

Объединенный институт ядерных исследований

На правах рукописи

Жабицкая Евгения Игоревна

МЕТОД АСИНХРОННОЙ
ДИФФЕРЕНЦИАЛЬНОЙ ЭВОЛЮЦИИ
ДЛЯ ЧИСЛЕННОГО ИССЛЕДОВАНИЯ
МНОГОПАРАМЕТРИЧЕСКИХ МОДЕЛЕЙ
ФИЗИЧЕСКИХ СИСТЕМ

Специальность: 05.13.18 — математическое моделирование,
численные методы и комплексы программ

Диссертация на соискание ученой степени
кандидата физико-математических наук

Научный руководитель:
доктор физико-математических наук
Земляная Елена Валериевна
Научный руководитель:
доктор физико-математических наук
Киселев Михаил Алексеевич

Дубна – 2016

Оглавление

Введение	4
В.1. Глобальная минимизация	4
В.2. Актуальность темы диссертации	6
В.3. Цели и задачи диссертационной работы	9
В.4. Научная новизна	10
В.5. Теоретическая значимость и практическая ценность работы . .	11
В.6. Положения, выносимые на защиту	12
В.7. Достоверность и апробация результатов	14
В.8. Публикации и личный вклад автора в работу	15
В.9. Структура и объем работы	16
1. Метод асинхронной дифференциальной эволюции	18
1.1. Сравнение асинхронного и классического методов ДЭ	18
1.2. Оптимальные параметры метода АДЭ	27
1.3. Асинхронная дифференциальная эволюция с рестартом	38
1.4. Асинхронная дифференциальная эволюция с кроссовером, задаваемым адаптивной корреляционной матрицей	45
1.5. Программная реализация параллельного АДЭ-алгоритма	50
Основные результаты, представленные в главе 1	59
2. Оценка параметров пион-нуклонной амплитуды в рамках микроскопической модели пион-ядерного рассеяния с использованием АДЭ	61
2.1. Постановка задачи, основные формулы	63
2.2. Описание комплекса программ	67
2.3. Обсуждение численных результатов	72
2.4. Результативность АДЭ по сравнению с другими методами оптимизации	80
Основные результаты, представленные в главе 2	85

3. Анализ структуры однослойных везикул ДМФХ по данным МУРР в рамках модифицированной модели РФФ с использованием АДЭ	87
3.1. Постановка задачи, основные формулы	88
3.2. Описание комплекса программ	94
3.3. Обсуждение численных результатов	95
3.4. Результативность АДЭ по сравнению с другими методами оптимизации	100
3.5. Оценка ускорения АДЭ при параллельных вычислениях	104
Основные результаты, представленные в главе 3	106
Заключение	109
Список публикаций по теме диссертации	111
Список цитируемой литературы	115
Приложения	124
П. I. Используемые в работе тестовые функции из СЕС-2005 [78] .	124
П. II. Доказательства теорем 1.2.1 — 1.2.3 из раздела 1.2 диссертации	125
П. III. Результативность АДЭ-АСМ метода	136

Введение

В диссертации разработаны эффективные модификации метода дифференциальной эволюции (ДЭ)¹ [75] — одного из часто применяемых алгоритмов глобальной минимизации функции действительных переменных. Это метод асинхронной дифференциальной эволюции (АДЭ)² [7, 8] и метод АДЭ с рестартом и с адаптивной корреляционной матрицей (АДЭ-АКМ)³ [9, 10]. С их помощью проведено численное исследование многопараметрических моделей двух физических систем [1–6, 11–21]. Первая — это модель пион-ядерного рассеяния на основе микроскопического оптического потенциала⁴ и релятивистского волнового уравнения [1, 2, 15–18]. Вторая система — это обобщенная модель разделенных формфакторов⁵, адаптированная для анализа структуры везикулярных систем⁶ по данным малоуглового синхротронного рентгеновского рассеяния⁷ [3, 4, 19–21]. Предложенные новые методы минимизации позволили успешно провести исследование задач, рассмотренных в диссертации, они могут также использоваться в других научных исследованиях.

В.1. Глобальная минимизация

Оптимизационные задачи естественно возникают в науке и прикладных исследованиях, когда необходимо определить вектор параметров $\vec{\theta}^*$, минимизирующий⁸ целевую функцию $f(\vec{\theta})$:

$$f(\vec{\theta}^*) \leq f(\vec{\theta}) \in \mathbb{R}, \quad \forall \vec{\theta} \in \Omega \subset \mathbb{R}^D, \quad \vec{\theta} = \{\theta_j\}_{j=0, \dots, D-1}, \quad (\text{B.1})$$

$$\text{или} \quad \vec{\theta}^* = \underset{\vec{\theta} \in \Omega \subset \mathbb{R}^D}{\text{Argmin}} f(\vec{\theta}). \quad (\text{B.2})$$

¹Engl.: Differential Evolution (DE).

²Engl.: Asynchronous Differential Evolution (ADE).

³Engl.: Asynchronous Differential Evolution with Adaptive Covariance Matrix (ADE-ACM).

⁴Engl.: Microscopic Optical Potential (MOP).

⁵Engl.: Separated Form Factor (SFF).

⁶В данной работе под везикулами понимаются близкие к сферическим наноструктуры с оболочкой, представляющей собой липидную мембрану.

⁷Engl.: Small Angle X-ray Scattering (SAXS).

⁸Если искомым оптимум в задаче соответствует нахождению максимума целевой функции $g(\theta)$, то оптимизационная задача сводится к минимизации целевой функции $f(\theta) = -g(\theta)$.

Здесь \mathbb{R} — множество всех вещественных чисел; \mathbb{R}^D — вещественное векторное пространство размерности D ; Ω — вещественное пространство поиска решения размерности D .

Существует множество методов решения задачи (В.1) [23, 24]. К ним относятся, в первую очередь, *детерминированные прямые методы многомерной оптимизации* [33, 38, 43, 53, 54, 59]. При решении гладких, дифференцируемых унимодальных задач эффективны *детерминированные методы первого и второго порядков* [27, 28, 30–32, 36–38, 40, 42, 44, 48, 55, 60], обладающие свойством быстрой сходимости. В других случаях для многомерной локальной оптимизации часто применяются методы *случайного поиска*.

В отличие от задач локальной минимизации, задачи на глобальный минимум (ГМ) требуют других подходов. В случае многоэкстремальной и/или негладкой целевой функции $f(\vec{\theta})$ глобальный минимум (В.1) может быть найден с помощью методов *многомерной глобальной оптимизации*, например, *методов повторяющегося поиска* [22, 51], *стохастических методов оптимизации*, таких как метод Монте-Карло [52], и *эволюционных алгоритмов* (ЭА) ⁹ [29, 34, 42, 49, 50, 56, 57, 61, 65, 71–73, 75].

В эволюционных алгоритмах используется итерационный процесс, имитирующий развитие биологических популяций, который моделирует базовые положения теории биологической эволюции: процессы отбора, мутации и воспроизводства. Набор агентов, или векторов в пространстве поиска Ω , принято называть популяцией. Каждому агенту соответствует значение его пригодности в окружающей среде, равное соответствующему значению целевой функции f . Новые агенты порождаются посредством мутации и рекомбинации (кроссовера). В результате отбора менее пригодные агенты отбрасываются. Таким образом, популяция эволюционирует в соответствии с правилами отбора и целевой функцией, задаваемой окружающей средой [35, 47, 70–73, 75–77]. Дифференциальная эволюция (ДЭ) [68, 70–73, 75–77] — алгоритм прямого поиска решения задачи (В.1) — относится к классу эволюционных алгоритмов¹⁰ и подходит для минимизации недифференцируемых, нелинейных, мультимодальных функций от многих аргументов. Инженеры и исследователи успешно применяют ДЭ для решения оптимизационных задач в различных областях науки. Развитие этого метода

⁹Engl.: Evolutionary Algorithm (EA).

¹⁰См. также часть 1.1.1, с. 19 – 21 диссертации.

описано в книгах [67, 70], а также в обзорах [65, 66].

Большинство эволюционных алгоритмов могут быть разделены на две большие группы: *популяционные* (generational) алгоритмы, которые обновляют всю популяцию на каждой итерации и *квазистационарные* (incremental, steady-state) алгоритмы¹¹, которые одновременно обновляют только часть потенциальных решений. Традиционно метод ДЭ реализуют в виде популяционного алгоритма. В данной работе впервые операции мутации, рекомбинации (кроссовера) и отбора, характерные для классической ДЭ, задействованы в предельной форме квазистационарного алгоритма. Таким образом, сформулированы методы асинхронной дифференциальной эволюции [7], АДЭ с рестартом и адаптивным кроссовером [9, 10].

В.2. Актуальность темы диссертации

Как разработанные в диссертации методы, так и результаты решения конкретных задач, полученные с использованием этих методов, востребованы и актуальны.

Актуальность разработки и компьютерной реализации методов АДЭ и АДЭ-АКМ. Новые методы актуальны для многих физических задач, исследование которых требует глобальной минимизации. В частности, это относится к задачам, решенным в диссертации. В самом деле, параметры микроскопического оптического потенциала для модели пион-ядерного рассеяния и параметры модели разделенных формфакторов, описывающей дифференциальные сечения малоуглового рентгеновского рассеяния, оцениваются путем минимизации взвешенной квадратичной невязки между предсказаниями теоретической модели и экспериментальными данными. Зависимость этой невязки от вектора искомых параметров задает целевую функцию. Рассчитываемые целевые функции обладают сложным профилем и имеют несколько локальных минимумов. Это приводит к тому, что при использовании методов локального поиска [37, 45, 53]¹² вероятность определения глобального минимума в области допустимых значений параметров систем низкая, в то время как вероятность нахождения оптимального решения методами ДЭ значительно выше [5, 19]. С другой

¹¹В близком подклассе эволюционных алгоритмов — подклассе генетических алгоритмов — альтернативой традиционному поколенческому подходу является использование квазистационарного (steady-state) подхода [62, 63].

¹²Например, метода симплексов [53] или квазиньютоновского метода с формулой Давидона-Флетчера-Пауэлла [37] для обновления гессиана, реализованных в программе MINUIT [45].

стороны, определение параметров, отвечающих за структуру полидисперсной популяции везикул ДМФХ, по данным малоуглового синхротронного рассеяния в модели разделенных формфакторов требует глобальной минимизации многоэкстремальной функции с 9 частично коррелированными переменными. Эта задача требует существенных объемов вычислений, при этом особое значение приобретает развитие алгоритмов, обеспечивающих их эффективную параллельную реализацию на многопроцессорных системах. Реализованный на базе АДЭ-минимизации комплекс программ позволяет эффективно распараллелить вычисления, более чем в 50 раз уменьшив время расчетов.

Предложенные в диссертации алгоритмы асинхронной дифференциальной эволюции обладают рядом преимуществ по сравнению с другими популярными подходами. Они (i) не требуют вычисления производных; (ii) позволяют получить сравнимые со стандартной ДЭ скорость и вероятность сходимости¹³ к глобальному минимуму [7, 8] и в ряде случаев улучшить эти показатели [9, 10]; (iii) позволяют легко и эффективно распараллелить вычисления за счет равномерной и полной загрузки доступных вычислительных узлов при расчетах в параллельном режиме [7, 12].

Таким образом, разработка и программная реализация методов АДЭ актуальны и востребованы как для рассматриваемых в диссертации задач, так и для многих других научных исследований.

Актуальность микроскопической модели пион-ядерного рассеяния. Исследованию пион-ядерного рассеяния посвящено значительное число теоретических и экспериментальных работ [87]. Тем не менее, задача построения реалистичного, теоретически обоснованного пион-ядерного потенциала, необходимого для моделирования не только упругого рассеяния пионов, но и более сложных пион-ядерных взаимодействий, остается актуальной.

Большинство известных работ по расчетам пион-ядерных дифференциальных сечений основано на применении тех или иных феноменологических форм потенциала либо на использовании микроскопического потен-

¹³ Под оценкой средней скорости сходимости понимается величина, обратная среднему времени, необходимому для достижения минимума с заданной точностью. Так как расчет целевой функции требует значительно больших затрат компьютерного времени, чем операции самого алгоритма, оценка средней скорости сходимости может быть аппроксимирована величиной, обратной среднему количеству вычислений целевой функции, необходимому для достижения минимума с заданной точностью. Под оценкой вероятности сходимости понимается отношение числа успешных сошедшихся к глобальному минимуму попыток к общему числу попыток минимизации функции при помощи данного алгоритма.

циала Кислингера [98, 99], где учитывается вклад в пион-ядерное рассеяние s -, p - и d -волн πN -амплитуды. Все они имеют 6 и более свободных параметров.

В диссертации используется гибридный подход [84], сочетающий трехпараметрическую модель микроскопического оптического потенциала (МОП) на базе высокоэнергетического приближения (ВЭП)¹⁴ и расчет наблюдаемых характеристик на основе релятивистского волнового уравнения (РВУ) [1, 2, 15]. Как показывают расчеты, качество описания экспериментальных данных в рамках трехпараметрической модели сопоставимо с результатами более сложных моделей с большим числом параметров [2]. При этом параметры микроскопической модели имеют простой смысл и характеризуют динамику изменения амплитуды рассеяния пионов на внутриядерных нуклонах в зависимости от энергии пионов [2]. Это позволяет исследовать эффект влияния ядерной среды на процесс рассеяния.

Актуальность моделирования структуры везикулярных систем на основе фосфолипидов обусловлена практическими приложениями в области фармакологии и косметологии, где везикулы (наносферы) используются в качестве переносчиков лекарств и других активных компонент. Кроме того, исследование бислоя оболочки везикул играет важную роль в структурной биологии и биофизике, поскольку позволяет получить новую информацию о структуре и свойствах биологических мембран.

Для анализа структуры полидисперсных везикулярных систем развит и успешно используется метод разделенных формфакторов (РФФ). Ранее этот метод применялся [117, 129] для анализа данных малоуглового рассеяния нейтронов (МУРН)¹⁵. В диссертации метод разделенных формфакторов адаптирован для моделирования структуры полидисперсной популяции однослойных везикул димиристоилфосфатидилхолина (ДМФХ)¹⁶ по данным малоуглового рентгеновского рассеяния (МУРР) [3, 4, 19–21]. Предложенная в работе модификация РФФ-подхода для обработки МУРР-спектров учитывает флуктуации параметров модели бислоя оболочки везикул и использует более сложный, по сравнению с нейтронным рассеянием, профиль распределения плотности длины рассеяния поперек мембраны.

¹⁴Engl.: High-energy approximation (HEA).

¹⁵Engl.: Small Angle Neutron Scattering (SANS).

¹⁶Engl.: Dimiristoilphosphatydilcholin (DMPC).

Это позволило достичь лучшего соответствия теоретической модели экспериментальным данным.

В.3. Цели и задачи диссертационной работы

Цели диссертации — это разработка, обоснование, компьютерная реализация и тестирование эффективных вычислительных методов глобальной минимизации, создание комплексов проблемно-ориентированных программ и численное исследование микроскопической модели пион-ядерного рассеяния и модели разделенных формфакторов, описывающей структуру везикулярной системы.

Задачи диссертации:

Разработать и исследовать новые эффективные алгоритмы глобальной минимизации.

- Разработать на базе метода ДЭ метод асинхронной дифференциальной эволюции (АДЭ), метод асинхронной дифференциальной эволюции с рестартом и кроссовером, определяемым адаптивной корреляционной матрицей (АДЭ-АКМ).
- Аналитическими методами получить ограничения на управляющие параметры алгоритмов ДЭ и АДЭ, необходимые для предупреждения преждевременной сходимости¹⁷ алгоритма.
- Путем методических расчетов для ряда тестовых функций оценить скорость и вероятности сходимости предложенных методов АДЭ; сравнить с классическим вариантом ДЭ и другими методами минимизации.

Провести численное исследование микроскопической модели пион-ядерного упругого рассеяния.

- Разработать комплекс программ, реализующий АДЭ-подгонку параметров микроскопической модели пион-ядерного упругого рассеяния по экспериментальным данным.
- Определить параметры микроскопического оптического потенциала по экспериментальным дифференциальным сечениям упругого рассеяния пионов на ядрах ^{28}Si , ^{40}Ca , ^{54}Ni и ^{208}Pb в диапазоне кинетических энергий налетающих пионов от 130 до 291 МэВ.

¹⁷ Преждевременная сходимость — это потеря разнообразия в популяции, которой оперирует эволюционный алгоритм. Эта потеря разнообразия приводит к искусственному ограничению на пространство поиска возможных решений, доступных алгоритму в процессе эволюции.

- Провести анализ влияния ядерной среды на характеристики пион-нуклонной амплитуды.

Провести численное исследование структуры везикулярных систем.

- Разработать модификацию модели разделенных формфакторов для учета флуктуаций параметров липидного бислоя в везикулярных системах.
- Разработать комплекс программ, реализующий параллельную АДЭ-подгонку параметров модели везикулярной системы по данным малоуглового рентгеновского рассеяния в режиме параллельных вычислений.
- Провести численное исследование модели, описывающей структуру полидисперсной популяции везикул ДМФХ в 40% растворе сахарозы.

Оценить эффективность применения разработанных методов минимизации.

- Сравнить на примере задачи о пион-ядерном рассеянии и задачи об исследовании параметров везикулярной системы эффективность новых алгоритмов (вероятность нахождения оптимального решения) по сравнению с другими методами минимизации, использовавшимися ранее для решения аналогичных задач.
- Оценить ускорение вычислений, достигаемое при решении перечисленных выше задач, при расчетах с использованием различного количества параллельных вычислительных узлов.

В.4. Научная новизна

В работе впервые операции мутации, кроссовера и отбора, характерные для классической ДЭ, задействованы в предельной форме квазистационарного алгоритма. Таким образом, сформулирован метод асинхронной дифференциальной эволюции [7], в котором на каждом шаге обновляется только один из членов популяции.

Предложен метод асинхронной дифференциальной эволюции с рестартом [9]. Метод приводит размер популяции в соответствие со сложностью решаемой задачи.

Впервые на основе проведенного аналитического исследования получены ограничения на управляющие параметры для ряда стратегий методов АДЭ и ДЭ [8], которые являются необходимым условием предупреждения преждевременной сходимости алгоритмов.

Предложена новая адаптивная схема для операции кроссовера [10], в рамках которой эта операция модифицирована и зависит от парных корреляций между параметрами задачи. Алгоритм асинхронной дифференциальной эволюции с кроссовером, задаваемым адаптивной корреляционной матрицей (АДЭ-АКМ), со встроенной адаптивной схемой выбора масштабирующего фактора [10] и процедурой рестарта с автоматическим увеличением размера популяции [9] не требует настройки параметров пользователем и превосходит другие известные в литературе адаптивные варианты ДЭ по скорости и вероятности сходимости.

Впервые на основе тестовых расчетов получены оценки для скорости и вероятности сходимости методов АДЭ [8], АДЭ с рестартом [9] и АДЭ-АКМ [10] для тестовых функций из наборов тестовых задач СЕС-2005¹⁸ [5] и ВВОВ-2012¹⁹ [10] при различных значениях управляющих параметров алгоритмов.

Впервые метод АДЭ применен для численного исследования микроскопической модели пион-ядерного рассеяния [1, 16–18]. Впервые трехпараметрическая модель на основе микроскопического оптического потенциала и релятивистского волнового уравнения Шредингера использована для расчета дифференциальных сечений упругого рассеяния пионов на ядрах ^{28}Si , ^{40}Ca , ^{58}Ni , ^{208}Pb в диапазоне кинетической энергии налетающих пионов от 130 до 291 МэВ [2, 15]. На этой основе получены и проанализированы характеристики пион-нуклонной амплитуды в ядерной среде, выполнено их сравнение с соответствующими параметрами рассеяния пионов на свободных нуклонах [2].

Впервые метод АДЭ применен для анализа структуры везикулярных систем [21]. Учтены флуктуации параметров модели бислоя, позволившие применить метод разделенных формфакторов для анализа данных МУРР [3–5, 19, 20]. Впервые на этой основе получены параметры, характеризующие структуру полидисперсной популяции однослойных везикул ДМФХ в 40% растворе сахарозы [4, 20].

В.5. Теоретическая значимость и практическая ценность работы

Показано, что методы АДЭ и АДЭ-АКМ могут быть эффективно использованы для нахождения глобального минимума функции действи-

¹⁸См. [78].

¹⁹См. [41].

тельных переменных [1, 3, 5–10].

Показано, что теоретический подход на основе трехпараметрической микроскопической модели обеспечил согласующиеся с экспериментальными данными значения дифференциальных сечений упругого пион-ядерного рассеяния и позволил проанализировать характеристики пион-нуклонной амплитуды в ядерной среде [2, 15].

Показано, что модифицированная модель разделенных формфакторов с учетом флуктуаций бислоя применима для исследования структуры полидисперсных везикулярных систем на основе данных МУРР [4, 19].

Созданы проблемно-ориентированные комплексы программ (для оценки параметров модели пион-ядерного рассеяния и для оценки параметров везикулярных систем), которые в настоящее время используются для дальнейших численных исследований [1, 3, 5, 15].

На основе проведенного исследования получены численные результаты, которые являются физически значимыми при моделировании пион-ядерного рассеяния [2, 15] и исследовании везикулярных систем [4, 19]. В частности, значения параметров, полученные в [2], используются для моделирования неупругого пион-ядерного рассеяния [102]. Параметры, определяющие структуру ДМФХ, позволили на основе сравнительного анализа сделать заключение о везикулярной структуре фосфолипидной транспортной наносистемы [121].

Разработанные методы глобальной минимизации АДЭ и АДЭ-АКМ могут быть использованы в различных областях науки, а также для прикладных исследований.

В.6. Положения, выносимые на защиту

Разработана модификация алгоритма ДЭ — метод асинхронной дифференциальной эволюции (АДЭ) [7].

- Асинхронизация позволила повысить эффективность параллельной реализации и тем самым ускорить поиск глобального минимума пропорционально количеству задействованных узлов при расчетах на параллельных системах, использующих несколько десятков вычислительных узлов [5–7, 11, 12].
- Получены аналитические ограничения на управляющие параметры алгоритма для стратегий АДЭ и ДЭ [8].

- Разработан метод АДЭ с новым типом адаптивного кроссовера, учитывающим парные корреляции между аргументами целевой функции [10]. Предложенный новый алгоритм способен идентифицировать коррелированные группы переменных и благодаря этому эффективно решать минимизационные задачи с разделяемыми, неразделяемыми и частично разделяемыми переменными.
- АДЭ с адаптивным кроссовером, адаптивной схемой выбора масштабирующего фактора [10] и автоматическим увеличением размера популяции посредством рестарта [9] не требует подгонки параметров алгоритма пользователем [10].
- АДЭ с кроссовером, задаваемым адаптивной корреляционной матрицей, превосходит другие адаптивные варианты ДЭ по скорости и вероятности сходимости на наборах тестовых задач СЕС-2005²⁰ [5] и ВВОВ-2012²¹ [10].

С использованием АДЭ разработан проблемно-ориентированный комплекс программ, проведен численный анализ данных по дифференциальным сечениям упругого рассеяния заряженных пионов на ядрах ^{28}Si , ^{40}Ca , ^{58}Ni , ^{208}Pb при различных кинетических энергиях пионов $T = 130 \div 290$ МэВ в области энергии возбуждения пион-нуклонного $\Delta(1232)P_{33}$ -резонанса [1, 15, 16].

- Определены параметры пион-нуклонной амплитуды в трехпараметрическом микроскопическом оптическом потенциале, используемом для расчета дифференциальных сечений пион-ядерного рассеяния [2, 15, 17, 18]. Эти параметры характеризуют динамику изменения амплитуды рассеяния пионов на внутриядерных нуклонах (нуклонах «ядерной материи») в зависимости от энергии пионов [2].
- Использование АДЭ-минимизации позволило увеличить вероятность определения глобального минимума невязки между предсказаниями теоретической модели и экспериментальными данными, характеризующейся наличием нескольких локальных минимумов, по сравнению с методами SIMPLEX и MIGRAD²², так как невязка характеризуется наличием нескольких локальных минимумов. Таким образом, бóльшая

²⁰См. [78].

²¹См. [41].

²²MINUIT-реализация [46] метода симплексов (метод Нелдера — Мида [53]) и квазиньютоновского метода с формулой Дэвидона-Флетчера-Пауэлла [30, 31, 37, 38, 55] для обновления гессиана.

часть расчетов, где ранее требовались постоянный контроль и пере-
проверки, была автоматизирована.

С использованием АДЭ разработан комплекс программ и проведено
численное исследование модели малоуглового синхротронного рассеяния
на полидисперсной популяции везикул ДМФХ в 40% растворе сахарозы.

- Предложена модификация метода разделенных формфакторов, кото-
рая учитывает флуктуации параметров, описывающих структуру би-
слоя оболочек везикул. Модифицированный метод позволил описать
экспериментальные данные по рассеянию гамма-квантов в диапазоне
векторов рассеяния q от 0.06 до 4 nm^{-1} [3, 5].
- Определены параметры везикулярной системы, и на основе сопоставле-
ния с результатами анализа спектров нейтронного рассеяния сделаны
выводы о влиянии концентрации сахарозы на структуру везикулярной
системы ДМФХ [4, 19].
- Использование АДЭ-минимизации позволило эффективно распаралле-
лить вычисления, требующие глобальной минимизации многоэкстре-
альной функции с 9 частично-коррелированными переменными и зна-
чительных затрат компьютерного времени, более чем в 50 раз умень-
шив время расчетов, увеличив при этом вероятность определения гло-
бального минимума с 45% (с использованием минимизатора на основе
квазиныютоновского метода) до 90% [19].

