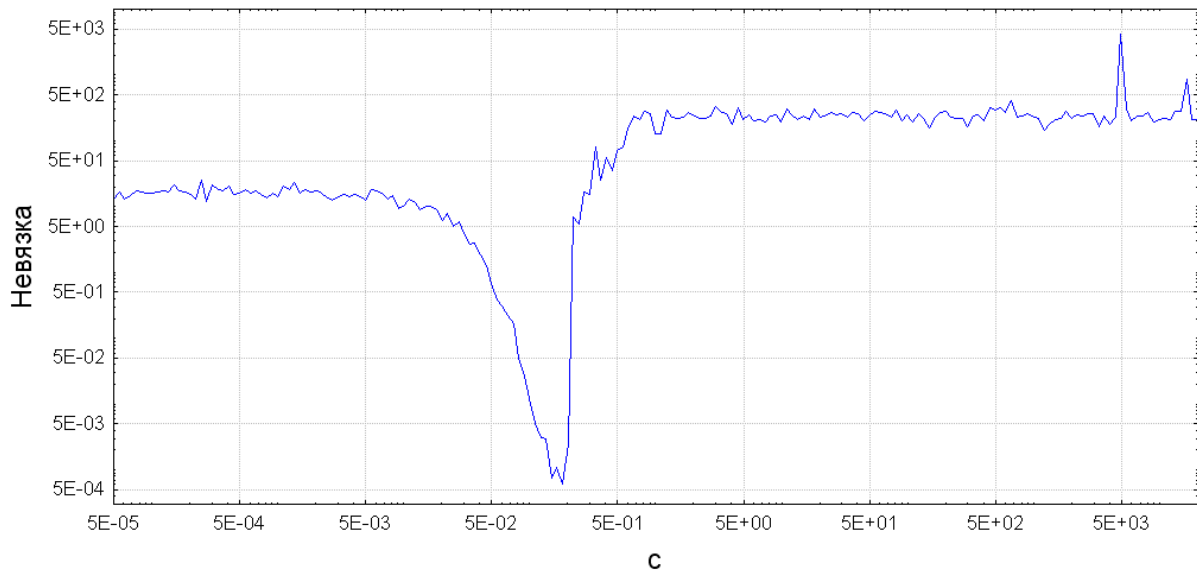


График получен усреднением по 256 запускам метода. Виден ярко выраженный минимум при $T_0 \approx 10^{-5}$. Однако зачастую такой минимум является не единственным. Ниже приведен график зависимости невязки от c на той же задаче при $T_0 = 1$:



При $c \approx 0.15$ также наблюдается локальный минимум.

Зачастую подбор оптимальной точки из пространства \mathcal{P} делается вручную. Однако этот подход занимает очень много времени, и не позволяет учесть многие эффекты, связанные с большим количеством минимумов, особенно на более сложных задачах.

Я же предлагаю подход на основе метода отжига. Обозначим $\xi(x)$ значение невязки, получаемое методом отжига при параметрах, соответствующих точке $x \in \mathcal{P}$. Ясно, что $\xi(x)$ — случайная величина при фиксированном x . Пусть $f(x) = \mathbf{E}\xi(x)$ — функция среднего значения семейства случайных величин ξ . Тогда измерение $\xi(x)$ можно рассматривать как измерение $f(x)$ с некоторой помехой $v(x)$, причём $\mathbf{E}v(x) = 0$. Таким образом, для поиска оптимального набора параметров годится любой метод, оптимизирующий функцию, измеряемую с помехами, в частности, метод отжига.

При применении метода отжига для нахождения параметров полезно учитывать следующие свойства:

- Координату для параметра, описывающего стартовую температуру, полезно задавать логарифмически, т.е. строить зависимость f от параметра $a_0 = \ln T_0$. Это существенно ускоряет поиск оптимальных параметров.
- Координаты для c_i в сверхбыстром отжиге можно задавать как и логарифмически, так и линейно.
- Допустимый диапазон для c_i , как правило, от 10^{-6} до 100.

В случае поиска оптимальных параметров отжига для какого-либо бесконечного набора задач удобно выбрать “базисный” набор, на котором будет искаться минимум суммарной или максимальной невязки.

7.1 Пример

Данная технология использовалась для подбора оптимальных параметров для решения задачи о пересечении двух кривых второго порядка на плоскости. С помощью программы, исходные коды которой можно найти в приложении, были подобраны значения $c = 1.938 \cdot 10^{-5}$, $T_0 = 1.712 \cdot 10^{-2}$, которые позволяют решать задачу с любыми

случайными коэффициентами, задающими невырожденные кривые второго порядка, в пределах от -10 до 10 , с получением точности не менее 10^{-8} при нескольких повторных запусках 10000 итераций метода отжига (для большинства кривых повторных запусков вообще не требуется). Для кривых, вырождающихся в прямые, точность может составлять 10^{-4} .