В.7. Достоверность и апробация результатов

Эффективность предложенных алгоритмов глобальной минимизации
подтверждена вычислениями на основе тестовых наборов задач СЕС-2005
и ВВОВ-2012, а также сравнительными расчетами с применением дру-
гих часто используемых методов минимизации. Достоверность и обос-
нованность аналитических оценок, касающихся ограничений на управляю-
щие параметры АДЭ, подтверждена тестовыми расчетами. Достоверность
и обоснованность полученных численных результатов в рамках исследо-
вания многопараметрических моделей физических систем подтверждены
их согласием с экспериментальными данными и теоретическими оценками.
Предсказания разработанных моделей не противоречат опубликованным
результатам других авторов.

Результаты представлены автором на научных семинарах:

- Научный семинар РУДН «Математическое моделирование», 18 апреля 2012 г.
- Научный семинар ЛИТ ОИЯИ «Семинар по вычислительной физике», 17 апреля 2014 г.

на российских и международных научных конференциях:

- International Conference on Mathematical Modeling and Computational Physics (ММСР 2011), 2011, Stara Lesna, Slovakia.
- XIX международная конференция «Математика. Компьютер. Образование» (МКО-2012), 2012, ОИЯИ, Дубна, Россия.
- XVI конференция молодых учёных и специалистов ОМУС-2012, 2012, ОИЯИ, Дубна, Россия.
- Fifth Conference on Numerical Analysis and Applications (НАА-2012), 2012, University of Rousse, Lozenetz, Bulgaria.
- Современные проблемы прикладной математики и информатики (МРАМCS-2012), 2012, Дубна, Россия.
- International Conference on Mathematical Modeling and Computational Physics (ММСР 2013), 2013, JINR, Dubna, Russia.
- 39th Meeting of the PAC for Condensed Matter Physics, 2014, Dubna, Russia.
- XXI международная конференция «Математика. Компьютер. Образование» (МКО-2014), 2014, Дубна, Россия.
- Современные проблемы прикладной математики и информатики (МРАМCS-2014), 2014, Дубна, Россия.

В.8. Публикации и личный вклад автора в работу

Основные результаты диссертации опубликованы в 21 научной работе. Из них 6 работ [1–6] опубликованы в российских и 4 работы [7–10] — в иностранных рецензируемых изданиях.

Диссертант в сотрудничестве с коллегами и соавторами из ОИЯИ и других научных центров участвовал в математической постановке рассмотренных в работе задач, в проверке и улучшении соответствующих математических моделей, в разработке методов их численного исследования, в анализе и интерпретации получаемых численных результатов. В разработку представленных в диссертации вычислительных схем и комплексов

программ, в получение численных результатов, в анализ их точности и достоверности автором внесен определяющий вклад.

Работы [8, 15, 19] подготовлены лично автором. Результаты работ [6, 7, 9, 11–14] получены с определяющим вкладом автора. В работах [1, 16] все расчеты выполнены автором на основе разработанных им алгоритмов и программ. В работах [2, 17, 18] автором выполнены все расчеты параметров трехпараметрической модели микроскопического оптического потенциала. В работе [3] автором предложена модификация метода разделенных формфакторов для учета флуктуаций внутренних параметров структуры липидного бислоя везикул, систематизированы анализируемые конфигурации, выполнены все расчеты и проведено сравнение конфигураций на предмет наилучшего соответствия экспериментальным данным. В работах [4, 19–21] для описания экспериментальных данных синхротронного рассеяния использована предложенная автором модифицированная модель разделенных формфакторов, автором выполнены все расчеты, касающиеся анализа данных малоуглового рентгеновского рассеяния. В работах [19, 21] все расчеты выполнены автором.

В.9. Структура и объем работы

Диссертационная работа состоит из введения, трех глав, заключения и трех приложений.

В первой главе представлен метод асинхронной дифференциальной эволюции (АДЭ), метод АДЭ с рестартом и адаптивным кроссовером (АДЭ-АКМ) в сравнении с ДЭ. Сделаны эмпирические оценки скорости и вероятности сходимости метода АДЭ и проведено сравнение полученных результатов с классическим вариантом ДЭ. Представлены параллельные версии АДЭ. Приведены оценки ожидаемого ускорения параллельных вычислений и результаты эмпирических оценок на примере тестовых задач.

Во второй главе описана модель пион-ядерного рассеяния на основе трехпараметрического МОП и релятивистского волнового уравнения. Представлен комплекс программ, составленный для оценки параметров микроскопического оптического потенциала по экспериментальным дифференциальным сечениям упругого рассеяния пионов на ядрах. Приведены результаты расчетов. По полученным результатам сделан ряд выводов о характере рассеяния пионов на нуклонах, связанных в ядре. Представ-

лены результаты распараллеливания АДЭ-расчетов на основе технологии *Open Multi-Processing* (OMP).

Третья глава посвящена численному исследованию структуры полидисперсной везикулярной системы ДМФХ с применением метода АДЭ. Представлена модель РФФ, модифицированная для учета флуктуаций параметров структуры бислоя оболочки везикул. Представлены результаты расчетов для ряда конфигураций, описывающих структуру бислоя по данным МУРР. Проведено сравнение различных конфигураций для описания структуры оболочки везикул как с учетом, так и без учета флуктуаций параметров структуры бислоя. Получены оценки для параметров модифицированной и немодифицированной модели РФФ. Показано, что учет флуктуаций позволяет значительно улучшить согласие между предсказаниями теории и экспериментальными данными по МУРР на популяции везикул. Представлены оценки ускорения реальных параллельных вычислений, проведенных с использованием технологии *Message Passing Interface* (MPI) на параллельном кластере ЛИТ ОИЯИ.

Диссертационная работа содержит **3 приложения**. В первом приложении приведен вывод ограничений на управляющие параметры алгоритма. Во втором приложении дано описание тестовых функций из *IEEE Congress on Evolutionary Computation CEC-2005* [78]. В третьем приложении представлены результаты тестирования новых алгоритмов на наборе тестовых функций ВВОВ-2012.

Список цитируемой в диссертации литературы содержит 129 наименований. Работа изложена на 139 страницах, содержит 45 рисунков и 19 таблиц.

1. Метод асинхронной дифференциальной эволюции

Алгоритм дифференциальной эволюции (ДЭ) [67–77] относится к классу эволюционных алгоритмов, использующих итерационный процесс, похожий на рост и развитие биологических популяций [26, 35, 39, 47]. Метод ДЭ успешно применяется при поиске решений минимизационных задач, для которых целевая функция негладкая и/или многоэкстремальная.

Класс эволюционных алгоритмов, к которому принадлежит ДЭ, содержит два больших подмножества: подмножество *популяционных*, обновляющих всю популяцию на каждой итерации, и подмножество *квазистационарных* алгоритмов, одновременно обновляющих только часть потенциальных решений. Традиционно метод ДЭ относится и используется в форме популяционного алгоритма.

В данной работе операции мутации, кроссовера и отбора, характерные для ДЭ, впервые задействованы в предельной форме квазистационарного алгоритма. Сформулированный новый алгоритм, в котором на каждом шаге обновляется только один из членов популяции, назван асинхронной дифференциальной эволюцией (АДЭ) [7]. Также в работе представлен метод АДЭ с рестартом (АДЭ-Р) [9], АДЭ с кроссовером, использующим адаптивную корреляционную матрицу (АДЭ-АКМ) [10].

В этой главе новые методы АДЭ и АДЭ-АКМ сравниваются с классической ДЭ. Получены оценки для скорости и вероятности сходимости, ускорения при параллельных вычислениях. Проведены аналитические и эмпирические исследования областей сходимости и допустимых значений параметров алгоритмов. Получены ограничения, которым должны удовлетворять параметры алгоритмов для предотвращения преждевременной сходимости¹. Сделаны оценки ускорения расчетов на параллельных вычислительных системах.

1.1. Сравнение асинхронного и классического методов ДЭ

Постановка задачи. Будем рассматривать задачу (В.1) поиска глобального минимума $\vec{\theta}^* = \{\theta_j\}_{j=0, \dots, D-1}$ целевой функции

¹Преждевременная сходимость — это потеря разнообразия в популяции, которой оперирует эволюционный алгоритм. Эта потеря разнообразия приводит к искусственному ограничению на пространство поиска возможных решений, доступных алгоритму в процессе эволюции.

$f(\vec{\theta}): \Omega \subset \mathbb{R}^D \rightarrow \mathbb{R}$:

$$f(\vec{\theta}^*) \leq f(\vec{\theta}) \in \mathbb{R}, \forall \vec{\theta} \in \Omega \subset \mathbb{R}^D. \quad (\text{B.1})$$

Здесь Ω — вещественное пространство поиска решения размерности D , возможно, с нелинейными ограничениями на допустимые значения параметров.

1.1.1. Метод дифференциальной эволюции

Алгоритм дифференциальной эволюции — это стохастический алгоритм поиска минимума функций действительных переменных. ДЭ — эволюционный алгоритм, основанный на простой последовательности вычислительных итераций: инициализации популяции, а затем циклически повторяющихся операциях мутации, скрещивания (кроссовера) и отбора. Схема классического варианта ДЭ представлена на рис. 1.1, слева.

Алгоритм оперирует *популяцией* размера N_p . Каждый член популяции представлен вектором $\vec{\theta}$ в пространстве параметров. В качестве *начальной* популяции ($g=0$) классическая ДЭ создает N_p векторов в области поиска Ω , случайно распределенных в этой области² (инициализация популяции). Для каждого члена популяции, называемого *целевым* ($\vec{\theta}_i$), алгоритм проводит операции мутации и кроссовера с целью получить *пробный* вектор. Дифференциальная эволюция использует масштабируемую разность между случайно выбранными различными членами популяции для компоненты *мутантного* вектора:

$$\vec{v}_{i,g} = \vec{\theta}_{r,g} + F(\vec{\theta}_{p,g} - \vec{\theta}_{q,g}), \quad (1.1)$$

где $F > 0$ — масштабирующий фактор, $\vec{\theta}_r$ — базовый вектор. ДЭ также называют «алгоритмом с адаптивной мутацией». Если популяция широко распределена в пространстве поиска, то мутантный вектор, скорее всего, будет существенно отличаться от базового вектора, в противном случае, если популяция сконцентрирована в некоторой области, то мутации будут небольшими. Таким образом, алгоритм не использует каких-либо предположений

²Чаще всего — в соответствии с равномерным распределением.

Классический алгоритм ДЭ [70, 75]

```

//Инициализация популяции
//  $P_{\vec{\theta}} = \{\vec{\theta}_i\}_{i=0, N_p-1}$ ,  $\vec{\theta}_i = \{\theta_{i,j}\}_{j=0, D-1}$ 
do { // цикл по всем членам популяции
  for ( $i = 0$ ;  $i < N_p$ ;  $i = i + 1$ ) {
    //Мутация:
     $\vec{v}_i \leftarrow \vec{\theta}_r + F(\vec{\theta}_p - \vec{\theta}_q)$ ; //  $r \neq p \neq q$  — случ. индексы мутантный вектор
    //Кроссовер (рекомбинация):
    for ( $j = 0$ ;  $j < D$ ;  $j = j + 1$ )
       $u_{j,i} \leftarrow \begin{cases} v_{i,j}, & \text{если } \text{rand}(0, 1) \leq C_r; \\ \theta_{i,j}, & \text{если } \text{rand}(0, 1) > C_r. \end{cases}$  // пробный вектор

    //Отбор:
     $\vec{\theta}_i^{g+1} \leftarrow \begin{cases} \vec{u}_i & \text{если } (f(\vec{u}_i) < f(\vec{\theta}_i^g)); \\ \vec{\theta}_i^g & \text{если } (f(\vec{u}_i) \geq f(\vec{\theta}_i^g)). \end{cases}$ 
  }  $g \leftarrow g + 1$ ; // переход к новому поколению
} while (пока не выполнены критерии выхода);

```

Асинхронный алгоритм ДЭ (АДЭ) [7]

```

//Инициализация популяции
//  $P_{\vec{\theta}} = \{\vec{\theta}_i\}_{i=0, N_p-1}$ ,  $\vec{\theta}_i = \{\theta_{i,j}\}_{j=0, D-1}$ 
do { // выбор целевого вектора
   $i = \text{ChooseTargetVector}()$ ; // выбор цел. вект.  $\vec{\theta}_i$ 
  //Мутация:
   $\vec{v}_i \leftarrow \vec{\theta}_r + F(\vec{\theta}_p - \vec{\theta}_q)$ ; //  $r \neq p \neq q$  — случ. индексы мутантный вектор
  //Кроссовер (рекомбинация):
  for ( $j = 0$ ;  $j < D$ ;  $j = j + 1$ )
     $u_{j,i} \leftarrow \begin{cases} v_{i,j}, & \text{если } \text{rand}(0, 1) \leq C_r; \\ \theta_{i,j}, & \text{если } \text{rand}(0, 1) > C_r. \end{cases}$  // пробный вектор

  //Отбор:
   $\vec{\theta}_i \leftarrow \begin{cases} \vec{u}_i & \text{если } (f(\vec{u}_i) < f(\vec{\theta}_i)); \\ \vec{\theta}_i & \text{если } (f(\vec{u}_i) \geq f(\vec{\theta}_i)). \end{cases}$ 
} while (пока не выполнены критерии выхода);

```

Рис. 1.1. Сравнение классического (ДЭ, слева) и асинхронного (АДЭ, справа) алгоритмов дифференциальной эволюции

о функции плотности вероятности для распределения координат мутантного вектора.

Биномиальное скрещивание (рекомбинация, кроссовер) заключается в том, что каждая из координат формируемого *пробного* вектора с некоторой вероятностью C_r берется из мутантного вектора, а в противном случае — из целевого:

$$u_{i,j} = \begin{cases} v_{i,j}, & \text{если } \text{rand}(0, 1) \leq C_r, \text{ или } j = j_{\text{rand}}; \\ \theta_{i,j}, & \text{иначе.} \end{cases} \quad (1.2)$$

Для того чтобы гарантировать отличие пробного вектора от целевого, по крайней мере одна случайно выбранная координата j_{rand} берется из мутантного вектора. Если значение целевой функции для пробного вектора $\vec{u}_{i,g}$ меньше, чем для целевого $\vec{\theta}_{i,g}$, то пробный вектор заменяет своего «родителя» (целевой вектор) в следующем поколении. В противном случае целевой вектор остается в популяции на следующую итерацию.

С момента изобретения в 1995 г. [73, 75] метод дифференциальной эволюции стал очень популярным алгоритмом поиска глобального минимума без использования производных. Большинство последних достижений по развитию и усовершенствованию этого метода представлено в книгах [67, 70] и обзоре [65]. Метод ДЭ применяется для решения различных оптимизационных задач, обзор приложений можно найти в [66].

Параллельные реализации ДЭ базируются в основном на том факте, что значение целевой функции для каждого члена пробной популяции может вычисляться независимо от остальных. Единственная фаза в алгоритме, которая требует взаимодействия с остальными членами популяции, — это переход между поколениями. Первые попытки распределить вычисления между разными компьютерами, соединенными локальной компьютерной сетью, были предприняты Лампиненом [66]. В его методе вся популяция хранится в мастер-процессе, который выбирает членов популяции для мутации (стыковки) и отправляет пробные векторы для вычисления целевой функции slave-процессорам. Схема распараллеливания Лампинена способна также преодолевать недостатки, связанные с неоднородностью в скорости подчиненных процессов. Однако с возрастанием количества используемых вычислительных узлов ускорение параллельных вычислений в этой и других схемах, распараллеливающих классический вариант ДЭ,

начинает значительно отставать от числа задействованных вычислительных узлов.

1.1.2. Алгоритм АДЭ

Метод АДЭ, так же как и метод ДЭ, оперирует популяцией P_θ , состоящей из N_p векторов $\vec{\theta}$ в пространстве параметров Ω :

$$P_\theta = \{ \vec{\theta}_i \}, \vec{\theta}_i \in \Omega \subset \mathbb{R}^D, i = 0, 1, \dots, N_p - 1. \quad (1.3)$$

Начальная популяция инициализируется N_p векторами, случайно выбранными из Ω .

В классическом варианте ДЭ [70] (рис. 1.1, слева) выполняется цикл, на каждом шаге которого ко всем членам текущей популяции применяются операции мутации, рекомбинации (кроссовера) и отбора для формирования следующего поколения. Смена поколения в ДЭ происходит синхронно для всех членов популяции.

В новом алгоритме АДЭ [7] операторы мутации, кроссовера и отбора применяются для членов популяции без синхронизации по поколениям (рис. 1.1, справа).

После инициализации в цикле алгоритма выбирается *целевой* вектор $\vec{\theta}_i$. Для выбранного целевого вектора операции мутации, кроссовера и отбора проводятся так же, как и в классическом варианте ДЭ. Если значение целевой функции для пробного вектора оказывается лучше с точки зрения заданного критерия (меньше при минимизации), чем значение для целевого вектора, он сразу заменяет родителя в популяции.

Таким образом, в алгоритме АДЭ нет обязательного для классической ДЭ перебора всех членов популяции (целевых векторов). Целевые векторы, для которых будут осуществляться операции мутации, кроссовера и отбора, выбираются из популяции независимо, по одному, что открывает дополнительные возможности для распараллеливания и ускорения вычислений. Сравнение классического и асинхронного алгоритмов приведено на рисунках 1.1 и 1.2.

Описанная последовательность этапов выбора целевого вектора, мутации, рекомбинации и отбора повторяется в цикле до тех пор, пока не будет выполнен один из критериев остановки (достижение минимума с заданной точностью, диагностика стагнации алгоритма, превышение максимально допустимого количества итераций).

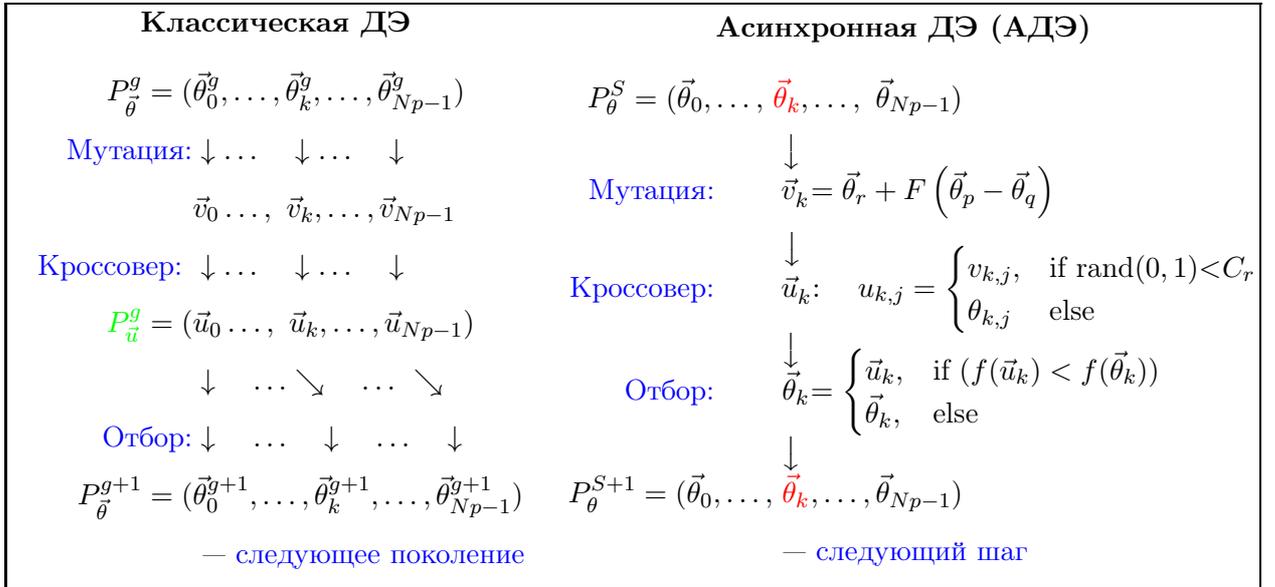


Рис. 1.2. Сравнение классического (ДЭ) и асинхронного (АДЭ) алгоритмов дифференциальной эволюции

Способ выбора очередного целевого вектора является специфической чертой АДЭ. Другими факторами, определяющими конкретную стратегию поиска глобального минимума в рамках АДЭ, являются способ выбора базового вектора, число разностных векторов и тип кроссовера. Для идентификации различных вариантов АДЭ введены обозначения DE/a/b/n/c, расширяющие принятую в [70] символику³. Здесь *a* соответствует способу выбора целевого вектора: например, это может быть случайный («rand») или худший, то есть член популяции, имеющий значение целевой функции, максимальное по популяции («worst»). Символ *b* отвечает за способ выбора базового вектора: это может быть как случайный («rand») или лучший («best») член популяции, так и другие стратегии. Числу разностных векторов, которые добавляются к базовому при формировании мутантного вектора, соответствует *n*. Под *c* закодирован тип кроссовера. Обычно используется биномиальное (равномерное) скрещивание («bin»).

Предложенный алгоритм прост в применении, имеет малое количество контрольных параметров (N_p , F и C_r), применим для решения задач большой ($D = 10 \dots 100 \dots$) размерности. В АДЭ не используются производные, поэтому с его помощью можно решать недифференцируемые, многоэкстремальные задачи.

Очевидное преимущество предложенного алгоритма заключается

³В [70] предложена и обычно используется DE/b/n/c-классификация стратегий, где *b* — способ выбора базового вектора, *n* — число разностных векторов, которые добавляются к базовому при формировании мутантного вектора, *c* — тип кроссовера.

в том, что улучшенный новый вектор принимает участие в дальнейшей эволюции сразу же, без задержки по времени. Так как цикл по всем членам популяции отсутствует, концепция *поколений*, ключевая в классической ДЭ, в АДЭ не применяется.

Отметим, что асинхронизация метода достигнута без введения в алгоритм новых параметров.

1.1.3. Сравнение скорости и вероятности сходимости АДЭ и ДЭ

Показатели скорости и вероятности сходимости алгоритма АДЭ [7]⁴ протестированы на наборе тестовых функций СЕС-2005 [78], сформулированных для специальной сессии, посвященной минимизации функций действительных переменных. Здесь приведены четыре функции, которые представляют разные классы минимизационных задач: f_1 — гиперсфера (униmodalная, разделяемые переменные), f_6 — функция Розенброка (немного локальных минимумов, неразделяемые переменные), f_9 — функция Растригина (многоэкстремальная, разделяемые переменные) и f_{11} — повернутая функция Вейерштрасса (многоэкстремальная, неразделяемые переменные). Нахождение решения для функции Розенброка и повернутой функции Вейерштрасса является более сложной задачей для минимизации в силу того, что параметры задачи неразделяемые. Результаты для всех вышеупомянутых функций представлены для размерности $D = 10$.

Задача считалась решенной, если алгоритм достигал глобального минимума с заранее определенной точностью ε : $f(\vec{\theta}) - f(\vec{\theta}^*) < \varepsilon$. Для гиперсферы требуемая точность была установлена на уровне $\varepsilon = 10^{-6}$. Для остальных функций $\varepsilon = 10^{-2}$. Ограничение на максимальное количество вычислений функции в СЕС-2005 установлено на уровне $N_{\text{FE}}^{\text{max}} = 10^5$.

Одно измерение соответствует усреднению результатов n_t запусков алгоритма. Для каждого измерения оценка вероятности успеха P_{succ} определяется как количество попыток, сошедшихся к глобальному минимуму (n_{succ}), деленное на полное количество запусков алгоритма:

$$P_{\text{succ}} = \frac{n_{\text{succ}}}{n_t}.$$

Среднее количество вычислений функции $\langle N_{\text{FE}} \rangle$ рассчитано для попыток, сошедшихся к глобальному минимуму.

⁴Результаты, представленные в этом разделе, опубликованы в [7].

В таблице 1.1 представлено сравнение результативности разных стратегий АДЭ с соответствующими стратегиями ДЭ при нахождении минимума 10-мерной функции Розенброка. Для всех стратегий использованы масштабирующий фактор $F = 0,9$ и кроссовер $C_r = 0,9$. Размер популяции N_p выбирался различным. Заметим, что для решения выбранной задачи с вероятностью, близкой к единице, семейство стратегий DE/rand/1/bin требует меньшего размера популяции, чем семейство, использующее DE/best/1/bin-стратегию. Для фиксированного N_p семейство DE/best/1/bin-стратегий характеризуется более высокой скоростью сходимости. В случае относительно небольшого размера популяции N_p скорости сходимости ДЭ и АДЭ сопоставимы, однако вероятности сходимости далеки от единицы. С увеличением размера популяции увеличивается вероятность сходимости. При достаточно большом размере популяции вероятность сходимости АДЭ близка к единице. Скорость и вероятность сходимости для асинхронной стратегии DE/rand/rand/1/bin близки к тем же показателям классического варианта DE/rand/1/bin: преимущества, даваемые при немедленном включении лучшего вектора-кандидата в эволюцию скомпенсированы ускоренной потерей разнообразия в популяции. Не удивительно, что асинхронная стратегия DE/rand/best/1/bin показывает лучшие результаты, чем DE/best/1/bin: последняя генерирует пробные векторы вокруг текущего лучшего вектора в течение жизни всего поколения, что ведет к потере популяционного разнообразия. Асинхронная стратегия DE/worst/best/1/bin — это хороший вариант относительно быстрой и надежной стратегии для нахождения локального минимума.

В таблице 1.2 проведено полученное в работе [7] сравнение результативности стратегии DE/worst/best/1/bin с классическими вариантами DE/rand/1/bin и DE/best/1/bin. Кроссовер C_r для всех стратегий выбран равным 0,1 для разделяемых функций и 0,9 для неразделяемых задач. Расчеты были проведены для различных размеров популяции: $N_p = 10, 15, 20, 25, 30, 50, 70, 100, 150$. Представлены результаты, потребовавшие наименьших затрат компьютерного времени (наименьшего среднего количества вычислений функции $\langle N_{FE} \rangle$), с наибольшей вероятностью сходимости. Известно, что DE/rand/1/bin обладает лучшей поисковой способностью, в то время как DE/best/1/bin быстрее сходится к минимуму. Вероятность и скорость сходимости новой асинхронной стратегии DE/worst/best/1/bin конкурентны по сравнению с классическими страте-

Таблица 1.1. Сравнение скорости и вероятности сходимости для классических и асинхронных стратегий ДЭ на примере 10-мерной функции Розенброка. $F = 0,9$, $C_r = 0,9$, $N_{FE}^{max} = 10^5$, $n_t = 25$ [7]

N_p	ДЭ		АДЭ			
	rand/1/bin		rand/rand/1/bin		worst/rand/1/bin	
	$\langle N_{FE} \rangle$	P_{succ}	$\langle N_{FE} \rangle$	P_{succ}	$\langle N_{FE} \rangle$	P_{succ}
15	1,94e+04	0,80	2,06e+04	0,80	4,19e+04	0,76
40	6,30e+04	0,96	6,52e+04	1,00	4,36e+04	1,00
N_p	best/1/bin		rand/best/1/bin		worst/best/1/bin	
	$\langle N_{FE} \rangle$	P_{succ}	$\langle N_{FE} \rangle$	P_{succ}	$\langle N_{FE} \rangle$	P_{succ}
40	2,26e+04	0,76	2,07e+04	0,88	1,53e+04	0,68
100	5,29e+04	0,88	5,26e+04	1,00	3,76e+04	1,00

Таблица 1.2. Сравнение скорости и вероятности сходимости для ДЭ и АДЭ для четырех представленных функций из СЕС-2005. $N_{FE}^{max} = 10^5$, $n_t = 25$ [7]

	ДЭ						АДЭ		
	rand/1/bin			best/1/bin			worst/best/1/bin		
	N_p	$\langle N_{FE} \rangle$	P_{succ}	N_p	$\langle N_{FE} \rangle$	P_{succ}	N_p	$\langle N_{FE} \rangle$	P_{succ}
f_1	10	3700	1,0	10	3212	1,0	10	2962	1,0
f_6	15	2,43e+4	0,80	15	1,60e+4	0,80	25	1,26e+4	0,92
f_6	30	4,36e+4	0,96	150	8,71e+4	0,92	70	2,67e+4	1,0
f_9	15	3973	1,0	30	6270	1,0	25	5403	1,0
f_9				15	4627	0,72	10	4607	0,48
f_{11}	20	5,78e+4	0,68	40	8,52e+4	0,40	50	4,65e+4	0,72
f_{11}	25	9,25e+4	0,72	70	1,45e+5	0,52	70	7,22e+4	0,76

гиями.

В табл. 1.3 приведены полученные в работе [9] скорости и вероятности сходимости для асинхронных стратегий DE/rand/rand/1/bin и DE/worst/best/1/bin. И если в СЕС-2005 максимально допустимое количество вычислений целевой функции фиксировалось на уровне $N_{FE}^{max} = 10^5$, в этих расчетах оно было увеличено до $N_{FE}^{max} = 10^7$. Здесь каждое испытание — это $n_t = 100$ запусков алгоритма. Представлены средние количества вычислений функции $\langle N_{FE} \rangle$ и медианы N_{FE} для сошедшихся к минимуму попыток. Расчеты проводились для размеров популяции $N_p = 10, 15, 20, 30, 50, 70, 100, 150, 200, 600$. В таблице зафиксированы результаты с наибольшей вероятностью сходимости, для которых потребовались наименьшие затраты компьютерного времени (наименьшее среднее количество вычислений функции $\langle N_{FE} \rangle$).

Таблица 1.3. Сравнение скорости и вероятности сходимости стратегий АДЭ для четырех представленных функций из CEC-2005, при увеличенном максимально допустимом количестве вычислений целевой функции. $N_{FE}^{max} = 10^7$, $n_t = 100$ [9]

	DE/rand/rand/1/bin				DE/worst/best/1/bin			
	N_p	$\langle N_{FE} \rangle$	Медиана N_{FE}	P_{succ}	N_p	$\langle N_{FE} \rangle$	Медиана N_{FE}	P_{succ}
f_1	10	3641	3633	1,00	10	2986	2977	0,98
					15	4344	4333	1,00
f_6	20	2,38e+04	2,32e+04	0,92	50	1,92e+04	1,84e+04	0,85
	50	8,30e+04	8,20e+04	0,98	150	5,65e+04	5,59e+04	0,85
	70	1,28e+05	1,25e+05	1,00	600	2,12e+05	2,11e+05	0,91
f_9	15	3963	4012	0,93	20	4317	4289	0,92
	30	7976	7988	1,0	50	10138	10108	1,0
f_{11}	20	0,406e+05	0,381e+05	0,59	50	0,347e+05	0,333e+05	0,55
	50	2,939e+05	2,913e+05	0,83	100	0,870e+05	0,842e+05	0,73
	100	1,168e+06	1,122e+05	0,90	200	1,952e+05	1,887e+05	0,84

Сравнение ускорения параллельных вычислений АДЭ и ДЭ на примере тестовых функций приведено в подразделе 1.5.3.

Выводы по разделу 1.1

Итак, мы описали метод асинхронной дифференциальной эволюции, предназначенный для решения задач нахождения глобального минимума. В рамках нового метода сформулированы новые асинхронные стратегии с хорошей поисковой способностью и быстрой сходимостью. Новые стратегии показывают конкурентоспособные результаты по сравнению с традиционно используемыми вариантами классического алгоритма ДЭ.

1.2. Оптимальные параметры метода АДЭ

Вероятность, с которой популяционный алгоритм достигнет глобального минимума, зависит от его способности сохранять в процессе эволюции разнообразие в популяции, таким образом, поддерживая потенциальную возможность достигнуть любой точки из пространства решений. Это необходимо, чтобы предотвратить преждевременную сходимость алгоритма. Операция отбора обычно приводит к уменьшению разброса векторов в популяции, поэтому ожидаемый разброс в популяции должен возрастать в результате операций мутации и кроссовера. Мерилом популяционного разнообразия в нашем случае может служить

статистическая вариация $\text{Var}(P_\theta)_j = \overline{\theta_j^2} - \overline{\theta_j}^2$ (ур. 1.4), рассчитанная по всем членам популяции. В этом разделе анализируется ожидаемая вариация популяции после операций мутации и кроссовера. Исходя из полученных выражений для ожидаемой вариации популяции после операций мутации и кроссовера, получим соотношения между контрольными параметрами $\{N_p, F, C_r\}$ алгоритма и выражения для «фактора расширения» популяции. На контрольные параметры выведены ограничения, необходимые для предотвращения преждевременной сходимости алгоритма. Выполнено сравнение результатов для классических стратегий (DE/rand/1/bin и DE/best/1/bin) с результатами для асинхронных стратегий (DE/rand/rand/1/bin и DE/rand/best/1/bin). Проведено сравнение полученных аналитических ограничений на параметры алгоритма с эмпирическими оценками областей сходимости для тестовых функций.

Методы ДЭ и АДЭ [65, 70] содержат только три контрольных параметра: размер популяции N_p , масштабирующий фактор F и вероятность кроссовера C_r . Во многих работах [65, 66, 70] эмпирически анализируются наиболее подходящие сочетания параметров алгоритма. Захариэ (Zaharie) в своей работе [79] показала, что существует нижний предел для F , необходимый для того, чтобы избежать преждевременной сходимости алгоритма для стратегии DE/rand/1/bin. В этом разделе представлен подобный аналитический анализ для других классических стратегий ДЭ и для ряда асинхронных стратегий АДЭ.

Метод ДЭ оперирует популяцией фиксированного размера N_p :

$$P_\theta = \left\{ \vec{\theta}_0, \vec{\theta}_1, \dots, \vec{\theta}_l, \dots, \vec{\theta}_{N_p-1} \right\}, \quad \vec{\theta}_l \in \Omega \subset \mathbb{R}^D, \quad l = \overline{0, N_p - 1}.$$

Популяция инициализируется случайно выбранными из Ω элементами. После инициализации начинается итерационный процесс. В ДЭ на каждой итерации (называемой поколением) появляется новая популяция (рис. 1.1, слева). В АДЭ (рис. 1.1, справа) операции мутации, кроссовера и отбора применяются к каждому члену популяции независимо. Асинхронизация влияет на ход итерационного процесса, на популяционное разнообразие и на возможную преждевременную сходимость.

1.2.1. Аналитические оценки границ области сходимости АДЭ

Проанализируем влияние операций мутации и кроссовера на дисперсию популяции (для каждой компоненты). Подобный анализ был прове-

ден для класса эволюционных стратегий в [64] в контексте анализа их поисковой способности, а затем в [79] для ДЭ для стратегии rand/1/bin. Мы же проведем анализ для стратегии best/1/bin ДЭ и для стратегий rand/rand/1/bin и rand/best/1/bin АДЭ, а также ряда существенно асинхронных стратегий АДЭ.

Изменение членов популяции от шага к шагу происходит стохастическим образом. Мы вычислим *ожидаемую* дисперсию: $E \left(\text{Var} \left(\vec{\theta} \right) \right)$, где E — это усреднение по популяции, а дисперсию популяции представляет вектор размерности D :

$$\text{Var} (P_{\theta}) = \left\{ \text{Var} (P_{\theta})_j \right\} = \left\{ \text{Var} \left(\vec{\theta} \right)_0, \dots, \text{Var} \left(\vec{\theta} \right)_{D-1} \right\}, \quad (1.4)$$

$$\text{Var} \left(\vec{\theta} \right)_j = \text{Var} (\theta_j) = \overline{\theta_j^2} - \overline{\theta_j}^2, \quad \overline{\theta_j^2} = \frac{1}{N_p} \sum_{l=0}^{N_p-1} \theta_{l,j}^2, \quad \overline{\theta_j} = \frac{1}{N_p} \sum_{l=0}^{N_p-1} \theta_{l,j}.$$

В рассматриваемых вариантах алгоритма операции мутации и кроссовера действуют независимо на каждую из компонент. Поэтому в доказательствах можно ограничиться покомпонентным анализом. Для краткости, когда нет принципиальной разницы между вектором и его отдельной компонентой, в записи будем опускать знак вектора и индекс компоненты: $\text{Var} \left(\vec{\theta} \right)_j \equiv \text{Var} (\theta)$.

Следующие *положения*, обычно применяемые в алгоритме ДЭ, не учитываются в доказательстве, представленном в работе [79], где ожидаемая вариация популяции после операций мутации и кроссовера оценена для стратегии rand/1/bin:

П1: $r \neq l, p \neq l, q \neq l$ — векторы, из которых строится мутантный вектор, не совпадают с целевым.

П2: Хотя бы одна компонента $v_l - v_{l,m}$ мутирует.

В диссертации доказательства для стратегий DE/rand/1/bin и DE/best/1/bin проведены с учетом П1. Положение П2 частично учитывается следующим образом: $(1 - C_r)^D$ — вероятность того, что ни одна из координат не была возмущена, $1/D$ — вероятность возмущения данной координаты, D — размерность задачи. Тогда эффективный кроссовер (полная вероятность возмущения данной координаты, если задана величина C_r) больше номинального⁵ на величину $\frac{(1-C_r)^D}{D}$. Считаем, что

⁵Точнее, эффективный кроссовер больше номинального на величину $\frac{(1-C_r)^D}{D} e^{\left(\frac{D-1}{D}\right)^D}$.

в представленных ниже теоремах речь идет именно об эффективном значении кроссовера ($C_r^{\text{eff}} = C_r + \frac{(1-C_r)^D}{D}$).

Так же, как и в работе [79], рассмотрим вариант, когда масштабирующий фактор для каждой координаты j и члена популяции l может быть представлен в виде

$$\widetilde{F}_l^j = F^j \xi_l^j > 0, \quad \xi_l^j \sim N(0, 1), \quad (1.5)$$

где случайные величины ξ_l^j независимы и одинаково распределены⁶. В классической ДЭ масштабирующий фактор \widetilde{F}_l^j — константа, он одинаков для всех координат j и членов популяции l :

$$\widetilde{F}_l^j = F > 0, \quad l = \overline{0, N_p - 1}, j = \overline{0, D - 1}.$$

Основной теоретический результат этого раздела заключается в выводе следующих теорем:

Теорема 1.2.1 (ДЭ: DE/rand/1/bin). Пусть $P_\theta = \{\vec{\theta}_0, \dots, \vec{\theta}_{N_p-1}\}$ — текущая популяция; $P_v = \{\vec{v}_0, \dots, \vec{v}_{N_p-1}\}$ — популяция, состоящая из мутантных векторов, сформированных по классической стратегии ДЭ (DE/rand/1/bin); $P_u = \{\vec{u}_0, \dots, \vec{u}_{N_p-1}\}$ — популяция, состоящая из пробных векторов, сформированных по этой стратегии, полученная путем скрещивания членов популяций P_θ и P_v . Тогда, как в случае постоянного масштабирующего фактора F^7 , так и в случае масштабирующего фактора \widetilde{F}_l^j , определяемого соотношением (1.5), для стратегии DE/rand/1/bin классической ДЭ выполнены следующие утверждения⁸:

а) после мутации ожидаемая j -я компонента вариации популяции равна

$$\begin{aligned} E(\text{Var}(P_v)_j) &= (1 + \zeta^j) \text{Var}(P_\theta)_j = \\ &= \left(1 + 2F^{j^2} - \frac{1}{N_p} + \frac{1}{N_p(N_p - 1)^2} \right) \text{Var}(P_\theta)_j; \end{aligned} \quad (1.6)$$

б) после мутации и кроссовера ожидаемая j -я компонента вариации по-

⁶В работе [79] рассмотрен случай с F_l^i , заданным (1.5) только для стратегии DE/rand/1/bin.

⁷В случае постоянного масштабирующего фактора F здесь и далее в формулах вместо F^j стоит F , индекс j в них можно опустить везде и рассматривать $\text{Var}(P)$ как векторную величину.

⁸Если не учитывать П1: $r \neq l, p \neq l, q \neq l$, то в [79]: $E(\text{Var}(P_v))_j = \left(1 + 2F^{j^2} - \frac{1}{N_p} \right) \text{Var}(P_\theta)_j$; $E(\text{Var}(P_u))_j = \left(1 + 2F^{j^2} C_r - \frac{C_r}{N_p} (2 - C_r) \right) \text{Var}(P_\theta)_j$.

пуляции равна

$$\begin{aligned} E(\text{Var}(P_u)_j) &= (1 + \alpha^j) \text{Var}(P_\theta)_j = \\ &= \left(1 + 2F^{j^2} C_r - \frac{C_r}{N_p - 1} \left(2 - \frac{N_p}{N_p - 1} C_r \right) \right) \text{Var}(P_\theta)_j. \end{aligned} \quad (1.7)$$

Теорема 1.2.2 (ДЭ: DE/best/1/bin). Пусть $P_\theta = \{\vec{\theta}_0, \dots, \vec{\theta}_{N_p-1}\}$ – текущая популяция; $P_v = \{\vec{v}_0, \dots, \vec{v}_{N_p-1}\}$ – популяция, состоящая из мутантных векторов для ДЭ (DE/best/1/bin); $P_u = \{\vec{u}_0, \dots, \vec{u}_{N_p-1}\}$ – популяция, состоящая из пробных векторов для ДЭ, полученная путем скрещивания членов популяций P_θ и P_v . Как в случае постоянного масштабирующего фактора F , так и для F_l^j , заданного соотношением (1.5), для стратегии DE/best/1/bin выполнены следующие утверждения:

а) после мутации ожидаемая дисперсия компонент популяции равна

$$E(\text{Var}(P_v))_j = (1 + \zeta^j) \text{Var}(P_\theta)_j = 2F^{j^2} \text{Var}(P_\theta)_j; \quad (1.8)$$

б) после мутации и кроссовера ожидаемая дисперсия компонент популяции равна

$$\begin{aligned} E(\text{Var}(P_u))_j &= (1 + \alpha^j) \text{Var}(P_\theta)_j = \\ &= \left(1 + 2F^{j^2} C_r - C_r - \frac{C_r - C_r^2}{N_p} \right) \text{Var}(P_\theta)_j + \frac{(N_p - 1)(C_r - C_r^2)(\theta_{b,j} - \bar{\theta}_j)^2}{N_p}, \end{aligned} \quad (1.9)$$

где $\theta_{b,j}$ – j -я координата лучшего вектора θ_b в популяции.

Теорема 1.2.3 (АДЭ: асинхронные стратегии).

Пусть $P_\theta = \{\vec{\theta}_0, \dots, \vec{\theta}_{N_p-1}\}$ – текущая популяция;

$P_{v_k} = \{\vec{\theta}_0, \dots, \vec{\theta}_{k-1}, \vec{v}_k, \vec{\theta}_{k+1}, \dots, \vec{\theta}_{N_p-1}\}$ – популяция после замены одного из векторов (k -го) мутантным вектором для одной из стратегий АДЭ.

$P_{u_k} = \{\vec{\theta}_0, \dots, \vec{\theta}_{k-1}, \vec{u}_k, \vec{\theta}_{k+1}, \dots, \vec{\theta}_{N_p-1}\}$ – популяция, после замены одного из векторов пробным вектором, полученная путем скрещивания членов популяций P_θ и P_v . Как в случае постоянного масштабирующего фактора F , так и для \tilde{F} , заданного соотношением (1.5), выполнены следующие утверждения:

а) после мутации ожидаемая дисперсия компонент популяции равна⁹

⁹Где $\theta_{b,j}$ и $\theta_{w,j}$ – соответственно j -е координаты лучшего и худшего векторов в популяции.

$$E(\text{Var}(P_{\vec{v}_k})_j) = (1 + \tilde{\zeta}^j) \text{Var}(P_\theta)_j =$$

$$\begin{aligned} \text{rand/rand/1/bin:} & \left(1 + \frac{2F^{j^2}}{N_p} - \frac{2}{N_p^2}\right) \text{Var}(P_\theta)_j; \\ \text{rand/best/1/bin:} & \left(1 + \frac{2F^{j^2}}{N_p} - \frac{N_p+1}{N_p^2}\right) \text{Var}(P_\theta)_j + \frac{N_p-1}{N_p^2} (\theta_{b,j} - \bar{\theta}_j)^2; \\ \text{worst/rand/1/bin:} & \left(1 + \frac{2F^{j^2}}{N_p} + \frac{N_p-1}{N_p^2}\right) \text{Var}(P_\theta)_j - \frac{N_p+1}{N_p^2} (\theta_{w,j} - \bar{\theta}_j)^2; \\ \text{worst/best/1/bin:} & \left(1 + \frac{2F^{j^2}}{N_p}\right) \text{Var}(P_\theta)_j - \frac{N_p+1}{N_p^2} (\theta_{w,j} - \bar{\theta}_j)^2 + \\ & + \frac{N_p-1}{N_p^2} (\theta_{b,j} - \bar{\theta}_j)^2 + \frac{2}{N_p^2} (\theta_{b,j} - \bar{\theta}_j) (\theta_{w,j} - \bar{\theta}_j); \end{aligned}$$

б) после мутации и кроссовера ожидаемая дисперсия компонент популяции равна

$$E(\text{Var}(P_{\vec{u}_k})_j) = (1 + \tilde{\alpha}^j) \text{Var}(P_\theta)_j =$$

$$\begin{aligned} \text{rand/rand/1/bin:} & \left(1 + \frac{2F^{j^2}C_r}{N_p} - \frac{2C_r}{N_p^2}\right) \text{Var}(P_\theta)_j; \\ \text{rand/best/1/bin:} & \left(1 + \frac{2F^{j^2}C_r}{N_p} - \frac{N_p+1}{N_p^2}C_r\right) \text{Var}(P_\theta)_j + \frac{N_p-1}{N_p^2}C_r (\theta_{b,j} - \bar{\theta}_j)^2; \\ \text{worst/rand/1/bin:} & \left(1 + \frac{2F^{j^2}C_r}{N_p} + \frac{N_p-1}{N_p^2}C_r\right) \text{Var}(P_\theta)_j - \frac{N_p+1}{N_p^2}C_r (\theta_{w,j} - \bar{\theta}_j)^2; \\ \text{worst/best/1/bin:} & \left(1 + \frac{2F^{j^2}C_r}{N_p}\right) \text{Var}(P_\theta)_j - \frac{N_p+1}{N_p^2}C_r (\theta_{w,j} - \bar{\theta}_j)^2 + \\ & + \frac{N_p-1}{N_p^2}C_r (\theta_{b,j} - \bar{\theta}_j)^2 + \frac{2C_r}{N_p^2} (\theta_{b,j} - \bar{\theta}_j) (\theta_{w,j} - \bar{\theta}_j). \end{aligned}$$

Доказательства теорем приведены в приложении П. II.

Таким образом, в теоремах 1.2.1, 1.2.2 и 1.2.3 получены ожидаемые значения компонент дисперсии популяции после мутации и кроссовера: $E(\text{Var}(P_{\vec{u}})_j)$ для классической ДЭ и $E(\text{Var}(P_{\vec{u}_k})_j)$ для асинхронных вариантов ДЭ. Ожидаемые значения выражены в терминах N_p , F , C_r , первоначальной дисперсии популяции $\text{Var}(P_\theta)_j$ и координат определенного базового ($\vec{\theta}_b$ — лучший член популяции) и/или определенного целевого ($\vec{\theta}_w$ — худший член популяции) вектора в тех случаях, когда они явно используются в стратегии.

Так как отбор обычно приводит к уменьшению разнообразия (дисперсии) популяции, то для того чтобы избежать преждевременной сходимости алгоритма, мутация и кроссовер должны приводить к увеличению дисперсии:

$$E(\text{Var}(P_u)_j) = (1 + \alpha^j) \text{Var}(P_\theta)_j > \text{Var}(P_\theta)_j \quad \text{— для ДЭ,} \quad (1.10)$$

$$E(\text{Var}(P_{u_k})_j) = (1 + \tilde{\alpha}^j) \text{Var}(P_\theta)_j > \text{Var}(P_\theta)_j \quad \text{— для АДЭ.} \quad (1.11)$$

Параметры N_p , F , C_r должны быть выбраны таким образом, чтобы коэффициенты расширения популяции были больше единицы: $(1 + \alpha^j) > 1$ и $(1 + \tilde{\alpha}^j) > 1$.