В качестве базисного набора использовались 6 пар кривых, подобранных вручную, и 3 случайных:

1. $f_1(x, y) = x^2 + y^2 - 1$, $f_2(x, y) = x^2 + y^2 + 2x$.
2. $f_1(x, y) = xy - 1$, $f_2(x, y) = x^2 + y$.
3. $f_1(x, y) = 2x^2 + y^2 - 3$, $f_2(x, y) = x^2 + 2y^2 - 3$.
4. $f_1(x, y) = xy - 1$, $f_2(x, y) = y^2 + x$.
5. $f_1(x, y) = x^2 - y^2 - 1$, $f_2(x, y) = x^2 - 2xy + y^2 - x - y$.
6. $f_1(x, y) = x^2 - y$, $f_2(x, y) = x^2 + y^2 - 1$.
7. $f_1(x, y) = 3x^2 - 8xy + 6x + 10y + 10$, $f_2(x, y) = 3x^2 - xy - 8y^2 - x + 10y + 4$.
8. $f_1(x, y) = x^2 + 10xy + 3y^2 + 2x - 7y + 10$, $f_2(x, y) = 5x^2 + xy + 9y^2 + 3x - 3y - 10$.
9. $f_1(x, y) = x^2 - 4xy + y^2 + 9y$, $f_2(x, y) = -9x^2 - 4xy + y^2 - 8x - 3y - 4$.

В качестве результата работы внутреннего метода отжига брался максимум невязки после 32 запусков отжига на каждой паре кривых. Всего было сделано 640 итераций внешнего отжига, во внутреннем отжиге делалось по 10000 итераций.

8 Заключение

Метод отжига является весьма эффективным алгоритмом случайного поиска глобального минимума. Версия алгоритма сверхбыстрого отжига сходится значительно быстрее других методов, поэтому в практических задачах её использование, видимо, наиболее целесообразно.

Пока мало изучены возможности метода отжига для решения других задач, сводящихся к оптимизационным. В этой работе частично рассмотрен вопрос нахождения корней функции. Для этой задачи в некоторых случаях метод сверхбыстрого отжига может оказаться весьма эффективным.

Надеюсь, что в ближайшее время будут получены теоретические результаты относительно скорости сходимости метода отжига. К сожалению, на настоящий момент единственные известные результаты касаются статистической гарантии сходимости (см. [5–7], [12]).

9 Список литературы

1. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of State Calculations by Fast Computer Machines,” J. Chemical Physics, 21, 6, pp. 1087–1092, June, 1953.

2. S. Kirkpatrick, C. D. Gelatt Jr., and M. P. Vecchi, "Optimization by Simulated Annealing," *Science*, 220, pp. 671–680, 1983.
3. L. Ingber, "Simulated Annealing: Practice versus theory", *Mathematical and Computer Modelling*, 18(11), pp. 29–57, 1993.
4. L. Ingber, B. Rosen, "Genetic Algorithms and Very Fast Simulated Reannealing: A Comparison", *Mathematical and Computer Modelling*, 16(11), pp. 87–100, 1992.
5. S. Geman, D. Geman, "Stochastic relaxation, Gibbs distribution and the Bayesian restoration in images", *IEEE Trans. Patt. Anal. Mac. Int.* 6(6), pp. 721–741, 1984.
6. H. H. Szu, R. L. Hartley, "Fast Simulated Annealing", *Physical Letters A*, 122, pp. 157–162, 1987.
7. L. Ingber, "Very fast simulated re-annealing", *Mathematical and Computer Modelling*, 12, pp. 967–973, 1989.
8. X. Yao, "A New Simulated Annealing Algorithm", *International Journal of Computer Mathematics*, 56, pp. 161–168, 1995.
9. L. Ingber, "Adaptive simulated annealing (ASA): Lessons learned", *Journal "Control and Cybernetics"*, 1995
10. B. Cohen, "Training Synaptic Delays in a Recurrent Neural Network", Thesis submitted towards the degree "Master of Science" at Tel-Aviv University, 1994
11. Ф. Уоссермен, "Нейрокомпьютерная техника: Теория и практика", Перевод на русский язык, Ю. А. Зуев, В. А. Точенов, 1992.
12. X. Yao "Simulated Annealing with Extended Neighbourhood", *International Journal of Computer Mathematics*, 40, pp. 169–189, 1991
13. L. Ingber, R.P. Mondescu, "Optimization of Trading Physics Models of Markets", *IEEE Trans. Neural Networks*, 12(4), pp. 776–790, 2001
14. M. C. Forman, A. Aggoun, M. McCormick, "Simulated Annealing for Optimisation and Characterisation of Quantisation Parameters in Integral 3D Image Compression" *The Institute of Mathematics and its Applications / Horwood*, pp. 393–413, 2000
15. M. C. Forman, "Compression of Integral Three Dimensional Television Pictures", Ph.D. Thesis at De Montfort University Leicester, United Kingdom, 2000
16. L. Ingber, J.K. Wilson "Statistical mechanics of financial markets: Exponential modifications to Black-Scholes", *Mathematical Computer Modelling* 31(8/9), pp. 167–192, 2000
17. C. Jeong and M. Kim, "Fast Parallel Simulated Annealing for Traveling Salesman Problem on SIMD Machines with Linear Interconnections", *Parallel Computing*, 17, pp. 221–228, 1991.
18. X. Yao, "Call Routing by Simulated Annealing", *International Journal of Electronics*, October 1995