Введем обозначения:

$$\beta_b^j \equiv \frac{\theta_{bj} - \bar{\theta}_j}{\sqrt{\text{Var}(P_{\theta_j})}}, \quad \beta_w^j \equiv \frac{\theta_{wj} - \bar{\theta}_j}{\sqrt{\text{Var}(P_{\theta_j})}}, \quad (1.12)$$

и перепишем результаты теорем 1.2.1, 1.2.2 и 1.2.3 в следующем виде:

Классические стратегии ДЭ: $E(\text{Var}(P_{\bar{u}})_j) = (1 + \alpha^j)\text{Var}(P_{\theta})_j =$	
rand/1/bin [79]:	$\left(1 + 2F^{j^2}C_r - \frac{2C_r}{N_p-1} + \frac{N_p C_r^2}{(N_p-1)^2}\right) \text{Var}(P_{\theta})_j;$
best/1/bin:	$\left(1 + 2F^{j^2}C_r - C_r - \frac{C_r - C_r^2}{N_p} + \frac{(N_p-1)(C_r - C_r^2)}{N_p} \beta_b^{j^2}\right) \text{Var}(P_{\theta})_j.$
Стратегии асинхронной ДЭ: $E(\text{Var}(P_{\bar{u}_k})_j) = (1 + \tilde{\alpha}^j)\text{Var}(P_{\theta})_j =$	
rand/rand/1/bin:	$\left(1 + \frac{2F^{j^2}C_r}{N_p} - \frac{2C_r}{N_p^2}\right) \text{Var}(P_{\theta})_j;$
rand/best/1/bin:	$\left(1 + \frac{2F^{j^2}C_r}{N_p} - \frac{N_p+1}{N_p^2}C_r + \frac{N_p-1}{N_p^2}C_r\beta_b^{j^2}\right) \text{Var}(P_{\theta})_j;$
worst/rand/1/bin:	$\left(1 + \frac{2F^{j^2}C_r}{N_p} + \frac{N_p-1}{N_p^2}C_r - \frac{N_p+1}{N_p^2}C_r\beta_w^{j^2}\right) \text{Var}(P_{\theta})_j;$
worst/best/1/bin:	$\left(1 + \frac{2F^{j^2}C_r}{N_p} - \frac{N_p+1}{N_p^2}C_r\beta_w^{j^2} + \frac{N_p-1}{N_p^2}C_r\beta_b^{j^2} + \frac{2C_r}{N_p^2}\beta_w^j\beta_b^j\right) \text{Var}(P_{\theta})_j.$

В классической ДЭ за одно поколение значения функции вычисляются для N_p пробных векторов. В АДЭ за один шаг вычисляется значение функции для одного пробного вектора. Для корректного сравнения ограничений для АДЭ с ограничениями для классической ДЭ мы должны рассчитать, каким будет *фактор расширения* популяции для АДЭ после N_p шагов, и сопоставить его фактору расширения популяции для одного перехода между поколениями для ДЭ. Тогда $\tilde{\alpha}^j$, соответствующее α^j , можно вычислить из соотношения $1 + \alpha^j = (1 + \tilde{\alpha}^j)^{N_p}$ следующим образом: $\tilde{\alpha}^j = \sqrt[N_p]{1 + \alpha^j} - 1 \approx \alpha^j / N_p$, а для выполнения условий (1.10) и (1.11) необходимо потребовать $\alpha^j > 0$. Так как N_p зафиксировано, а эффективный кроссовер $C_r \in (0, 1]$, эти условия удобно записать в виде ограничений на масштабирующий фактор:

$$F^{j^2} > F_{\text{low}}^{j^2} \left(N_p, C_r, \alpha^j, \beta_b^j, \beta_w^j \right), \quad (1.13)$$

в котором значения $F_{\text{low}}^{j^2}$ выражены через N_p , C_r и α^j . Здесь α^j — фактор расширения популяции после N_p вычислений функции. Значения $F_{\text{low}}^{j^2}$ для каждой из рассматриваемых стратегий будут следующими¹⁰:

¹⁰Если не учитывать П1 (см. с. 29), то значения для F_{low} для классических стратегий ДЭ будут:

DE/rand/1/bin [79]:	$F_{\text{low}}^2 = \frac{\alpha}{2C_r} + \frac{1}{N_p} - \frac{C_r}{2N_p};$
DE/best/1/bin:	$F_{\text{low}}^2 = \frac{\alpha}{2C_r} + \frac{N_p+1}{2N_p} - \frac{C_r}{2N_p} - \frac{N_p-1}{2N_p}(1 - C_r)\beta_b^2.$

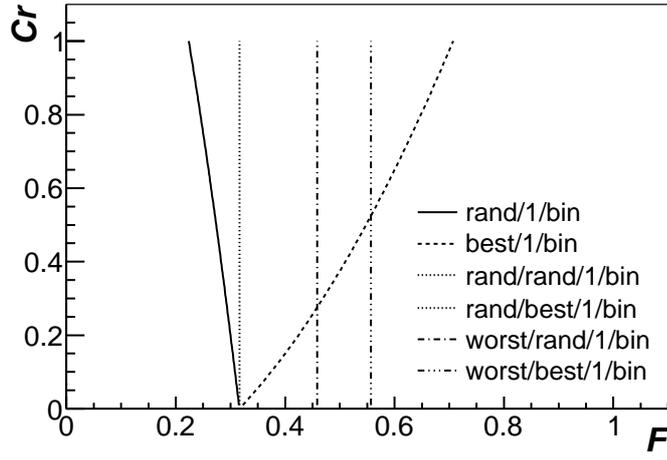


Рис. 1.3. $F_l(C_r, \alpha=0, N_p=10)$ -ограничения для различных стратегий ДЭ и АДЭ [8]

Классические стратегии ДЭ:

rand/1/bin	$F_{\text{low}}^j = \frac{\alpha^j}{2C_r} + \frac{1}{N_p-1} - \frac{N_p C_r}{2(N_p-1)^2};$
best/1/bin:	$F_{\text{low}}^j = \frac{\alpha^j}{2C_r} + \frac{1}{2} + \frac{1-C_r}{2N_p} - \frac{N_p-1}{2N_p}(1-C_r)\beta_b^{j^2}.$

Стратегии асинхронной ДЭ (АДЭ):

rand/rand/1/bin:	$F_{\text{low}}^j = \frac{N_p \alpha^j}{2C_r} + \frac{1}{N_p} \approx \frac{\alpha^j}{2C_r} + \frac{1}{N_p};$
rand/best/1/bin:	$F_{\text{low}}^j \approx \frac{\alpha^j}{2C_r} + \frac{N_p+1}{2N_p} - \frac{N_p-1}{2N_p}\beta_b^{j^2};$
worst/rand/1/bin:	$F_{\text{low}}^j \approx \frac{\alpha^j}{2C_r} - \frac{N_p-1}{2N_p} + \frac{N_p+1}{2N_p}\beta_w^{j^2};$
worst/best/1/bin:	$F_{\text{low}}^j \approx \frac{\alpha^j}{2C_r} + \frac{N_p+1}{2N_p}\beta_w^{j^2} - \frac{N_p-1}{2N_p}\beta_b^{j^2} + \frac{\beta_b^j \beta_w^j}{2N_p}.$

Таким образом, нижняя граница для масштабирующего фактора F_{low}^j зависит от размера выбранной популяции: с ростом размера популяции F_{low}^j снижается. Увеличение требуемого значения фактора расширения популяции α^j сдвигает нижнюю границу для масштабирующего фактора в сторону больших F . При снижении величины эффективного кроссовера нижняя граница масштабирующего фактора сдвигается в сторону больших F . Различие в результатах для различных стратегий заключается в слагаемых, зависящих от координат определенного базового ($\vec{\theta}_b$ — лучший член популяции) и/или определенного целевого ($\vec{\theta}_w$ — худший член популяции) векторов популяции в тех случаях, когда они явно используются в стратегии. В практических приложениях эти члены могут быть напрямую рассчитаны в текущей популяции.

На рис. 1.3 показаны ограничения $F_{\text{low}}(C_r, \alpha=0, N_p=10)$ для двух классических и четырех асинхронных стратегий. Эти ограничения изображены с учетом того, что $E(\beta_b^2) \approx 1$, $E(\beta_w^2) \approx 1,2$ и $E(\beta_b \beta_w) \approx 0$.

Комбинации параметров, при которых ожидаемая дисперсия популяции увеличивается после мутации и кроссовера, соответствуют большим значениям F . Ограничение для стратегии DE/rand/rand/1/bin мало отличается от ограничения для DE/rand/1/bin. Ограничение для DE/rand/best/1/bin практически совпадает с ограничением для DE/rand/rand/1/bin, в то время как DE/best/1/bin имеет меньшую область сходимости при больших значений C_r . Критическое значение F_{low} для DE/worst/rand/1/bin больше, чем для DE/rand/rand/1/bin. Для DE/worst/best/1/bin необходимы еще большие значения F , чтобы избежать преждевременной сходимости. Различие в допустимых областях контрольных параметров (рис. 1.3) качественно обусловлено относительной стохастичностью стратегий.

1.2.2. Эмпирические оценки областей сходимости АДЭ

Зависимость скорости и вероятности сходимости различных стратегий АДЭ, решающих задачу f_1 из CEC-2005 (10-мерную сдвинутую гиперсферу), от параметров алгоритма представлены на рис. 1.4. Каждая ячейка гистограммы показывает вероятность сходимости и скорость сходимости, усредненные по 25 попыткам минимизации выбранной функции. Ограничение по максимальному числу вычислений функции для каждой попытки: $N_{\text{FE}}^{\text{max}}=10^5$. Числа в ячейках показывают оценку для вероятности сходимости $P_{\text{succ}} \geq 0,4$, которая вычислялась как отношение числа сошедшихся к глобальному минимуму попыток к общему числу попыток. Чёрно-белая палитра показывает среднее число вычислений функции, потребовавшееся для нахождения минимума (среди успешных попыток): $\langle N_{\text{FE}} \rangle$. Кривые показывают теоретические ограничения на $\{F, C_r\}$ для различных по величине факторов расширения популяции: $\alpha = 0$, $\alpha = 0,1$ и $\alpha = 0,5$.

Возрастание фактора расширения приводит к снижению скорости сходимости. Для $F > 1$ минимизация занимает больше времени, чем при $F < 1$. Для $F = 1$ вероятность успеха ниже из-за уменьшения разнообразия среди мутантных векторов.

При решении задачи f_1 (гиперсферы) как область сходимости, так и скорость сходимости для классической стратегии DE/rand/1/bin и ее асинхронного аналога DE/rand/rand/1/bin похожи (рис. 1.4, а, в). Область сходимости классической DE/best/1/bin-стратегии хорошо согласуется с теоретически рассчитанной областью увеличения диспер-

сии популяции после мутации и кроссовера (рис. 1.4, б). Асинхронная DE/rand/best/1/bin-стратегия (рис. 1.4, г) по сравнению с классическим аналогом DE/best/1/bin имеет более широкую область сходимости. Асинхронные стратегии DE/worst/rand/1/bin и DE/worst/best/1/bin (рис. 1.4, д, е) сходятся быстрее, но требуют бóльших значений для масштабирующего фактора F для предупреждения преждевременной сходимости.

Область сходимости для решения задачи f_6 (сдвинутая функция Розенброка), представленная на рис. 1.5, в основном лежит в области с $\alpha > 0,1$. Так как f_6 — задача с неразделяемыми переменными, то бóльшие значения C_r приводят к увеличению вероятности сходимости.

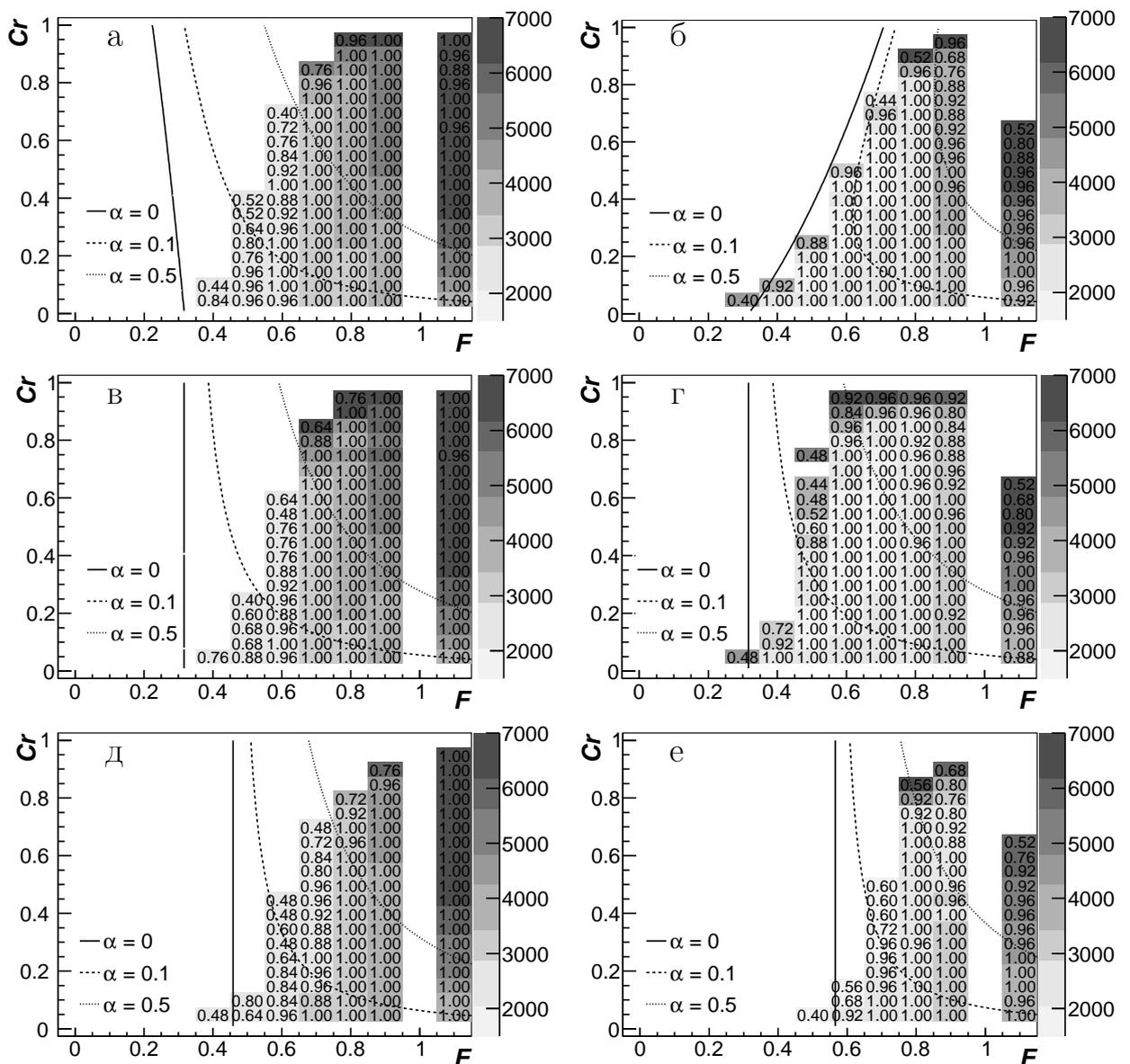


Рис. 1.4. Решение задачи f_1 из CEC-2005 (сдвинутая 10-мерная гиперсфера), $N_p = 10$, $N_{FE}^{\max} = 10^5$; а) DE/rand/1/bin, б) DE/best/1/bin, в) DE/rand/rand/1/bin, г) DE/rand/best/1/bin, д) DE/worst/rand/1/bin, е) DE/worst/best/1/bin [8]

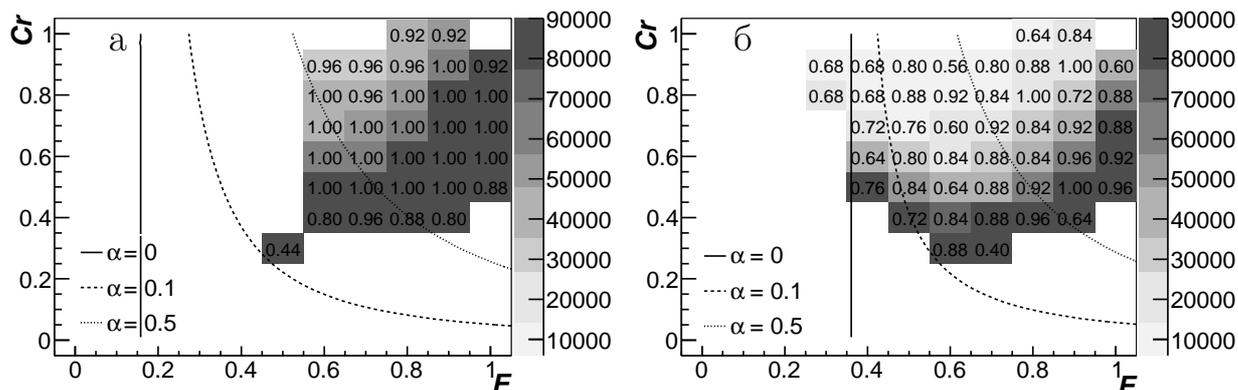


Рис. 1.5. Сдвинутая повернутая 10-мерная задача Розенброка (f_6). $N_{FE}^{\max} = 2 \times 10^5$: а) DE/rand/rand/1/bin, $N_p = 40$, б) DE/worst/best/1/bin, $N_p = 70$ [8]

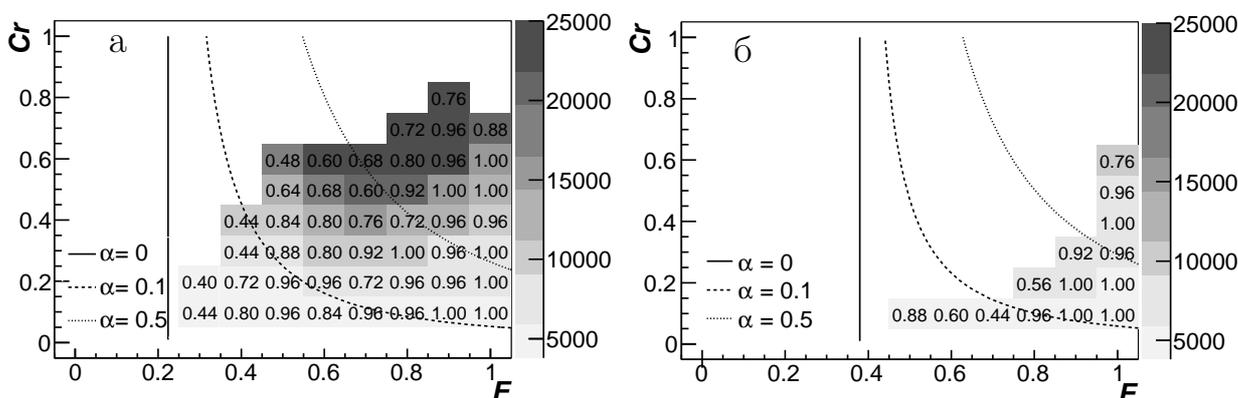


Рис. 1.6. Среднее количество вычислений функции $\langle N_{FE} \rangle$ при $P_{\text{succ}} \geq 0.4$ для сдвинутой функции Растригина (f_9 , $D = 10$): $N_{FE}^{\max} = 10^5$; а) DE/rand/rand/1/bin, $N_p = 20$, б) DE/worst/best/1/bin, $N_p = 25$. Числа показывают P_{succ} при $P_{\text{succ}} \geq 0.4$. α — показатель расширения популяции [8]

Ультрамногоэкстремальная задача f_9 (сдвинутая функция Растригина) является задачей с разделяемыми переменными. и для нее более предпочтительны малые значения кроссовера (рис. 1.6). Стратегия DE/rand/rand/1/bin (рис. 1.6, а) обладает хорошей поисковой способностью и находит минимум быстрее и с большей вероятностью, чем стратегия DE/worst/best/1/bin (рис. 1.6, б).

Сдвинутая повернутая функция Вейерштрасса (f_{11}) — ультрамногоэкстремальная, неразделяемая, в пределе — недифференцируемая. Область сходимости для этой функции небольшая (рис. 1.7). Оптимальный показатель расширения популяции $\alpha \sim 0.5$. Тем не менее, как стратегия DE/rand/rand/1/bin (рис. 1.7, а), так и стратегия DE/worst/rand/1/bin (рис. 1.7, б) успешно находят минимум с достаточной вероятностью при оптимально подобранных параметрах.

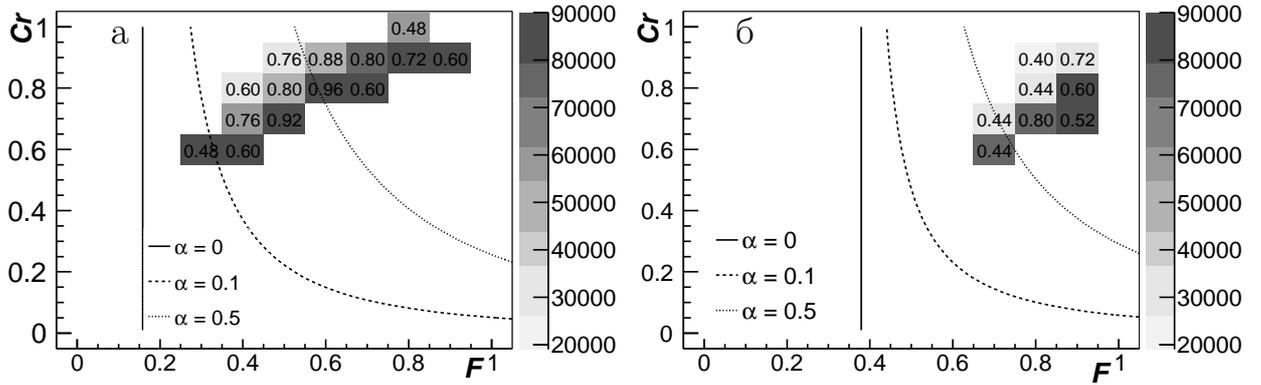


Рис. 1.7. Сдвинутая повернутая 10-мерная функция Вейерштрасса (f_{11}): $N_{FE}^{\max} = 2 \times 10^5$; а) DE/rand/rand/1/bin, $Np = 40$, б) DE/worst/best/1/bin, $Np = 50$ [8]

Выводы по разделу 1.2

Таким образом, для того чтобы снизить вероятность преждевременной сходимости ДЭ и АДЭ, нужно выбирать управляющие параметры алгоритма такими, чтобы они удовлетворяли соотношению (1.13): $F^{j^2} > F_{\text{low}}^{j^2} (N_p, C_r, \alpha^j)$. Впервые выведены аналитические выражения для величины граничного значения $F_{\text{low}}^{j^2}$ для классической стратегии DE/best/1/bin и для четырех различных асинхронных стратегий. Нижняя граница для масштабирующего фактора F_{low}^j зависит от размера выбранной популяции: с ростом размера популяции F_{low}^j снижается. Увеличение требуемого значения фактора расширения популяции α^j сдвигает нижнюю границу для масштабирующего фактора в сторону больших F . При снижении величины эффективного кроссовера нижняя граница масштабирующего фактора сдвигается в сторону больших F . Проведено сканирование области сходимости для этих стратегий на примере поиска минимума для функций f_1, f_6, f_9, f_{11} из CEC-2005 [78]. Показано, что аналитические ограничения не противоречат эмпирически найденным областям сходимости. Подтверждено, что для простых, разделяемых задач предпочтительнее использовать небольшие значения C_r .

1.3. Асинхронная дифференциальная эволюция с рестартом

Результаты анализа сходимости АДЭ для ряда тестовых функций представлены в таблицах 1.2 и 1.3. При этом в табл. 1.3, в отличие от условий теста CEC-2005 [78], максимально допустимое число вычислений функции было увеличено с 10^5 до 10^7 . Но даже если совсем не ограничивать максимально допустимое число вычислений функции, вероятность находж-

дения глобального минимума часто существенно меньше единицы как для классической ДЭ, так и для АДЭ.

Анализ не сошедшихся к глобальному минимуму попыток минимизации выявил, что причиной неудач становится, либо i) преждевременная сходимость либо ii) окончание поиска в локальном минимуме. Вероятность обеих неудач может быть существенно снижена, если выбрать достаточно большой размер популяции N_p . То есть для АДЭ (и ДЭ) вероятность сходимости к глобальному минимуму возрастает при увеличении размера популяции N_p . С другой стороны, поиск с бóльшим размером популяции может потребовать значительно бóльших затрат компьютерного времени: для бóльших популяций характерна медленная скорость сходимости (рис. 1.8).

Для практических задач необходимо либо адаптировать размер популяции согласно сложности решаемой задачи, либо проводить ряд попыток минимизации с различным размером популяции, что требует дополнительных усилий со стороны исследователя и дополнительных затрат компьютерного (CPU)-времени.

1.3.1. Алгоритм и критерии рестарта АДЭ

Для того чтобы сбалансировать хорошую скорость сходимости с высокой вероятностью сходимости, реализован метод АДЭ с рестартом

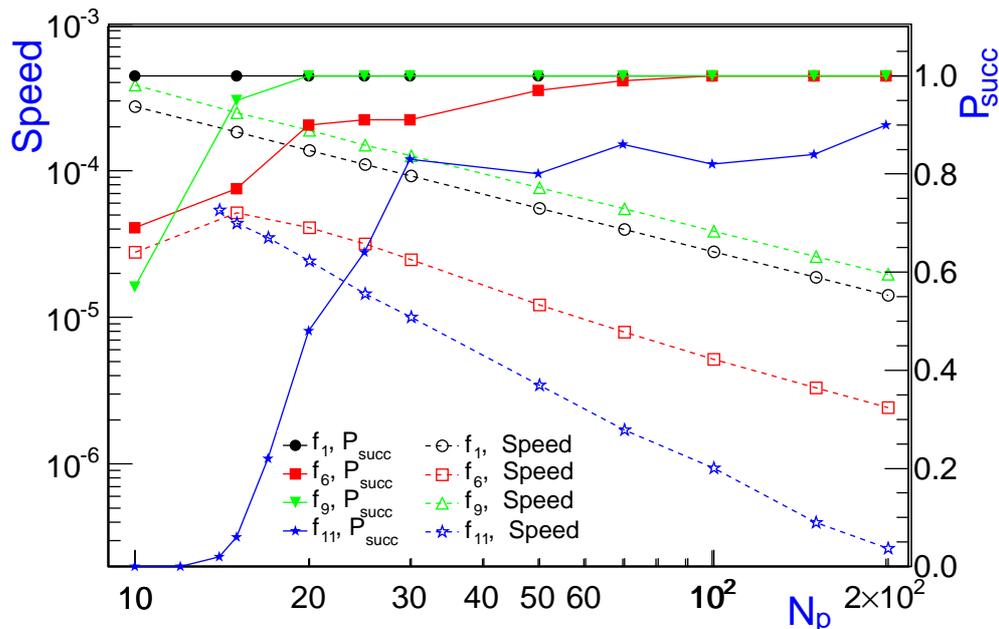


Рис. 1.8. Скорость сходимости $Speed = \langle N_{FE} \rangle^{-1}$ (пунктирные линии) и вероятность сходимости P_{succ} (сплошные) для DE/rand/rand/1/bin-стратегии метода АДЭ для f_1 , f_6 , f_9 и f_{11} размерности $D = 10$ тестовых функций из CEC-2005 [78]

(АДЭ-Р). Модифицированный алгоритм стартует с популяцией небольшого размера N_p^{\min} . Для численных экспериментов, представленных ниже, использовано $N_p^{\min}=10$. Когда в процессе поиска лучшего решения АДЭ-Р диагностирует стагнацию, алгоритм начинает расчет заново (рестартует), увеличивая при этом размер популяции в k раз (k — мультипликатор популяции).

Для диагностики стагнации алгоритм АДЭ-Р отслеживает максимальные разбросы координат членов популяции $\Delta\theta_j$ и значений целевой функции Δf :

$$\Delta\theta_j = \max_{i=0,\dots,N_p-1} \{\theta_{i,j}\} - \min_{i=0,\dots,N_p-1} \{\theta_{i,j}\}, \quad (1.14)$$

$$\Delta f = \max_{i=0,\dots,N_p-1} \{f_i\} - \min_{i=0,\dots,N_p-1} \{f_i\}. \quad (1.15)$$

Так как масштабирующий фактор F в операторе мутации (1.1) порядка единицы, разность в этом операторе должна быть больше, чем величина координат базового вектора, домноженная на «машинный эпсилон». В противном случае прибавление такой разности не повлечет за собой изменения в значениях координат из-за ошибок округления. Таким образом, в качестве критерия стагнации (и рестарта) предложено:

$$\exists j \in [0, D - 1] : \quad \Delta\theta_j < \varepsilon_\theta \max_i \{|\theta_{i,j}|\}. \quad (1.16)$$

В текущей реализации АДЭ использовано 8-байтное представление чисел с плавающей точкой (стандарт IEEE 754). Соответствующая величина машинного эпсилона около 10^{-16} . Типичное значение ε_θ должно превосходить 10^{-15} . Можно использовать бóльшие значения ε_θ . Тогда стагнация может быть диагностирована на более ранней стадии, что потенциально может ускорить сходимость алгоритма. Однако слишком большие значения ε_θ приведут к излишним рестартам. Необходимо отметить, что нижняя граница для ε_θ обусловлена машинной точностью и не зависит от минимизируемой функции.

Исходя из сравнения в операторе отбора существует другой критерий стагнации, основанный на разбросе значений целевой функции:

$$\Delta f < \varepsilon_f \max_{i=0,\dots,N_p-1} \{|f_i|\}. \quad (1.17)$$

Так же как и для ε_θ , нижний предел для ε_f обусловлен машинной точностью. Также следует отметить, что для важного класса задач, обобщенного метода наименьших квадратов, верхний предел для ε_f может быть оценен

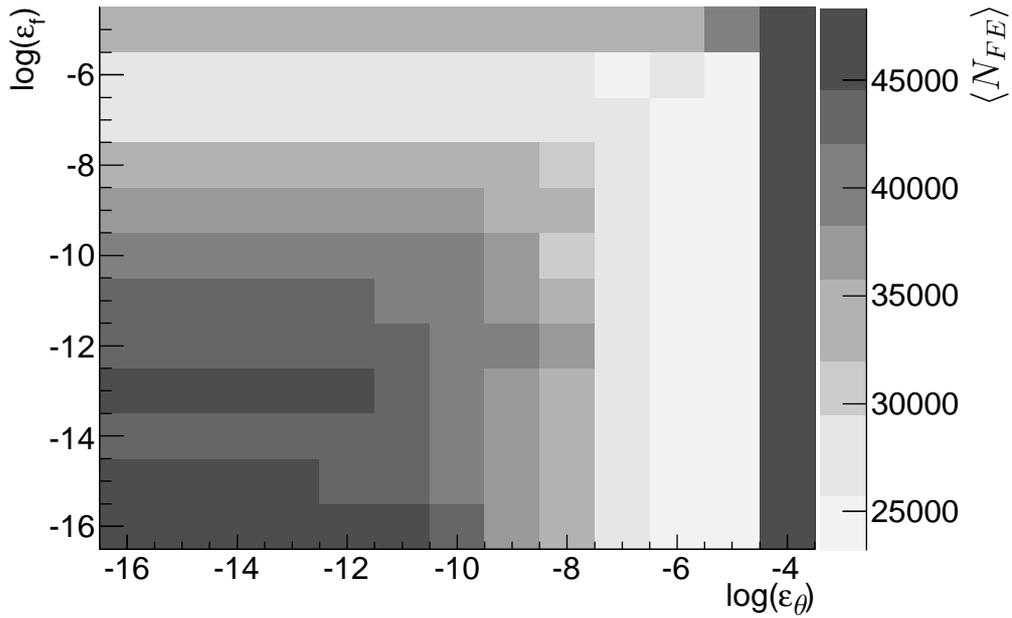


Рис. 1.9. Среднее количество вычислений функции $\langle N_{FE} \rangle$ в зависимости от ϵ_θ и ϵ_f . Решения задачи Розенброка f_6 , $D = 10$, стратегией DE/worst/best/1/bin метода АДЭ-Р, $k=2$, $\hat{P}_{\text{succ}} = N_{\text{succ}}/N_{\text{total}} = 1$ [9]

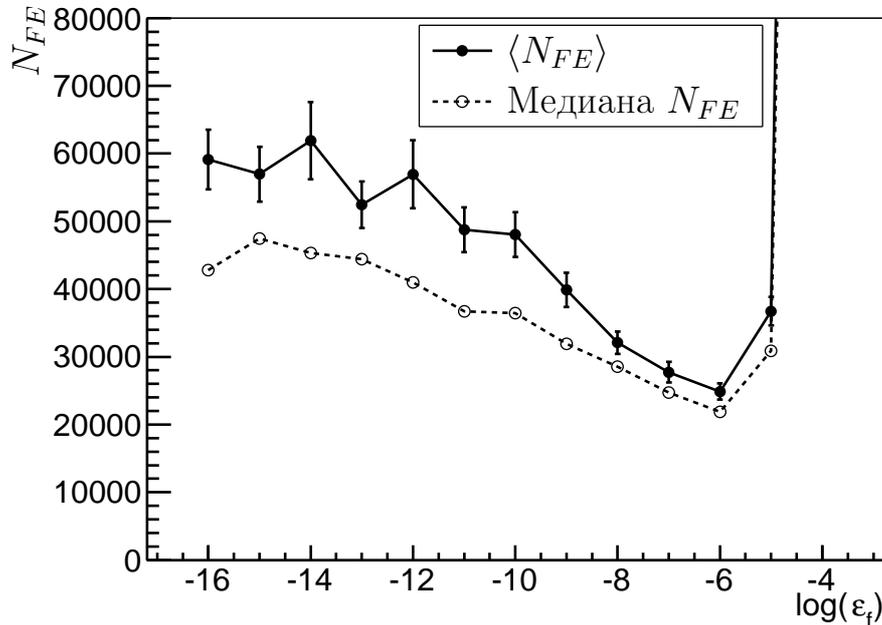


Рис. 1.10. Среднее $\langle N_{FE} \rangle$ и медианное количество вычислений функции в зависимости от ϵ_f без использования критерия (1.16). Решение задачи Розенброка f_6 , $D = 10$, стратегией DE/worst/best/1/bin метода АДЭ-Р, $k=2$, $P_{\text{succ}}=1$ [9]

a priori.

Таблица 1.4. Сходимость стратегий DE/rand/rand/1/bin и DE/worst/best/1/bin метода АДЭ-Р для четырех тестовых функций при $\varepsilon_\theta = 10^{-12}$ и различных ε_f [9]

	ε_f	DE/rand/rand/1/bin			DE/worst/best/1/bin			
		$\langle N_{FE} \rangle$	Медиана N_{FE}	P_{succ}	$\langle N_{FE} \rangle$	Медиана N_{FE}	P_{succ}	P_{succ}
f_1	10^{-12}	3641	3633	1,0	3070	2979	1,0	
f_6	10^{-7}	4,91e+04	4,20e+04	1,0	2,88e+04	2,30e+04	1,0	
	10^{-12}	5,11e+04	4,03e+04	1,0	4,09e+04	3,22e+04	1,0	
f_9	10^{-7}	5339	2977	1,0	5865	6913	1,0	
	10^{-12}	6121	2977	1,0	6942	8315	1,0	
f_{11}	10^{-7}	3,22e+05	0,75e+05	1,0	1,54e+05	1,14e+05	1,0	
	10^{-12}	2,02e+05	0,78e+05	1,0	1,78e+05	1,32e+05	1,0	

1.3.2. Зависимость сходимости от критериев рестарта

На рис. 1.9 представлена зависимость скорости сходимости для DE/worst/best/1/bin-стратегии как функции параметров применяемых критериев рестарта ε_θ и ε_f . На рис. 1.10 представлено среднее $\langle N_{FE} \rangle$ и медианное количество вычислений функции как функция ε_f из критерия (1.17) без использования критерия (1.16). В обоих случаях решалась задача Розенброка f_6 . Мультипликатор популяции k в этих расчетах равен двум. Для всех ячеек диаграмм $P_{succ} = 1,0$. Большие значения для ε_θ и ε_f приводят к более ранней диагностике стагнации эволюционного процесса, более раннему рестарту и более быстрой сходимости. Верхний предел для ε_f на уровне $4 \cdot 10^{-5}$ предопределен видом минимизируемой функции, в минимуме равной $f^* = 360$, и требуемой точностью на уровне 10^{-2} . Анализируя рис. 1.9, можно заключить, что при приближении значений ε_θ и ε_f к нижнему, не зависящему от задачи, а предопределенному машинной точностью пределу зависимость среднего числа вычислений функции от параметров ε_θ и ε_f слабая. Видно, что даже очень маленькие значения ε_θ и ε_f , равные $10^{-14} \dots 10^{-12}$, соотносятся с разумной скоростью сходимости. Таким образом, значения ε_θ и ε_f , равные машинному эпсилону, домноженному на фактор порядка $10^3 \dots 10^4$, обеспечивают удовлетворительную скорость сходимости.

Сравнение скорости сходимости стратегии DE/rand/rand/1/bin и DE/worst/best/1/bin метода АДЭ-Р для четырех тестовых функций при $\varepsilon_\theta = 10^{-12}$ и различных ε_f представлено в табл. 1.4.

Средние значения $\langle N_{FE} \rangle$ и медианы N_{FE} стратегии DE/rand/rand/1/bin для АДЭ в сравнении с АДЭ-Р представлены

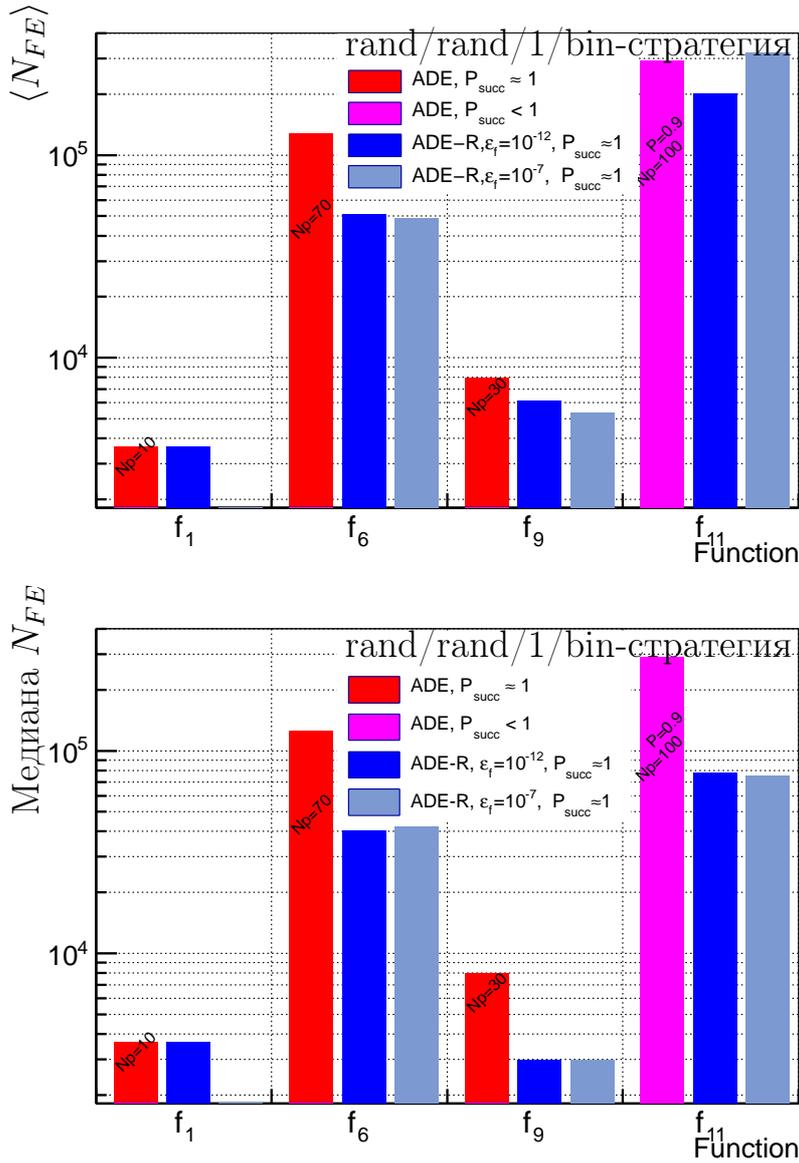


Рис. 1.11. Скорость сходимости стратегии DE/rand/rand/1/bin для АДЭ и АДЭ-Р. Для АДЭ: $N_p = N_p^{opt}$ зависит от целевой функции; $P_{succ} \approx 1$ для f_1 , f_6 , f_9 и $P_{succ} < 1$ для f_{11} . Для АДЭ-Р: $\epsilon_\theta = 10^{-12}$, $\epsilon_f = 10^{-12}$ и $\epsilon_f = 10^{-7}$; $N_p^{min} = 10$, $P_{succ} \approx 1 \forall f$. Среднее значение $\langle N_{FE} \rangle$ (сверху) и медиана N_{FE} (снизу) необходимого количества вычислений функции для достижения требуемой точности. Данные из [9]

на рис. 1.11. В АДЭ-Р для всех функций в качестве начального выбирался размер популяции $N_p^{min} = 10$. Мультипликатор популяции $k = 2$ использован для увеличения размера популяции после рестарта. Для всех функций достигнуто $\hat{P}_{succ} = N_{succ}/N_{total} = 1$ при полном количестве попыток $N_{total} = 100$. Сравнение проведено с использованием метода АДЭ без рестарта с размером популяции, «оптимально» подобранным для каждой из функций: выбран размер, обеспечивающий близкую к единице вероятность сходимости с наименьшими затратами компьютерного времени. При этом для функции f_{11} оценка вероятности сходимости к глобальному минимуму для АДЭ без рестарта равна $\hat{P}_{succ} = 0,9$ при популяции $N_p = 100$, в то время как для АДЭ-Р все попытки сошлись к глобальному минимуму (рис. 1.11). При решении задач Розенброка, Растргина и Вейерштрасса

с использованием метода АДЭ-Р требуется приблизительно в три раза меньше (медианная оценка) вычислений целевой функции, чем с методом АДЭ без рестарта. При этом для функции Вейерштрасса вероятность сходимости АДЭ-Р близка к единице, тогда как для АДЭ без рестарта эта вероятность сходимости составляет около 90%.

АДЭ-Р автоматически адаптирует размер популяции под сложность решаемой задачи. Стоит также отметить, что для сложных функций скорость сходимости могла быть быстрее при увеличении начального размера популяции N_p^{\min} .

1.3.3. Зависимость сходимости от мультипликатора популяции k

Зависимость скорости сходимости от мультипликатора популяции k также обсуждалась в нашей работе [9] (табл. 1.5, рис. 1.12). Разумными значениями для k являются значения $k \in (1,2 \dots 3,0)$, причем для большинства задач $k \in [1,5, 2,0]$ будет наилучшим. Меньшие значения k означают меньший темп роста популяции, они приводят к большему количеству необходимых рестартов. При больших значениях k размер популяции растет слишком быстро, что увеличивает время расчетов.

Таблица 1.5. Сходимость АДЭ-Р ($\varepsilon_\theta = 10^{-12}$, $\varepsilon_f = 10^{-12}$) для различных величин мультипликатора популяции k [9]

	k	DE/rand/rand/1/bin			DE/worst/best/1/bin		
		$\langle N_{FE} \rangle$	Медиана N_{FE}	P_{succ}	$\langle N_{FE} \rangle$	Медиана N_{FE}	P_{succ}
f_6	1,2	4,88e+04	3,93e+04	1,0	8,56e+04	7,81e+04	1,0
	1,5	4,85e+04	3,93e+04	1,0	4,97e+04	4,14e+04	1,0
	2,0	5,11e+04	4,03e+04	1,0	4,09e+04	3,22e+04	1,0
	3,0	5,76e+04	4,06e+04	1,0	4,60e+04	3,42e+04	1,0
	5,0	7,42e+04	4,06e+04	1,0	5,60e+04	3,51e+04	1,0
f_{11}	1,2	2,15e+05	1,79e+05	1,0	2,37e+05	2,18e+05	1,0
	1,5	1,75e+05	1,14e+05	1,0	1,64e+05	1,27e+05	1,0
	2,0	2,02e+05	0,78e+05	1,0	1,78e+05	1,32e+05	1,0
	3,0	2,99e+05	1,16e+05	1,0	2,11e+05	1,14e+05	1,0
	5,0	1,00e+06	3,15e+05	0,96	2,66e+05	0,55e+05	1,0

Выводы по разделу 1.3

В этом разделе описан метод асинхронной дифференциальной эволюции с рестартом. Получено, что новый метод АДЭ-Р обеспечивает быструю скорость сходимости одновременно с высокой вероятностью сходимости по сравнению с ДЭ и АДЭ без рестарта. При этом исследователю не нужно

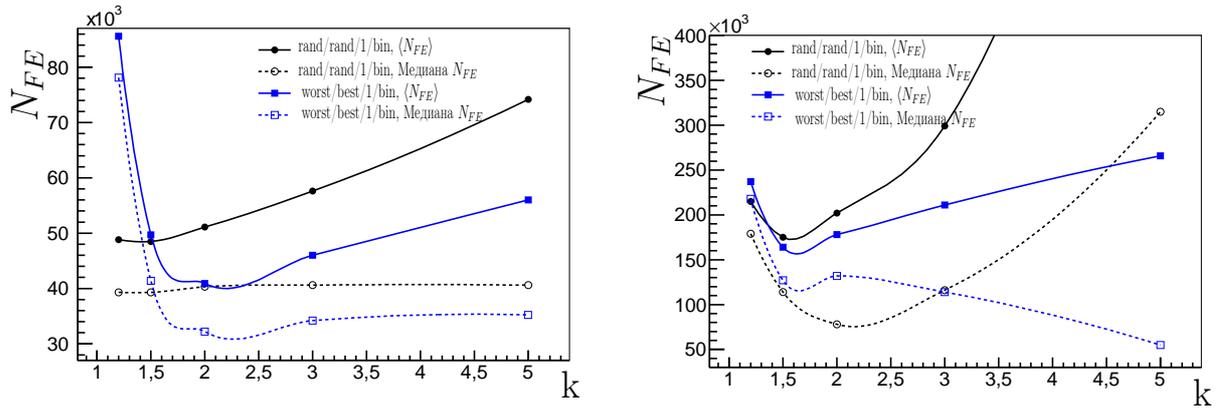


Рис. 1.12. Медиана (пунктирные линии) и среднее $\langle N_{FE} \rangle$ (сплошные) количество вычислений функции для стратегий DE/rand/rand/1/bin и DE/worst/best/1/bin метода АДЭ-Р ($\varepsilon_\theta = 10^{-12}$, $\varepsilon_f = 10^{-12}$) в зависимости от величины мультипликатора популяции k . Решение задачи Розенброка f_6 (слева). Решение задачи Вейерштрасса f_{11} (справа). Данные из [9]

подбирать «оптимальный» размер популяции. Проанализирована зависимость скорости и вероятности сходимости от величины параметров рестарта: критериев рестарта ε_x и ε_f и мультипликатора популяции k [9]. Получено, что для широкого класса задач зависимость скорости и вероятности сходимости от величины критериев рестарта ε_x и ε_f слабая. Значения критериев рестарта можно выбрать исходя из априорного знания о природе целевой функции (например, в случае минимизации невязки, являющейся χ^2 -распределением с m степенями свободы). В любом случае, величина критериев ε_x и ε_f должна быть на несколько порядков грубее, чем машинная точность представления вещественных чисел.

1.4. Асинхронная дифференциальная эволюция с кроссовером, задаваемым адаптивной корреляционной матрицей

Как известно, популяция в ДЭ адаптируется к ландшафту целевой функции [70]. Поэтому информация о корреляции между параметрами может быть получены из анализа ковариационной матрицы $cov(\theta_j, \theta_k)$, рассчитанной по координатам членов популяции.

Сформулируем алгоритм асинхронной дифференциальной эволюции с кроссовером, задаваемым адаптивной корреляционной матрицей (АДЭ-АКМ) [10]. Сформируем выборочную ковариационную матрицу членов те-

кущей популяции Q , элементы которой равны

$$q_{jk} = \frac{1}{N_p - 1} \sum_{i=0}^{N_p-1} (\theta_{ij} - \bar{\theta}_j) (\theta_{ik} - \bar{\theta}_k). \quad (1.18)$$

Отнормировав их на диагональные элементы, получаем выборочную корреляционную матрицу S :

$$s_{jk} = \frac{q_{jk}}{\sqrt{q_{jj}q_{kk}}}. \quad (1.19)$$

Так как важна скорость сходимости, предпочтительно иметь небольшую популяцию N_p . Однако в этом случае выборочная корреляционная матрица (1.19) не может быть надежной оценкой для фактической корреляционной матрицы. Вместо этого мы используем успешные шаги ДЭ для итеративного обновления оценки корреляционной матрицы — адаптивную корреляционную матрицу C :

$$C' = \nu C + (1 - \nu)S. \quad (1.20)$$

Коэффициент $0 < \nu < 1$ характеризует степень обучения адаптивной корреляционной матрицы. Если $\nu = 0$, обучения нет и используется выборочная корреляционная матрица. Большие значения ν связаны с расширением горизонта накопления информации. В численных экспериментах, описанных далее, мы использовали $\nu = 0,99$.

Исходя из адаптивной корреляционной матрицы, можно определить группы коррелированных переменных следующим образом. Алгоритм выбирает случайную координату m и порог (уровень) отсечения c_{thr} :

$$m = [N_p \cdot \text{rand}(0, 1)], \quad c_{\text{thr}} = \text{rand}(0, 1). \quad (1.21)$$

Анализируя все элементы c_{mj} адаптивной корреляционной матрицы C с фиксированным индексом m , алгоритм идентифицирует коррелированные переменные:

$$\{I_m\} = \forall j : |c_{mj}| > c_{\text{thr}}. \quad (1.22)$$

Набор коррелированных переменных $\{I_m\}$ определяет подпространство Ω_m в области поиска. Посредством модифицированной операции кроссовера все компоненты мутантного вектора v_i , которые входят в $\{I_m\}$, переходят в пробный вектор u_i , в то время как другие компоненты берутся

из целевого вектора θ_i :

$$u_{i,j} = \begin{cases} v_{ij}, & \text{если } j \in \{I_m\}, \\ \theta_{i,j}, & \text{если } j \notin \{I_m\}. \end{cases} \quad (1.23)$$

Так операция кроссовера (1.23) основана на адаптивной корреляционной матрице. Идентифицируя стратегии АДЭ, мы будем определять этот кроссовер символом «асм», например, DE/rand/rand/1/асм.

Со стороны кроссовера для текущей итерации (шага) алгоритм уменьшает размерность задачи до размерности подпространства Ω_m коррелированных переменных. В этом подпространстве применяется равномерный (биномиальный) кроссовер с $C_r = 1$. Следует отметить, что АСМ-кроссовер инвариантен относительно вращений в пределах Ω_m .

По построению (формулы (1.18) – (1.19)) коэффициенты c_{mj} в адаптивной корреляционной матрице отражают линейные корреляции между переменными. Чтобы минимизировать целевые функции с произвольными корреляциями между переменными, порог c_{thr} изменяется в пределах всего диапазона $[0, 1]$.

Масштабирующий фактор F_i следует адаптивному правилу [80]. Для каждой мутации (1.21) фактор F_i выбирается случайным образом из распределения Коши с коэффициентом сдвига μ_F и коэффициентом масштаба σ_F :

$$F'_i = \text{randc}(\mu_F, \sigma_F). \quad (1.24)$$

Если сгенерированное значение превышает F_{max} , оно приравнивается к F_{max} . Если $F'_i < F_{\text{min}}$, он генерируется заново. Если пробный вектор u_i выбран и заменяет родителя в популяции, коэффициент сдвига обновляется по следующей формуле:

$$\mu'_F = (1 - c_F)\mu_F + c_FL_2(F). \quad (1.25)$$

Здесь $L_2 = \frac{\sum_{i=0}^{N_p-1} F_i^2}{\sum_{i=0}^{N_p-1} F_i}$ — контрагармоническое среднее множества масштабирующих факторов, связанных с членами текущей популяции, c_F — скорость обучения для обновления коэффициента сдвига.

С помощью АСМ-кроссовера мы избавились от уровня кроссовера C_r в качестве управляющего параметра алгоритма АДЕ. Так как масштабирующий фактор F следует адаптивному правилу (1.24), а размер популяции N_p самоадаптируется к сложности целевой функции с помощью рестарта, асинхронная дифференциальная эволюция с адаптивной корреляционной

матрицей является квазипараметронезависимой с точки зрения пользователя.

1.4.1. Численные тесты и результаты

Результативность стратегий АДЭ оценена на наборе функций действительных аргументов из ВВОВ-2012 [41]. В качестве эталона протестированы два варианта стратегий DE/rand/rand/1/bin: с кроссовером $C_r = 0$ («b0») и $C_r = 0,9$ («b0.9»). Их производительность сравнивается с новой стратегией DE/rand/rand/1/asn («asn») и вариантом алгоритма JADE без архива («JADE») [80] с асинхронной DE/rand/rand/1/bin-стратегией. Алгоритм JADE использует адаптивную схему для кроссовера. Было показано, что он имеет более высокую производительность по сравнению с другими адаптивными DE-алгоритмами [69, 80]. Для всех четырех стратегий используются $F_{\min} = 0,1$, $F_{\max} = 1,0$, $c_F = 0,01$, $\sigma_F = 0,1$, начальные значения $\mu_F = 0,5$ и $N_p^{\min} = 10$. Если после очередного рестарта и увеличения размера популяции оказывается, что размер популяции превышает $20D$, иницируется независимый рестарт с размером N_p^{\min} . Максимальное число вычислений функции ограничивается $10^6 D$ для каждого прогона.

Численные результаты [10] представлены в приложении П. III на рис. III.1 и III.2 в соответствии с ВВОВ-2012 [41]. Попытка определялась как успешная и вычисления заканчивались, если алгоритм находил значение целевой функции ниже $f(\vec{\theta}^*) + 10^{-8}$. Ожидаемое время работы алгоритма (ERT) при ненулевой вероятности успеха рассчитывается как отношение числа вызовов целевой функции во всех испытаниях к числу успешных испытаний.

Для функций с разделяющимися переменными АДЭ-АКМ с адаптивной корреляционной матрицей показывает результаты, близкие к методу DE/rand/rand/1/bin с фиксированным кроссовером $C_r = 0$. Небольшое отставание следует отнести на счет определенной доли времени, затрачиваемого на «обучение» алгоритма ландшафту целевой функции. Для других функций DE/rand/rand/1/asn-стратегия является конкурентоспособной, по сравнению с АДЭ с $C_r = 0,9$, за исключением многоэкстремальных функций, где знание локальной структуры, представленное корреляционной матрицей, не играет существенной роли. Поведение ДЭ с фиксированным кроссовером существенно зависит от выбранного уровня C_r . Посредством АКМ-кроссовера мы исключили уровень кроссовера C_r из управляю-

щих параметров алгоритма АДЭ. АДЭ с кроссовером на основе адаптивной корреляционной матрицы сочетает в себе преимущества обеих схем с фиксированным кроссовером, поэтому может использоваться как для минимизации функций с разделяющимися, так и неразделяющимися и частично разделяющимися переменными. Это позволило решить все 24 тестовые задачи в пространстве с $D = 5$ и 18 из 24 задач при $D = 20$ в пределах выделенного количества вычислений целевой функции.

Анализ результатов тестов показывает, что большая часть плохоопределенных и многоэкстремальных функций могут быть решены только при помощи популяций ДЭ с относительно большим N_p . В то же время большое количество вычислений функции требуется еще до того, как процедура рестарта увеличит размер популяции в достаточной степени. Для сложных функций следует изначально выбирать большие значения N_p^{\min} или реализовать более эффективную процедуру рестарта.

АДЭ-АКМ и JADE показывают аналогичную производительность. JADE показывает лучшие результаты на плохо обусловленных функциях, в то время как для разделяемых и слабо структурированных многоэкстремальных функций АДЭ-АКМ имеет более высокую скорость сходимости. Здесь следует отметить, что набор функций ВВОВ-2012 содержит только полностью разделяемые и полностью неразделяемые задачи. Поэтому адаптивная схема, используемая в JADE, стремится назначить кроссовер C_r либо близким к нулю, либо к единице для соответствующих функций. Равномерное смешивание переменных в биномиальном кроссовере не имеет недостатков на этом наборе функций. АДЭ с адаптивной корреляционной матрицей может быть использована для минимизации более сложных функций с *частично* разделяющимися многомерными функциями, которые состоят из нескольких независимых (или слабо зависимых) субкомпонент размерности $D_j \geq 2$.

Таким образом, на основе корреляционного анализа разработан новый тип кроссовера. Этот кроссовер реализован в алгоритме асинхронной дифференциальной эволюции [10]. Получившийся новый алгоритм способен определить группы связанных переменных и, таким образом, эффективно решать разделяемые, частично разделяемые и неразделяемые минимизационные задачи без дополнительных усилий со стороны исследователей. Предлагаемый кроссовер в сочетании с адаптивной схемой для выбора масштабирующего фактора F и автоматическим увеличением размера

популяции посредством рестарта приводит с точки зрения пользователя к не требующему настройки параметров алгоритму. АДЭ-АКМ выполняет минимизацию на конкурентоспособном уровне по набору тестов функций ВВОВ-2012 [41]. АДЭ-АКМ, также как и АДЭ, успешно выполняет расчеты на параллельных вычислительных системах. Внутренняя простота алгоритма позволяет использовать его в качестве основы, в которую может быть включена любая стратегия ДЭ.

1.5. Программная реализация параллельного АДЭ-алгоритма

1.5.1. Комплекс программ, реализующий методы глобальной минимизации АДЭ

Комплекс программ, реализующий методы глобальной минимизации АДЭ и АДЭ-Р, создан на языке C++ с использованием технологии объектно-ориентированного программирования. Он состоит из блока, реализующего предложенные численные методы¹¹, и надстроек, реализующих запуск метода в различных режимах (рис. 1.13).

В режиме последовательных вычислений возможен прямой запуск из программного кода или с использованием технологии XML. В режиме па-

¹¹Заголовочные файлы (*.h) и соответствующие файлы с реализацией (*.cpp) — "ade_point", "cycle_array", "problem_definition", "parameter_list", "ade_settings", "ade_point_pool", "ade_iterator", "ade_criteria", "memory_utils", "ade_statu", "inequality_list" — содержат необходимые объявления и реализации элементов алгоритма.

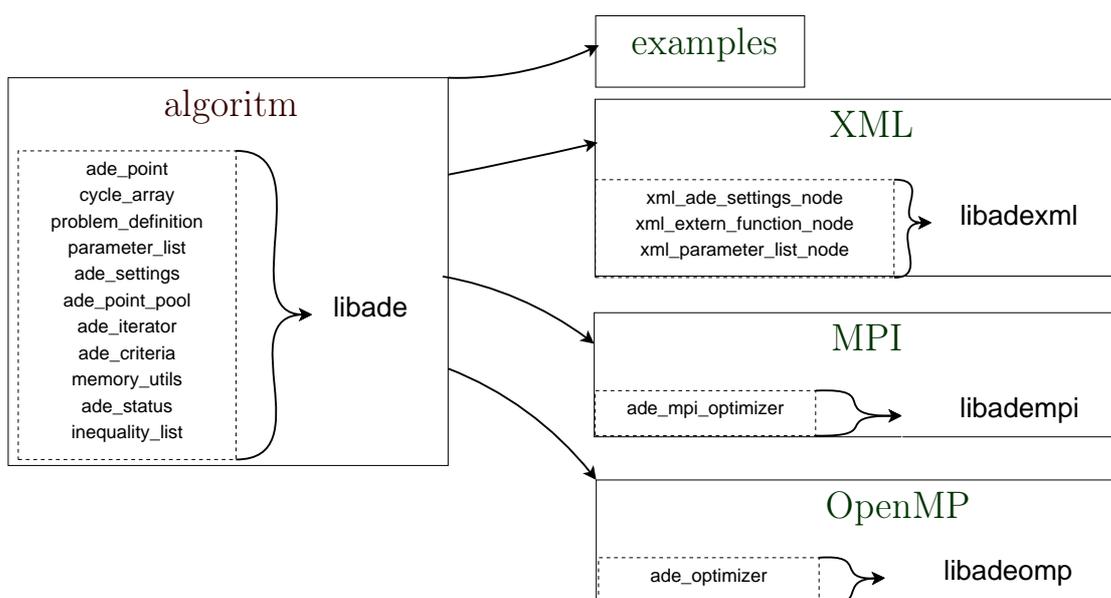


Рис. 1.13. Реализация АДЭ

раллельных вычислений — с использованием стандартов OpenMP¹² [7] или MPI¹³ [6, 10, 21]. Makefile позволяет корректно скомпилировать все компоненты комплекса, получив в результате набор библиотек¹⁴, которые в дальнейшем могут быть использованы для минимизации функций действительных переменных. Исходный код снабжен примерами, иллюстрирующими использование метода.

При минимизации целевых функций, требующих значительных вычислений, выгодно использовать одну из параллельных реализаций алгоритмов АДЭ и АДЭ-Р (рис. 1.14).

Далее в этом разделе диссертации описаны особенности параллельной реализации методов АДЭ и АДЭ-Р, представлены результаты оценки ускорения АДЭ и АДЭ-Р при параллельных вычислениях в сравнении с аналогичными результатами для ДЭ и несинхронизованной ДЭ с рестартом (ДЭ_{НС-Р}) на примере тестовых задач. Оценка значений ускорения АДЭ и АДЭ-Р при минимизации реальных многопараметрических моделей физических систем представлена в разделе 3.5 главы 3 диссертации.

1.5.2. Особенности параллельной реализации АДЭ и АДЭ-Р

Реализованы параллельные версии программы в стандартах OpenMP [7] и MPI [6, 10, 21]. Для распараллеливания использована модель ведущий/ведомый (рис. 1.14).

Ведущий процесс хранит данные о текущей популяции и проводит операции АДЭ. Значения целевой функции вычисляются N_{proc} параллельными ведомыми процессами. Мастер минимизации запрашивает пробные векторы из алгоритма и направляет их для расчета на ведомые процессоры. По мере расчета значений целевых функций результаты асинхронно передаются мастером в программу алгоритма. Класс АДЭ реализован таким образом, что для каждого пробного вектора отслеживается индекс соответствующего целевого вектора. Оператор отбора осуществляет сравнение вычисленного значения функции для пробного вектора с соответствующим значением для вектора текущей популяции с тем же индексом. Таким обра-

¹²OpenMP (Open Multi-Processing) — открытый стандарт для распараллеливания программ на многопроцессорных системах с общей памятью, реализованный на языках C, C++ и Фортран. Официальный сайт OpenMP: www.openmp.org.

¹³Message Passing Interface (MPI, интерфейс передачи сообщений) — программный интерфейс (API) для передачи информации, который позволяет обмениваться сообщениями между процессами, выполняющими одну задачу. Официальное описание стандарта MPI доступно на сайте www.mpi-forum.org.

¹⁴libade, libadeomp, libadempi

зом, в алгоритме корректно обрабатывается случай, когда первоначальный целевой вектор был замещен в популяции другим пробным вектором, рассчитанным с помощью более быстрого вычислительного потока. Освободившемуся ведомому процессору сразу же выставляется задание на расчет значения целевой функции следующего пробного вектора, что позволяет полностью и эффективно использовать все имеющиеся вычислительные мощности.

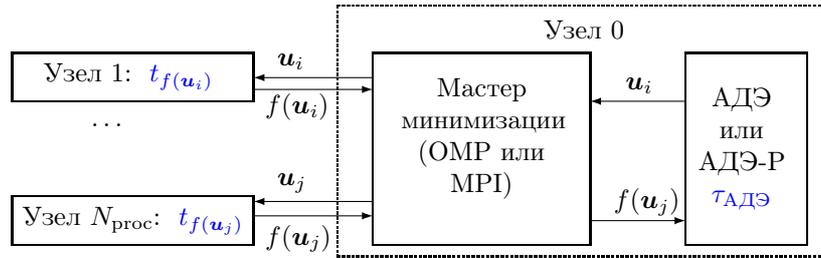


Рис. 1.14. Схема параллельной реализации алгоритма АДЭ [11]

1.5.3. Ускорение метода АДЭ при параллельных вычислениях

Сравнение ускорений при параллельных вычислениях для классической синхронной ДЭ и АДЭ выполнено для десятимерной задачи Розенброка из СЕС-2005 [78], с искусственно увеличенным временем расчета целевой функции [7] (рис. 1.15). Обычно в практических приложениях вычисление целевой функции — наиболее времязатратная часть расчетов. В проведенном тесте предполагалось, что время, необходимое для АДЭ-операций ($\tau_{\text{АДЭ}}$), пренебрежимо мало по сравнению с длительностью расчета целевой функции ($t_{f(u_i)}$):

$$t_{f(u_i)} \gg \tau_{\text{АДЭ}}. \quad (1.26)$$

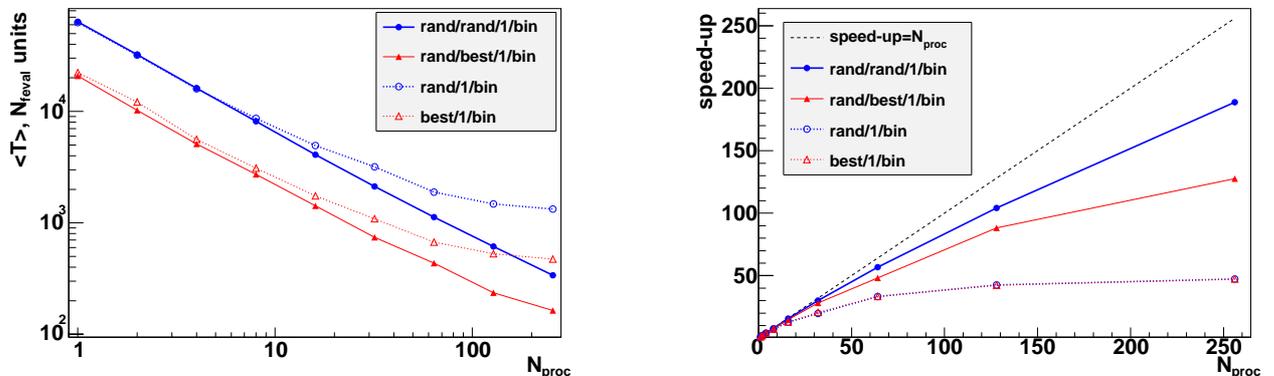


Рис. 1.15. Время решения задачи (слева) и соответствующее ускорение (справа) для параллельных вычислений с использованием различных стратегий классического и асинхронного вариантов ДЭ [7]. 10-мерная задача Розенброка

Время вычисления каждого значения целевой функции случайно и распределено согласно усеченному нормальному распределению¹⁵:

$$t_{f(\mathbf{u}_i)} \sim N(1, \sigma^2 | t > 0). \quad (1.27)$$

Здесь за единицу принято наиболее вероятное время вычисления функции. Разброс σ обусловлен различной производительностью ведомых процессов и/или различной сложностью вычисления значений целевой функции, соответствующих различным пробным векторам параметров.

Для АДЭ последовательность работы проста: как только ведомый процесс заканчивает вычисление целевой функции, он получает следующий целевой вектор для вычислений. Для классического синхронизированного варианта ДЭ ведомые процессы в первую очередь вычисляют значения целевой функции для пробных векторов, соответствующих еще не обработанным членам популяции. Если остались еще не занятые ведомые процессы, то они вычисляют значения целевых функций для заново сгенерированных пробных векторов, соответствующих случайно выбранным целевым векторам, для которых вычисление пробных векторов еще не закончено. Как только новое поколение полностью сформировано, все ведомые процессы останавливаются и алгоритм переключается на работу с новым поколением.

Распараллеливание модифицирует алгоритм, так как изменяется последовательность вычисления значений целевой функции для пробных векторов. Параллельная реализация стратегий АДЭ может потребовать отличного (по сравнению с реализацией той же стратегии в последовательной моде) количества вычислений целевой функции N_{FE} .

Время решения 10-мерной задачи Розенброка различными стратегиями параллельных реализаций АДЭ и классической синхронной ДЭ с $\sigma = 0,2$ представлено на рис. 1.15. Размер популяции зафиксирован: $N_p = 40$. Параллельные расчеты синхронизированными стратегиями ДЭ DE/rand/1/bin и DE/best/1/bin (пунктирные линии) довольно быстро выходят на уровень насыщения. В то же время соответствующие им асинхронные стратегии DE/rand/rand/1/bin и DE/rand/best/1/bin (сплошные линии) показывают более высокие значения для ускорения при параллель-

¹⁵Другими словами, вероятность $p(t)$ того, что время вычисления значения целевой функции $f(\mathbf{u}_i)$ составит t , есть $p(t) = \begin{cases} \phi\left(\frac{t-1}{\sigma}\right) / (1 - \Phi\left(-\frac{1}{\sigma}\right)), & t > 0, \\ 0, & t \leq 0, \end{cases}$ где $\phi(\cdot)$ – плотность вероятности стандартного нормального распределения, $\Phi(\cdot)$ – его функция распределения.

ных вычислениях.

Таким образом, для типичной минимизационной задачи была получена зависимость ускорения алгоритма АДЭ как функция доступного числа процессоров N_{proc} (рис. 1.15). Благодаря асинхронизации достигнуто лучшее ускорение по сравнению с классической синхронной дифференциальной эволюцией. Для параллельных вычислений с большим количеством параллельных процессов метод АДЭ значительно превосходит синхронизированный метод классической ДЭ.

1.5.4. Ускорение методов АДЭ-Р и несинхронизированной ДЭ со смешиванием поколений с рестартом (ДЭСП-Р)

В [6] оценки ускорения при параллельных вычислениях приведены для метода АДЭ с рестартом (про метод АДЭ-Р см п.1.3). АДЭ-Р стартует с популяции небольшого размера N_p^{\min} . Для рестарта задействованы критерии стагнации (1.16) и (1.17). При выполнении *критериев рестарта* алгоритм возобновляет вычисления с увеличенным в k раз размером популяции¹⁶.

Сравнение проводилось с параллельной ДЭ со смешиванием поколений [67] (ДЭСП), к которой применялся такой же метод рестарта (ДЭСП-Р). В этой реализации ДЭ освободившимся ведомым процессам поручается расчет пробных векторов, сгенерированных случайным образом по формуле (1.1) для еще не рассчитанных векторов из текущей популяции. Для каждого пробного вектора отслеживается индекс целевого вектора. Переход к следующему поколению осуществляется в момент, когда для каждого вектора текущей популяции существует хотя бы один пробный вектор с рассчитанной целевой функцией. Пробные векторы, рассчитанные позже перехода к следующему поколению, сравниваются с целевыми векторами с тем же индексом.

Скорость сходимости стратегий ДЭСП-Р и АДЭ-Р и ускорение для многопоточковых вычислений. В представленных ниже тестах, так же как и в тестах раздела 1.5.3, предполагается, что вычисление целевой функции требует значительных затрат компьютерного времени, а время, необходимое для выполнения операций ДЭ, пренебрежимо мало (1.26). Время вычисления каждого значения функции случайно и распределено

¹⁶В представленных ниже численных расчетах использованы $N_p^{\min} = 16$, мультипликатор популяции $k = 2$ и критерии рестарта: $\varepsilon_\theta = 10^{-12}$, $\varepsilon_f = 10^{-12}$.

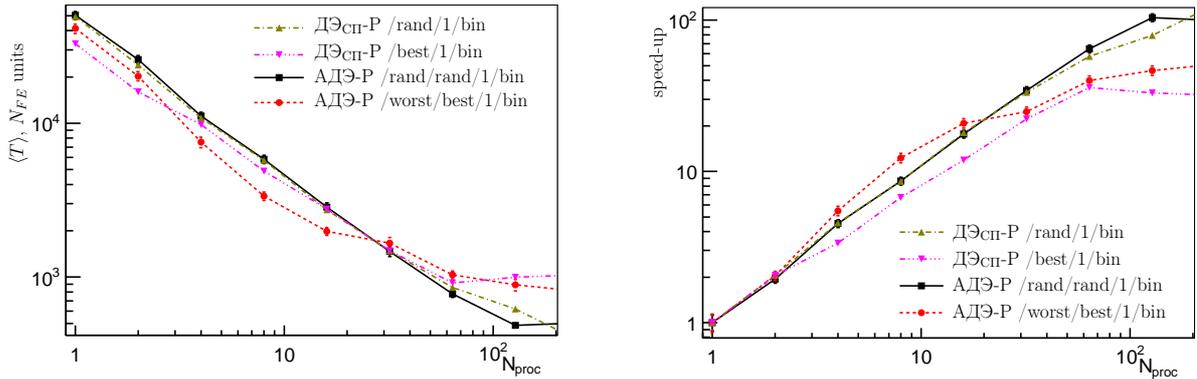


Рис. 1.16. Затраченное время (слева) и соответствующее ускорение (справа) для решения 10-мерной задачи Розенброка в многопоточном режиме при $\sigma = 0,2$ с использованием различных стратегий классического и асинхронного вариантов ДЭ_{СП}-Р с рестартом. Параметры методов АДЭ-Р и ДЭ_{СП}-Р: $F = 0,9$, $C_r = 0,9$, $N_p^{\min} = 16$, $k = 2$. Сходимость алгоритма к глобальному минимуму зарегистрирована в 100 случаях из 100 попыток для каждого измерения [9]

согласно усеченному нормальному распределению. За единицу принято наиболее вероятное время вычисления функции. Разброс σ обусловлен различной производительностью ведомых процессов и/или различной сложностью вычисления значений целевой функции, соответствующих различным пробным векторам.

Время, необходимое для решения задачи Розенброка (f_6 из набора тестовых функций СЕС-2005 [78]) различными стратегиями АДЭ-Р и ДЭ_{СП}-Р, и соответствующее ускорение при многопоточковых вычислениях представлены на рис. 1.16.

Для рассмотренных параллельных несинхронизованных стратегий классического ДЭ со смешиванием поколений с рестартом (ДЭ_{СП}-Р) (DE/rand/1/bin и DEbest/1/bin) и АДЭ-Р (DE/rand/rand/1/bin, DE/worst/best/1/bin, DE/linworst/linbest/1/bin, DE/linworst/best) ускорение остается близким к линейному с ростом количества вычислительных узлов до 50 и выше. Отметим, что для параллельной реализации классических стратегий ДЭ, DE/rand/1/bin и DE/best/1/bin (см. 1.5.3), оно в несколько раз ниже.

Для метода АДЭ-Р с ростом количества вычислительных узлов для стратегий DE/rand/rand/1/bin ускорение продолжает оставаться близким к линейному, однако для стратегии DE/worst/best/1/bin оно падает. Преимущество DE/rand/rand/1/bin обусловлено большей стохастичностью стратегии. Отметим, что ускорение, близкое к линейному, сохраняется для

числа задействованных процессоров, на порядок превышающего размер популяции.

- Ускорение АДЭ-Р близко к линейному до $N_{\text{proc}} \gtrsim 5N_p > N_p^{\min}$.
- Для $N_{\text{proc}} < N_p^{\min}$: стратегии с меньшей стохастичностью DE/best/1/bin, DE/worst/best/1/bin, DE/linworst/best/1/bin решают задачу быстрее.
- При $N_{\text{proc}} > N_p$ стратегии, сохраняющие разнообразие DE/rand/1/bin, DE/rand/rand/1/bin, сохраняют линейность ускорения. В результате обгоняют по скорости стратегии с меньшей стохастичностью.

Причины сверхускорения. Пока число потоков в среднем меньше, чем размер популяции ($N_{\text{proc}} \leq N_p^{\min} = 16$), ускорение для всех представленных стратегий практически линейно и даже немного превосходит N_{proc} .

Это объясняется тем, что при распараллеливании рассчитанные точки возвращаются в популяцию с запаздыванием (по сравнению с последовательной модой), что уменьшает вероятность вырождения популяции и увеличивает вероятность достижения минимума при меньшем N_p . Действительно, средний размер финишной популяции для $N_{\text{proc}} = 16$ для всех стратегий меньше, чем при однопоточных вычислениях (см. табл. 1.6). Дальнейшее увеличение количества процессоров приводит к тому, что для каждого целевого вектора рассчитывается в среднем (N_{proc}/N_p) пробных векторов, что неэффективно. Поэтому при дальнейшем росте количества задействованных процессоров для стратегий, лучше сохраняющих разнообразие в популяции, количество рестартов уменьшается. Для стратегии DE/rand/rand/1/bin

$$\langle N_p^{\text{fin}} \rangle_{N_{\text{proc}}=128} < \langle N_p^{\text{fin}} \rangle_{N_{\text{proc}}=1}.$$

Стратегии с меньшей стохастичностью, сохраняющие меньше разнообразия в популяции, рестартуют чаще и заканчивают вычисления с большим N_p^{fin} . Для стратегии DE/worst/best/1/bin

$$\langle N_p^{\text{fin}} \rangle_{N_{\text{proc}}=128} > \langle N_p^{\text{fin}} \rangle_{N_{\text{proc}}=1}.$$

Результативность стратегий АДЭ-Р при решении других задач из СЕС-2005 представлена в табл. 1.7. Видно, что при числе процессоров меньшем, чем средний размер популяции ($N_{\text{proc}} \sim 1$), стратегии направленного поиска, такие как DE/worst/best/1/bin, DE/linworst/best/1/bin, DE/linworst/linbest/1/bin, сходятся в среднем быстрее, чем стратегия

Таблица 1.6. Средний размер финишной популяции $\langle N_p^{\text{fin}} \rangle$ при разных N_{proc}

Стратегия	N_{proc}	1	16	128
	ADE-R/rand/rand/1/bin		14.2	13.6
ADE-R/worst/best/1/bin		26.7	22.9	36.5

DE/rand/rand/1/bin. Однако при числе задействованных процессов больше, чем средний размер популяции, эффективной оказывается лучше сохраняющая разнообразие в популяции стратегия DE/rand/rand/1/bin, для которой ускорение, сопоставимое с количеством задействованных вычислительных узлов, сохраняется вплоть до $N_{\text{proc}} \approx 100$. Стратегии DE/linworst/best/1/bin, DE/linworst/linbest/1/bin дают близкие к наилучшему результаты для любого числа вычислительных узлов как для функций с разделяемыми переменными, так и для многоэкстремальных функций с неразделяемыми переменными. Наибольшее ускорение при большом количестве вычислительных узлов получают решения более сложных задач, так как для таких задач задействованы популяции большего размера.

Таблица 1.7. Среднее время и ускорение при параллельных вычислениях стратегий АДЭ-Р при решении задач f_1 , f_6 , f_9 и f_{11} из СЕС-2005 [78] с увеличенным допустимым количеством вычислений функции (до 10^7) для разных N_{proc}

	Стратегия	N_{proc}	1		16		128		256	
		$\langle T \rangle$	$\langle T \rangle$	уск.	$\langle T \rangle$	уск.	$\langle T \rangle$	уск.	$\langle T \rangle$	уск.
f_1 : Hsph.	rand/rand/1/bin	3640	300	12	91	40	70.1	52		
	linworst/linbest/1/bin	3270	300	11	98	33	76.3	43		
	linworst/best/1/bin	3080	290	11	100	31	77.6	40		
	worst/best/1/bin	3070	600	5.1	280	11	235.2	13		
f_6 : Ros.	rand/rand/1/bin	50300	2850	18	480	104	504.0	100		
	linworst/linbest/1/bin	65300	3150	21	690	94	929.8	70		
	linworst/best/1/bin	39400	2490	16	660	59	1800 (480)	20 (80)		
	worst/best/1/bin	41300	1980	21	890	46	807.8	50		
f_9 : Rastr.	rand/rand/1/bin	6120	380	16	84.5	73	58.8	100		
	linworst/linbest/1/bin	5500	460	12	92	60	63.0	87		
	linworst/best/1/bin	6900	530	13	120	60	82.2	84		
	worst/best/1/bin	6900	940.2	7.4	390	18	238.3	29		
f_{11} : W.	rand/rand/1/bin	234000	14300	16	2000	120	2200	110		
	linworst/linbest/1/bin	191000	16300	12	2000	95	1100	170		
	linworst/best/1/bin	190000	11700	16	1800	105	980	190		
	worst/best/1/bin	203000	25100	8.1	11000	19	11000	19		

Итак, в этом разделе оценены скорости сходимости и ускорения при параллельных вычислениях для различных стратегий метода АДЭ-Р. Проведено сравнение с аналогичными характеристиками параллельной несинхронизированной ДЭ с аналогично реализованным рестартом (ДЭ_{СП}-Р). Представлены сравнительные характеристики различных асинхронных стратегий АДЭ-Р.

Выводы по разделу 1.5

В этом разделе представлены результаты тестов параллельной версии алгоритмов АДЭ и АДЭ-Р, реализованной с использованием модели ведущий/ведомый.

Получено, что метод АДЭ заметно превосходит классическую синхронизированную ДЭ при использовании большого количества параллельных процессоров (больше, чем размер популяции). Например, метод АДЭ, решая 10-мерную задачу Розенброка из СЕС-2005 с популяцией $N_p = 40$, позволяет получить результат в 180 раз быстрее, используя 256 параллельно работающих процессоров, чем используя один вычислительный узел. Аналогичное решение классической синхронизированной ДЭ сохраняет ускорение, близкое к линейному лишь при $N_{\text{проц}} \leq N_p$, а при использовании 256 параллельно работающих процессорах получает результат лишь в 48 раз быстрее, чем с использованием одного вычислительного узла.

Тестирование метода АДЭ-Р проводится при меньшем размере начальной популяции ($N_p^{\text{min}} = 10$). При этом средний размер конечной популяции $\langle N_p^{\text{fin}} \rangle$ также меньше «оптимального» N_p для АДЭ без рестарта, так как такой популяции оказывается достаточно для решения задачи с вероятностью, близкой к единице. С ростом количества задействованных вычислительных узлов для асинхронных стратегий с рестартом (ДЭ_{НС}-Р и АДЭ-Р) ускорение сопоставимо с количеством задействованных вычислительных узлов, в то время как для параллельной реализации классической синхронной ДЭ оно в несколько раз ниже.

Полученные значения для ускорения АДЭ-Р, сопоставимые с количеством задействованных процессоров $N_{\text{проц}}$, при $N_{\text{проц}} \approx 100$ (см. также раздел 3.5) говорят об эффективности применения АДЭ-Р на параллельных вычислительных системах.

Основные результаты, представленные в главе 1

Разработан метод асинхронной дифференциальной эволюции (АДЭ). Описано семейство методов АДЭ, проведены аналитические и эмпирические исследования свойств предложенных алгоритмов, описаны основные преимущества.

- Семейство представляемых методов включает в себя
 - Метод АДЭ [7]. В нем операции мутации, кроссовера и отбора, характерные для ДЭ, впервые задействованы в предельной форме квазистационарного (steady-state) алгоритма: на каждом шаге обновляется только один из членов популяции.
 - Метод АДЭ с рестартом [9, 11]. Он позволяет исключить размер популяции из числа управляющих параметров алгоритма. Автоматическое наращивание размера популяции в соответствии со сложностью решаемой задачи позволяет увеличить среднюю скорость и вероятность сходимости алгоритма при решении широкого класса задач.
 - Метод АДЭ-АКМ [10]. В этом методе предложен и впервые использован новый тип адаптивного кроссовера, учитывающий корреляции между параметрами. В нем также реализована адаптация масштабирующего фактора F . Это делает новый метод квазипараметронезависимым с точки зрения пользователя и позволяет находить решения широкого класса задач без дополнительных настроек параметров алгоритма.
 - Параллельные версии алгоритмов АДЭ, АДЭ-Р, АДЭ-АКМ.
- Аналитическое исследование методов заключается в нахождении границ допустимых значений управляющих параметров АДЭ.
 - Сформулирован математический подход к оценке допустимых значений параметров алгоритма (раздел 1.2). Получено, что необходимым условием предупреждения преждевременной сходимости популяции является выполнение соотношения (1.13) между управляющими параметрами (F и C_r при заданном N_p) алгоритма АДЭ.
 - Сканирование области сходимости для ряда стратегий на примере поиска минимума для функций f_1, f_6, f_9, f_{11} из СЕС-2005 [78]

показывает, что полученные аналитические ограничения не противоречат эмпирически найденным областям сходимости [8].

- Проведено эмпирическое исследование свойств алгоритма:
 - Получены оценки для скорости и вероятности сходимости методов АДЭ и АДЭ-Р для ряда тестовых функций из СЕС-2005 при различных значениях управляющих параметров (разделы 1.1.3 и 1.3).
 - Впервые проанализирована зависимость скорости сходимости от величины параметров рестарта метода АДЭ-Р (раздел 1.3).
 - Получены оценки для скорости и вероятности сходимости метода АДЭ-АКМ для тестовых функций ВВОВ-2012 и проведено сравнение с другими существующими алгоритмами (раздел 1.4).
- Преимущества методов АДЭ состоят в улучшении показателей скорости и вероятности сходимости по сравнению с ДЭ и в возможности эффективно применять методы в режиме параллельных вычислений,
 - Показано, что метод АДЭ имеет сопоставимые с классической ДЭ скорость и вероятность сходимости (раздел 1.1). Методы АДЭ-Р и АДЭ-АКМ улучшают показатели скорости и вероятности сходимости по сравнению с ДЭ (раздел 1.3). Вероятность сходимости метода АДЭ-АКМ находится на конкурентном уровне среди лучших из существующих адаптивных методов ДЭ, особенно для задач и частично разделяющимися переменными (раздел 1.4).
 - Показано, что новый метод АДЭ хорошо приспособлен для параллелизации (см раздел 1.5, а также раздел 3.5). Получено, что методы АДЭ и АДЭ-Р заметно превосходят по скорости классическую синхронизированную ДЭ и показывают ускорение, сопоставимое с количеством задействованных процессоров N_{proc} , при использовании большого количества параллельных процессоров — больше, чем размер популяции: $N_{\text{proc}} \approx 50$ (см. также раздел 3.5).

Таким образом, использование методов АДЭ и АДЭ-Р для расчетов как в последовательной моде, так и на параллельных вычислительных системах эффективно и перспективно.

Основные результаты данной главы опубликованы автором в работах [1, 6–13, 15].

2. Оценка параметров пион-нуклонной амплитуды в рамках микроскопической модели пион-ядерного рассеяния с использованием АДЭ

В этой главе представлено численное исследование микроскопической модели упругого пион-ядерного рассеяния (ММПЯУР) с использованием метода АДЭ. Микроскопическая модель пион-ядерного упругого рассеяния, построенная с использованием микроскопического оптического потенциала (МОП) и релятивистского волнового уравнения (РВУ) Шредингера, используется для оценки параметров пион-нуклонной амплитуды по данным о дифференциальных сечениях упругого рассеяния пионов на ядрах при кинетических энергиях налетающих пионов $T = 130 \div 219$ МэВ, покрывающих область энергии возбуждения пион-нуклонного $\Delta(1232)P_{33}$ -резонанса с массой πN -пары 1232 МэВ¹ [1, 2, 11, 15–18].

Исследованию пион-ядерного рассеяния посвящено значительное число теоретических и экспериментальных работ [87, 96, 101, 109]. Тем не менее, задача построения реалистичного теоретически обоснованного пион-ядерного потенциала, необходимого для моделирования не только упругого рассеяния пионов, но и более сложных пион-ядерных взаимодействий, остается актуальной. Большинство известных работ по расчетам пион-ядерных дифференциальных сечений основано на применении тех или иных феноменологических форм потенциала либо на использовании микроскопического потенциала Кислингера [98, 99], где учитывается вклад в пион-ядерное рассеяние s -, p - и d -волн πN -амплитуды. Все они имеют 6 (шесть) или более свободных параметров этого потенциала, подгоняемых по экспериментальным данным. Другой подход использует высокоэнергетическое приближение (ВЭП) Глаубера – Ситенко многократного рассеяния нуклонов ядрами [82, 85, 91, 92, 94, 96, 101, 107, 109–111], где амплитуда пион-ядерного рассеяния строится на основе πN -амплитуды в виде гауссовской функции, с учетом перерассеяния пиона на нуклонах ядра. В ВЭП при расчете сечений в ядерных задачах имеются известные технические трудности, связанные с вычислением многомерных интегралов при учете перерассеяния налетающей частицы на нуклонах ядра. Поэтому в таком подходе ограни-

¹Также в литературе $\Delta(1232)P_{33}$ -резонанс иногда обозначается как (3,3)-резонанс.

чиваются в основном анализом рассеяния на легких ядрах, когда достаточно использовать волновые функции ядер осцилляторного базиса, что существенно упрощает расчеты. Другая трудность состоит здесь в необходимости учета искажения классической траектории интегрирования в кулоновском и ядерном поле, когда анализируются данные рассеяния на ядрах среднего и тяжелого веса, где эти поправки оказываются важными.

В диссертации для построения потенциала использован подход, разработанный ранее в [84] для моделирования каон-ядерного рассеяния. Согласно этому подходу микроскопический оптический потенциал пион-ядерного рассеяния строится на основе интеграла свертки (фолдинга) двух функций: формфактора распределения плотности точечных нуклонов ядра-мишени, которые известны из данных независимых экспериментов, и амплитуды пион-нуклонного рассеяния. Расчет сечений ведется на основе решения соответствующего релятивистского волнового уравнения с использованием построенного потенциала. Амплитуда пион-ядерного рассеяния характеризуется тремя параметрами, значения которых подгоняются по экспериментальным данным. Тем самым появляется возможность получить новую информацию о механизме пион-нуклонного рассеяния в ядерной среде, а именно: как изменяются параметры рассеяния пионов на нуклонах, которые не являются свободными, а находятся в ядерной среде.

Оптимизационные задачи, возникающие при оценке параметров микроскопического оптического потенциала (МОП), обладают сложным профилем, имеют несколько локальных минимумов, и, возможно, профиль целевой функции не везде является гладким. Это может приводить к определенным трудностям при использовании методов поиска локального минимума: завершение расчетов в одном из второстепенных минимумов, диагностирование сходимости метода в точке, не являющейся минимумом, или даже отсутствие сходимости метода. Поэтому в диссертации для оптимизации параметров пион-нуклонной амплитуды использован метод АДЭ.

В начале главы изложены основные формулы, описывающие пион-ядерное упругое рассеяние в приближении МОП и процедуру расчета дифференциальных сечений упругого рассеяния π -мезонов на ядрах.

Далее приводится описание программного комплекса для нахождения параметров МОП по экспериментальным данным о дифференциальных сечениях упругого рассеяния π -мезонов на ядрах. Программный комплекс оптимизирует параметры модели для наилучшего соответствия экспери-

ментальным данным. Путем минимизации невязки между теоретически полученными сечениями и с экспериментальными данными подгоняются три параметра пион-нуклонной амплитуды: полное сечение πN -рассеяния, отношение реальной к мнимой части πN -амплитуды рассеяния вперед и параметр наклона, которые в дальнейшем сравниваются с соответствующими параметрами амплитуды рассеяния на свободных нуклонах.

В программном комплексе использован метод глобальной минимизации АДЭ. В следующих разделах обсуждаются преимущества АДЭ при таких оценках и ускорение параллельных вычислений для OpenMP-реализации программного комплекса.

В конце главы обсуждаются полученные результаты: i) рассчитанные дифференциальные сечения упругого рассеяния π -мезонов на ряде ядер при энергиях от 130 до 290 МэВ, согласующиеся с экспериментальными данными; ii) полученные параметры πN -амплитуды рассеяния пионов на нуклонах ядерной среды; iii) сравнение сходимости метода АДЭ и метода симплексов и квазиньютоновского метода с формулой Давидона – Флетчера – Пауэлла для апдейта гессиана при оценке параметров МОП.

2.1. Постановка задачи, основные формулы

Микроскопический оптический потенциал. За основу расчетов взят микроскопический оптический потенциал, полученный в [82, 86] на основе теории многократного высокоэнергетического рассеяния частиц на составных системах [90, 94, 110] в оптическом пределе. В этом пределе механизм процесса интерпретируется как однократное рассеяние пионов на нуклонах ядра с πN -амплитудой, просуммированной по всем нуклонам, а распределение нуклонов в ядре задается функцией ядерной плотности. Обе эти величины известны, вообще говоря, из независимых данных: плотность — из электрон-ядерного рассеяния, а πN -амплитуда — из рассеяния пионов на водородной мишени. Действительно, поскольку непосредственно измерить рассеяние пионов на свободных нейтронах невозможно, то при анализе рассеяния на ядрах учитывается изоспиновая симметрия

$$f(\pi^{\pm}n) = f(\pi^{\pm}p), \quad (2.1)$$

что позволяет использовать только данные о рассеянии π^\pm -мезонов на свободных протонах. Итак, согласно [82] имеем

$$U_{\pi^\pm A}^H = -\frac{1}{\pi E} \sum_{N=p,n} \int_0^\infty j_0(qr) \rho_N(q) F_{\pi^\pm N}(q) q^2 dq. \quad (2.2)$$

Здесь² E — полная энергия пиона в с.ц.м.; $j_0(qr)$ — сферическая функция Бесселя; $\rho_p(q)$ и $\rho_n(q)$ — формфакторы функций распределения плотности точечных протонов и нейтронов ядра. Пион-нуклонная амплитуда рассеяния $F_{\pi^\pm N}$ в (2.2) представляется в виде

$$F_{\pi^\pm N}(q) = \frac{k}{4\pi} \sigma_{\pi^\pm N} [i + \alpha_{\pi^\pm N}] \cdot f_{\pi^\pm N}(q), \quad (2.3)$$

где $\sigma_{\pi^\pm N}$, $\alpha_{\pi^\pm N}$ и $\beta_{\pi^\pm N}$ есть параметры πN -амплитуды рассеяния, в которой формфактор

$$f_{\pi^\pm N}(q) = e^{-\frac{\beta_{\pi^\pm N} q^2}{2}} \quad (2.4)$$

определяет зависимость амплитуды от переданного импульса q . Для свободного πN -рассеяния эти параметры сведены в известные таблицы [89, 103]. Подставляя (2.3) в (2.2), получаем

$$\begin{aligned} U_{\pi^\pm A}^H &= V^H + iW^H = \\ &= -\frac{k}{4\pi^2 E} \sum_{N=p,n} \sigma_{\pi^\pm N} (i + \alpha_{\pi^\pm N}) \int_0^\infty j_0(qr) \rho_N(q) f_{\pi^\pm N}(q) q^2 dq. \end{aligned} \quad (2.5)$$

Здесь $E = \sqrt{k^2 + m_\pi^2}$ и $k = \frac{M_A k^{\text{lab}}}{\sqrt{(m_\pi + M_A)^2 + 2M_A T^{\text{lab}}}}$ — полная энергия и импульс пиона в с.ц.м., m_π — масса пиона, M_A — масса ядра, $k^{\text{lab}} = \sqrt{T^{\text{lab}} (T^{\text{lab}} + 2m_\pi)}$ — импульс, E^{lab} — полная, $T^{\text{lab}} = E^{\text{lab}} - m_\pi$ — кинетическая энергия пиона в лабораторной системе координат. Формфакторы функций распределения плотности точечных протонов и нейтронов ядра, $\rho_p(q)$ и $\rho_n(q)$, нормированы соответственно на Z и $(A - Z)$.

Учитывая соотношение изотопической симметрии (2.1) для ядер с близким числом протонов и нейтронов $Z \simeq (A - Z)$, можно предположить, что их функции распределения плотности одинаковы: $\rho_p = \rho_n = \frac{1}{2}\rho$, где ядерная плотность нуклонов $\rho(r)$ нормирована на атомную массу A . Обозначив скорость пиона как $\beta_c = k/E$, получим, что потенциал (2.5)

²Здесь и далее в этом разделе используется естественная система единиц (Natural units), в которой $\hbar = c = 1$; $\hbar = h/2\pi$ — редуцированная (или рационализированная, или приведённая) постоянная Планка или постоянная Дирака; h — постоянная Планка; c — скорость света в вакууме.

принимает вид

$$U_{\pi^{\pm}A}^H(r) = -\frac{\beta_c}{(2\pi)^2} \sigma_{\pi^{\pm}} (i + \alpha_{\pi^{\pm}}) \int_0^{\infty} j_0(qr) \rho(q) f_{\pi^{\pm}}(q) q^2 dq, \quad \text{где } f_{\pi^{\pm}}(q) = e^{-\frac{\beta_{\pi^{\pm}} \cdot q^2}{2}}. \quad (2.6)$$

Потенциал (2.6) зависит от параметров $\sigma_{\pi^{\pm}}$, $\alpha_{\pi^{\pm}}$ и $\beta_{\pi^{\pm}}$ πN -амплитуды рассеяния, которые с учетом изотопической симметрии (2.1) выражаются через параметры пион-протонной амплитуды:

$$\begin{aligned} \sigma_{\pi^+} &= \sigma_{\pi^+p}, & \alpha_{\pi^+} &= \alpha_{\pi^+p}, & \beta_{\pi^+} &= \beta_{\pi^+p}, \\ \sigma_{\pi^-} &= \sigma_{\pi^-p}, & \alpha_{\pi^-} &= \alpha_{\pi^-p}, & \beta_{\pi^-} &= \beta_{\pi^-p}. \end{aligned} \quad (2.7)$$

С учетом зарядовой симметрии в некоторых работах был использован еще более упрощенный подход. Например, в работах [89,103] приведены усредненные параметры πN -амплитуд рассеяния для ряда энергий в области (3,3)-резонанса:

$$\sigma = \frac{1}{2}[\sigma_{\pi^+p} + \sigma_{\pi^-p}], \quad \alpha = \frac{1}{2}[\alpha_{\pi^+p} + \alpha_{\pi^-p}], \quad \beta = \frac{1}{2}[\beta_{\pi^+p} + \beta_{\pi^-p}]. \quad (2.8)$$

Потенциал (2.6) с учетом зарядовой симметрии принимает вид

$$U^H(r) = -\frac{\beta_c}{(2\pi)^2} \sigma (i + \alpha) \int_0^{\infty} j_0(qr) \rho(q) f(q) q^2 dq, \quad \text{где } f(q) = e^{-\frac{\beta \cdot q^2}{2}} \quad (2.9)$$

и зависит от параметров σ , α и β πN -амплитуды рассеяния.

В работе [83] продемонстрировано, что использование потенциала (2.6) вместо (2.5) почти не сказывается на качестве расчета дифференциальных сечений упругого рассеяния.

В расчетах используется ядерная плотность $\rho(r)$ в виде симметризованной ферми-функции

$$\rho(r) = \rho_{SF}(r) = \rho_0 \frac{\sinh(R/a)}{\cosh(R/a) + \cosh(r/a)} \quad (2.10)$$

с параметрами радиуса R и диффузности a . Значения этих параметров приведены в табл. 2.1. Фактор

$$\rho_0 = \frac{3A}{4\pi R^3} \left[1 + \left(\frac{\pi a}{R} \right)^2 \right]^{-1} \quad (2.11)$$

вычисляется из условия нормировки $4\pi \int_0^{\infty} \rho(r) r^2 dr = A$. Для ядер среднего и тяжелого веса функция (2.10) практически совпадает по форме со стандартным ферми-распределением. Преимущество симметризованной ферми-функции состоит в том, что с ее помощью можно параметризовать также функции распределения плотности легких ядер, а также в том, что

Таблица 2.1. Параметры радиуса R и диффузности a

Мишень	A	Z	R , фм	a , фм	Ссылка	ρ_0 , фм $^{-3}$	$\langle r^2 \rangle^{1/2}$, фм	r_c , фм
^9Be	9	4	1,736	0,85	[105]	0,108	3,437	2,219
^{28}Si	28	14	2,98	0,533	[105]	0,192	3,042	1,293
^{28}Si	28	14	3,134	0,477	[85]	0,177	3,006	1,278
^{40}Ca	40	20	3,787	0,485	[105]	0,151	3,443	1,300
^{40}Ca	40	20	3,593	0,493	[85]	0,174	3,332	1,258
^{58}Ni	58	28	4,08	0,515	[97]	0,176	3,694	1,232
^{58}Ni	58	28	4,2	0,475	[88]	0,166	3,702	1,234
^{208}Pb	208	82	6,654	0,475	[105]	0,160	5,448	1,187

«Ссылка» = источник данных для R и a ; ρ_0 рассчитано в соответствии с R и a по формуле (2.11); $\langle r^2 \rangle^{1/2}$ — среднеквадратичный радиус ядра, рассчитанный в соответствии с R и a ; r_c — зарядовый радиус равномерно заряженной сферы с тем же среднеквадратичным радиусом

для нее формфактор известен в явном виде [81]:

$$\rho_{SF}(q) = -\rho_0 \frac{4\pi^2 a R}{q} \frac{\cos qR}{\sinh(\pi a q)} \left[1 - \left(\frac{\pi a}{R} \right) \coth(\pi a q) \tan(qR) \right]. \quad (2.12)$$

Релятивистское волновое уравнение. При расчете дифференциальных сечений упругого рассеяния используется подход, аналогичный разработанному в работе [84] для рассеяния K -мезонов ядрами. Для этого рассчитывается микроскопический оптический потенциал (2.6) вместе с кулоновским потенциалом равномерно заряженной сферы радиуса $R_c = r_c A^{1/3}$. Значения параметра r_c рассчитаны в соответствии со взятыми из литературы параметрами функции распределения протонов в ядре (см. табл. 2.1). Затем на основе полученного из уравнения Клейна — Гордона — Фока релятивистского волнового уравнения Шредингера

$$(\Delta + k^2) \psi(\vec{r}) = 2\bar{\mu} [U^H(r) + U_c(r)] \psi(\vec{r}), \quad (2.13)$$

где $\bar{\mu} = EM_A/(E + M_A)$ — релятивистская приведенная масса пиона в системе центра масс, E — полная энергия пиона, k — импульс, рассчитывается дифференциальное сечение упругого рассеяния с помощью программы DWUCK4 [100].

Подгонка параметров πN -амплитуды. Подгонка осуществляется минимизацией среднеквадратичного отклонения логарифма рассчитанного дифференциального сечения от соответствующих экспериментальных данных. Величина таких отклонений зависит от варьируемых параметров σ ,

Таблица 2.2. Параметры для расчетов оптического потенциала и дифференциальных сечений упругого рассеяния π -мезонов на ядрах

Мишень	T^{lab}	E	$\gamma^{(r)}$	$T_{\text{non}}^{\text{cm}}$	Мишень	T^{lab}	E	$\gamma^{(r)}$	$T_{\text{non}}^{\text{cm}}$
^{28}Si	130	267,56	1,908	187,68	^{40}Ca	130	268,16	1,915	188,53
^{58}Ni	130	268,59	1,920	189,15	^{208}Pb	130	269,30	1,928	190,15
^{28}Si	162	298,88	2,128	251,56	^{40}Ca	162	299,68	2,138	252,88
^{58}Ni	162	300,26	2,145	253,84	^{208}Pb	162	301,20	2,156	255,41
^{28}Si	180	316,46	2,252	290,53	^{40}Ca	180	317,38	2,263	292,16
^{58}Ni	180	318,05	2,271	293,36	^{208}Pb	180	319,14	2,285	295,31
^{28}Si	226	361,28	2,567	399,94	^{40}Ca	226	362,55	2,582	402,60
^{58}Ni	226	363,48	2,594	405,02	^{208}Pb	226	364,98	2,612	407,73
^{28}Si	230	365,17	2,594	410,12	^{40}Ca	230	366,47	2,610	412,89
^{58}Ni	230	367,43	2,622	414,92	^{208}Pb	230	368,97	2,640	418,22
^9Be	291	412,19	2,92	541,75	^{40}Ca	291	426,19	3,03	583,11
^{28}Si	291	424,36	3,01	578,43	^{208}Pb	291	429,72	3,07	592,16
^{58}Ni	291	427,54	3,05	586,55					

α и β пион-нуклонной амплитуды (2.3):

$$\chi^2 = f(\sigma, \alpha, \beta) = WRSS^{\text{lg}} = \sum_i \frac{(y_{i,\text{lg}} - \hat{y}_{i,\text{lg}}(\sigma, \alpha, \beta))^2}{s_{i,\text{lg}}^2}, \quad (2.14)$$

где $y_{i,\text{lg}} = \log\left(\frac{d\sigma}{d\Omega}\right)_i$ — экспериментальные значения логарифма дифференциальных сечений, а $\hat{y}_{i,\text{lg}} = \log\left[\frac{d\sigma}{d\Omega}(\sigma, \alpha, \beta)\right]_i$ — расчетные значения логарифма дифференциальных сечений, $s_{i,\text{lg}}^2$ — квадраты ошибок логарифмов экспериментальных значений.

Подгоняемыми значениями являются σ , α и β . Входными параметрами для расчетов также являются: кинетическая энергия налетающего π -мезона T^{lab} ; полная энергия пиона в систем ц.м. E ; фактор релятивизации $\gamma^{(r)} = \gamma \frac{M_A + m_\pi}{M_A + \gamma m_\pi}$, где $\gamma = \frac{E}{m_\pi}$; нерелятивистская кинетическая энергия $T_{\text{non}}^{\text{cm}} = \frac{k^2}{2\mu}$, где $\mu = \frac{M_A m_\pi}{M_A + m_\pi}$; атомная масса A и заряд Z ядра-мишени; радиус R и диффузность a ферми-распределения плотности точечных нуклонов в ядре. Значения этих параметров представлены в табл. 2.1 и 2.2.

2.2. Описание комплекса программ

Процесс расчета параметров МОП состоит из нескольких этапов и осуществляется по следующей схеме (рис. 2.1):

Скрипт «face4data» подготавливает все необходимые файлы для дальнейшей работы: вставляет в шаблоны значения параметров пионов и ядер, которые не будут варьироваться, копирует все необходимые

файлы в отдельную директорию.

Затем «main.xml» запускает минимизатор АДЭ, в котором параметры σ , α и β варьируются и расчет невязки (2.14) осуществляется скриптом

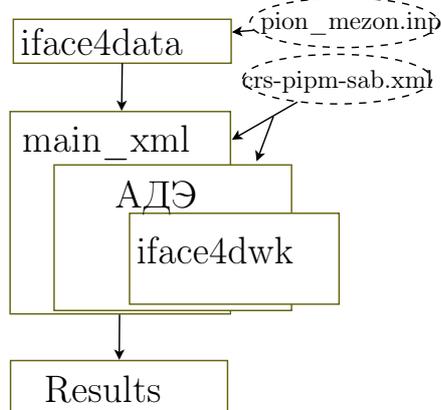


Рис. 2.1. Программный комплекс

«iface4dwk» для каждого расчета целевой функции:

- Пробный вектор (σ, α, β) записывается в файл «pipm-sab.NNNN.inp».
- Запускается модуль для расчета целевой функции «iface4dwk», использующий «pipm-sab.NNNN.inp» в качестве входных данных и выдающий результат в файл «pipm-sab.NNNN.out»).
- Значение целевой функции для пробного вектора (значение невязки) $f(\sigma, \alpha, \beta)$ считывается из файла «pipm-sab.NNNN.out» и передается в алгоритм.
- На стандартный out выдаются промежуточные результаты (через заданное число шагов): координаты минимума, значение в минимуме — и окончательные результаты расчетов (если получены): координаты минимума, значение в минимуме и оценки для доверительных интервалов.

Многократно запускаемая из АДЭ процедура «iface4dwk» осуществляет расчет целевой функции, выполняя последовательность действий, представленную на рис. 2.2:

- Получает варьируемые параметры (σ, α, β) из переданного алгоритмом файла и формирует входной файл «heapot_piplus.inp» для программы расчета МОП «heapot_piplus» [84] — модифицированной программы hea_total.f из библиотеки программ ОИЯИ (<http://www.jinr.ru/programs/jinrlib/hea>).
- Запускает программу «heapot_piplus», которая рассчитывает с заданными σ , α , β , A , Z , R , a и E микроскопический оптический потенциал МОП (σ, α, β) (вещественную и мнимую части) и формирует его в виде матриц $v5$ и $w5$.

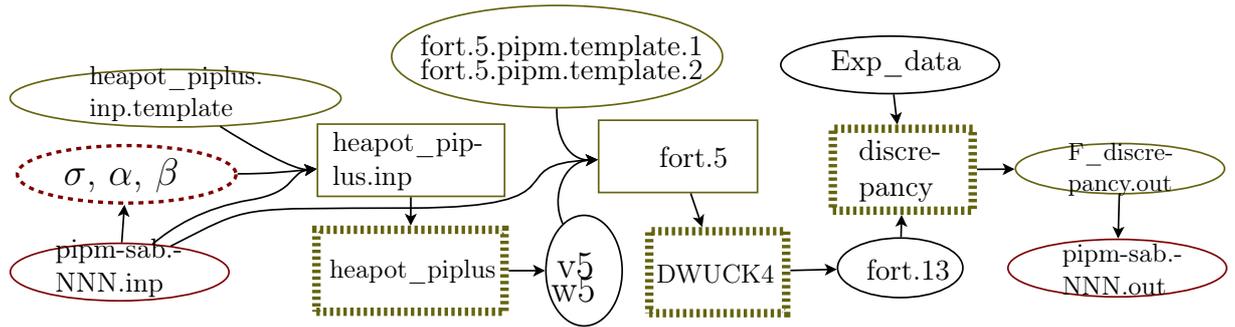


Рис. 2.2. Последовательность действий, осуществляемая скриптом «iface4dwb». «heapot_piplus» — программа для расчета МОП [84]; «DWUCK4» [100] — программа, рассчитывающая дифференциальное сечение упругого рассеяния на основе релятивизированного уравнения Шредингера (2.13); «discrepancy» — модуль для расчета невязки (2.14)

- Автоматически формируется стандартизованный входной файл «fort.5» для модифицированной программы DWUCK4 [100] из «fort.5.pipm.template.1» и «fort.5.pipm.template.2» и матриц v_5 и w_5 .
- Запускает модифицированную программу DWUCK4, которая реализует численное решение задачи рассеяния для уравнения Шредингера (2.13) на основе разложения волновых функций по сферическим гармоникам, вычисляет дифференциальные сечения упругого рассеяния и записывает результаты в файл.
- Запускает программу «discrepancy» (C++), которая считывает теоретические и экспериментальные значения сечений из файлов и рассчитывает взвешенную невязку (2.14) между теоретическими и экспериментально измеренными дифференциальными сечениями упругого рассеяния.
- Представляет результат в виде, удобном для дальнейшего использования в модуле «Алгоритм АДЭ».

Значения параметров σ , α и β , обеспечивающие минимум χ^2 :

$$(\sigma, \alpha, \beta) = \text{Argmin } f(\sigma, \alpha, \beta),$$

определяются с использованием алгоритма АДЭ.

Таким образом, разработанный комплекс программ помимо стандартного кода DWUCK4 и созданной ранее FORTRAN-программы «heapot_piplus» состоит из следующих компонент.

- 1 Библиотека метода АДЭ (ОМР-вариант) — модуль, написанный на языке C++. Заголовочные файлы использованных частей библиотеки: «ade_optimizer.h», «ade_settings.h», «parameter_list.h», «problem_definition.h».
- 2 Программа расчета взвешенной невязки «discrepancy» — модуль, написанный на языке C++.
- 3 Скрипт «iface4data», готовящий исходные данные к расчетам.
- 4 Скрипт «iface4dwk», для связки программ, осуществляющих различные этапы расчета целевой функции, в единую последовательность.
- 5 (опция) Скрипт, запускающий серию расчетов (для оценок для различных энергий и ядер).

Оценка ускорения АДЭ при параллельных вычислениях

Вычисления производились как в последовательном, так и в параллельном режимах. Распараллеливание вычислений организовано на основе технологии OpenMP³ для многопоточных приложений на многопроцессорных системах с общей памятью. Для распараллеливания использована так называемая модель ведущий/ведомый (см. рис. 1.14 в разделе 1.5.2).

Параметры, с которыми проводился расчет. Расчеты проводились со следующими параметрами АДЭ: $F = 0,8$, $C_r = 0,8$, и с первоначальной областью поиска глобального минимума: $\sigma \in [0, 20]$ фм², $\alpha \in [-2, 2]$, $\beta \in [0, 2]$ фм⁻². Требуемая точность локализации минимума (критерий выхода) задается как отношение разброса значений целевой функции (2.14) к максимуму модуля ее значения (см. критерий стагнации (1.17)). Выход из итерационного процесса происходит, когда упомянутая величина не превышает 10^{-4} :

$$e_f = \frac{\max_{i=0, \dots, N_p-1} \{f_i\} - \min_{i=0, \dots, N_p-1} \{f_i\}}{\max_{i=0, \dots, N_p-1} \{|f_i|\}} < \varepsilon_f = 10^{-4}. \quad (2.15)$$

Полученное ускорение при параллельных вычислениях. При вышперечисленных параметрах для расчетов рассеяния π^- на ядрах ^{28}Si при энергии $T^{\text{lab}} = 130$ МэВ на четырехъядерном процессоре Intel(R) Core(TM) i7-2630QM CPU@2.00GHz при четырехпоточном вычислении стратегия

³OpenMP (Open Multi-Processing) — открытый стандарт для распараллеливания программ на языках Си, Си++ и Фортран. www.openmp.org — официальный сайт OpenMP

DE/rand/rand/1/bin с $N_p = 20$ требует 2300 ± 250 вычислений целевой функции (2.14), что занимает 137 ± 18 с. Эта же стратегия в последовательном режиме вычислений для нахождения минимума с той же точностью требует 1875 ± 230 расчетов целевой функции, что занимает 231 ± 29 с. Таким образом, ускорение расчетов при использовании четырех процессоров по сравнению с однопоточным режимом составляет 1,7.

Потери реального времени при распараллеливании связаны с конкурированием параллельных процессов при работе с жестким диском. Действительно, обмен данными между отдельными модулями данной программы осуществляется посредством записи на жесткий диск — чтения с жесткого диска. Среднее время, затрачиваемое на одно вычисление в последовательной моде составляет $\tau_1^{\text{thread}} \approx 0,123$ с, а в параллельной — $\tau_1^{4 \text{ thread}} \approx 0,060$ с, что соответствует совокупному времени работы всех четырех ядер процессора 0,240 с.

Если бы обмен данными с жестким диском не ограничивал скорости расчетов, то ускорение оптимизационного алгоритма составило бы $speedup = \frac{4 * N_{thr}}{N_{1th}} = 3,3$. Отличие ускорения от количества задействованных процессоров обусловлено большим количеством вычислений целевой функции в параллельном режиме.

Таким образом, конкретная реализация решения реальных физических задач накладывает свои ограничения на скорость работы в параллельном режиме программного комплекса, использующего АДЭ. Так, при расчетах рассеяния пионов на ядрах была использована ОМР-версия параллельной реализации АДЭ, вычисления проводились на четырехъядерном персональном компьютере. Однако процесс расчета целевой функции производился программным комплексом, использующим ряд программ, которые, в свою очередь, обмениваются данными посредством записи на диск — чтения с диска. Это приводит к тому, что скорость обращения к диску становится ключевым фактором ограничения скорости при работе оптимизатора в параллельном режиме. Из-за этого реальное (астрономического) время расчетов сократилось всего в 1,7 раза, в то время как сам алгоритм мог обеспечить ускорение около 3,3.

2.3. Обсуждение численных результатов

Сравнение рассчитанных дифференциальных сечений упругого пион-ядерного рассеяния с экспериментальными данными. В работе [83] показано, что рассчитанное сечение с использованием потенциала (2.6) и подгонкой трех параметров (2.7) пион-нуклонной амплитуды дает результаты, согласующиеся с результатами расчета для потенциала (2.5) с шестью параметрами, рассеяния пионов отдельно на протонах и нейтронах [2]. Оба расчета близки друг к другу и отражают основные черты пион-нуклонного рассеяния. Однако, для того, чтобы добиться приемлемого согласия с экспериментальными данными с достаточно низкими значениями χ^2 отклонений, необходимо делать подгонку этих параметров, рассматривая πN -амплитуду как амплитуду рассеяния пионов на нуклонах, находящихся в ядерной среде.

Рис. 2.3 демонстрирует результаты подгонки параметров πN -амплитуд для дифференциальных сечений π^- -рассеяния (а) и π^+ -рассеяния (б) на ядрах ^{28}Si , ^{54}Ni , ^{208}Pb при $T^{lab} = 291$ МэВ по экспериментальным данным из [93]. Качество описания экспериментальных данных существенно улучшено по сравнению с расчетами, в которых использовались параметры рассеяния на свободных нуклонах [2]. В дальнейшем при сравнении с другими экспериментальными данными мы будем использовать только потенциал (2.6) и варьировать в подгонках три указанных параметра (2.7).

На рис. 2.4 и 2.6 приведены результаты подгонки к данным работ [106] и [95] для рассеяния π -мезонов на ядрах ^{28}Si и ^{40}Ca при энергиях $T^{lab} = 180$ МэВ и 130 МэВ.

На рис. 2.5 показаны в сравнении с подогнанными к ним сечениями экспериментальные данные из работы [104] сечения рассеяния π^\pm -мезонов на ядрах ^{28}Si , ^{58}Ni , ^{208}Pb при энергии $T^{lab} = 162$ МэВ. Следует отметить, что этот результат находится в согласии с приведенными в той же работе расчетами с использованием типичного в таких задачах шести- и более параметрического потенциала Кисслингера.

В табл. 2.3 приведены значения полученных таким образом параметров πN -амплитуды в ядерной среде («in-medium» параметров) и соответствующие статистические ошибки, обусловленные указанными в этих работах ошибками экспериментальных данных, оцененные исходя из условия

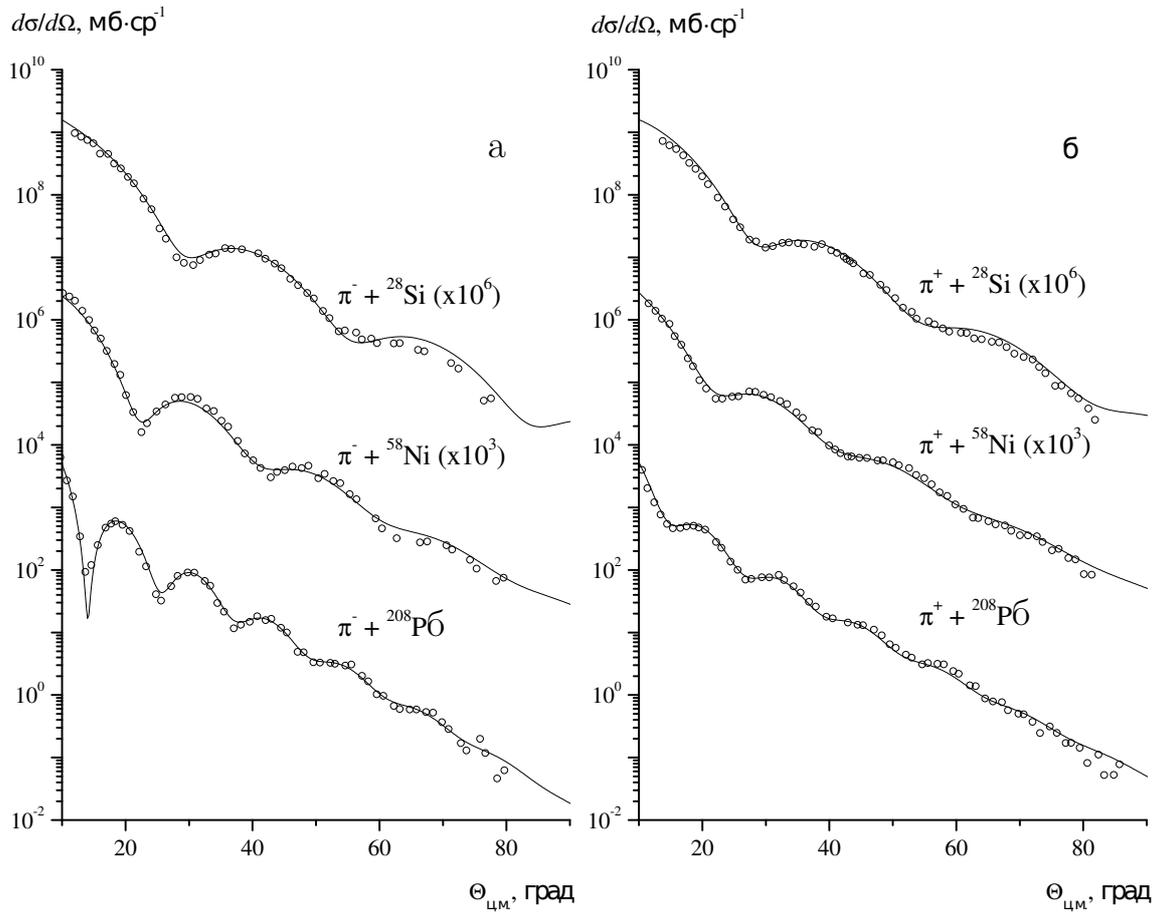


Рис. 2.3. Результаты расчета дифференциальных сечений π^- -рассеяния (а) и π^+ -рассеяния (б) на ядрах ^{28}Si , ^{54}Ni , ^{208}Pb при $T^{\text{lab}} = 291$ МэВ, рассчитанные с параметрами πN -амплитуды, подогнанными к экспериментальным данным из [93], и представленными в табл. 2.3 [2]

$\Delta\chi^2 = 1$, и χ^2 -отклонения в расчете на одну степень свободы.

По результатам приведенных сравнений можно прийти к заключению, что обсуждаемая трехпараметрическая модель при своей относительной простоте и небольшом числе подгоняемых параметров в целом обеспечивает согласие с экспериментальными данными, сопоставимое с результатами, полученными в рамках других подходов с большим числом параметров. Значения⁴ χ^2/k лежат главным образом в пределах от двух до шести, хотя, например, для рассеяния $\pi^+ + ^{208}\text{Pb}$ при $T^{\text{lab}} = 162$ МэВ, $\pi^\pm + ^{28}\text{Si}$ при $T^{\text{lab}} = 180$ МэВ они превышают 10. Относительно расчетов рассеяния на ^{208}Pb необходимо отметить, что в сформулированном подходе использовалось весьма удобное выражение (2.6) для оптического потенциала, которое, однако, как уже было отмечено, справедливо для рассеяния π -мезонов на ядрах с близкими числами протонов и нейтронов. И поскольку у ядра ^{208}Pb

⁴ $k = n - l$ — число степеней свободы, n — количество экспериментальных точек, l — количество переменных.

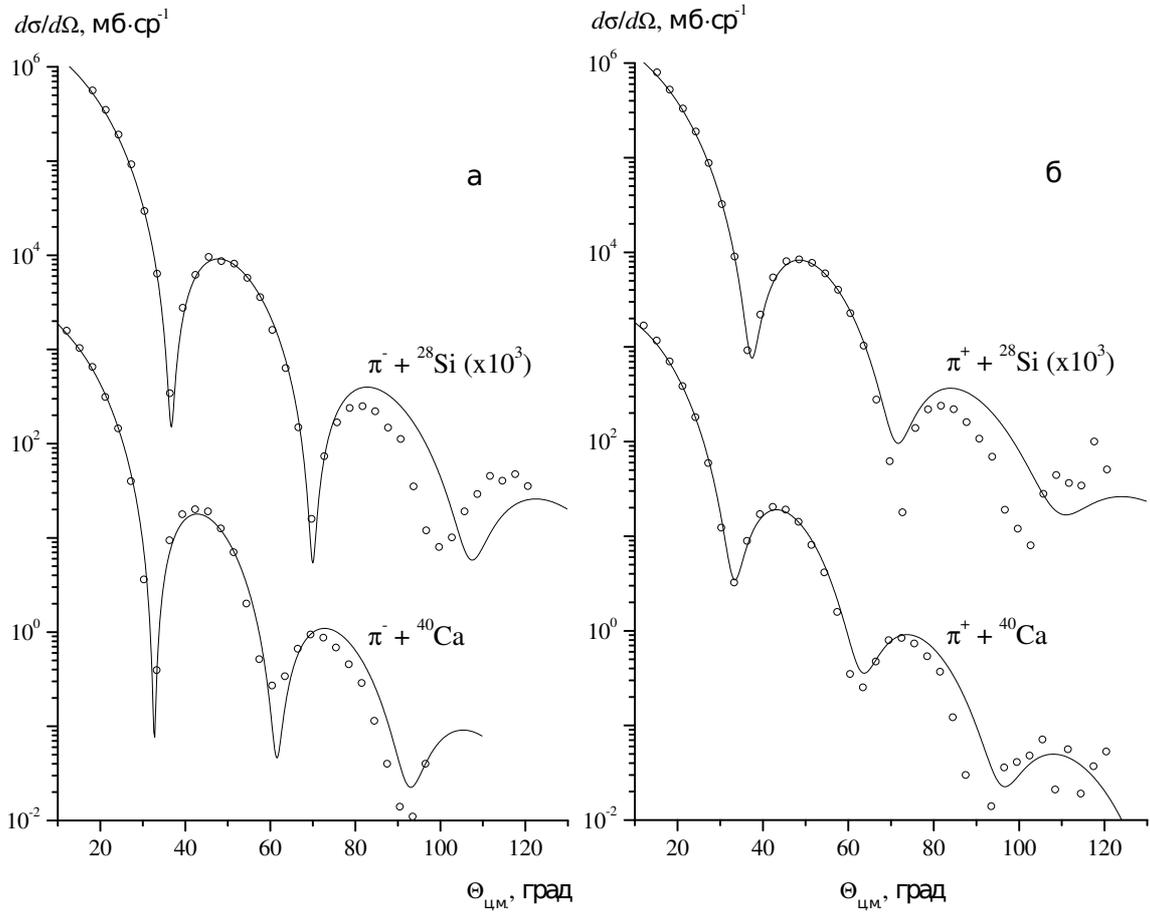


Рис. 2.4. То же, что на рис.2.3, но для ядер ^{28}Si , ^{40}Ca при $T^{lab} = 180$ МэВ. [2]. Экспериментальные данные — из [106] и [95]

имеется большой избыток из 44 нейтронов, то здесь требуется, вообще говоря, специальное рассмотрение.

Второе замечание состоит в том, что используемое выражение (2.4) для амплитуды πN -рассеяния в виде гауссовой функции справедливо в основном для области сравнительно больших энергий и малых углов рассеяния. Действительно, экспериментом установленное поведение дифференциального сечения πN -рассеяния характеризуется ростом при углах больше 90° (см., например, [108]), в то время как подгоняемое сечение пион-нуклонного рассеяния $\sim \exp(-\beta q^2)$ продолжает убывать в этой области углов. Для иллюстрации последствий такого несоответствия нами сделаны расчеты рассеяния пионов с отбрасыванием экспериментальных точек при углах $\Theta > 80^\circ$. На рис.2.7 эти результаты представлены для рассеяния $\pi^+ + ^{28}\text{Si}$ при $T = 180$ МэВ в сравнении с экспериментальными данными из [104, 106]. Сплошная кривая соответствует оптимальному значению χ^2 для полного набора экспериментальных данных (36 точек). Штриховая кривая построена для набора из 31 точки с отбрасыванием последних пя-

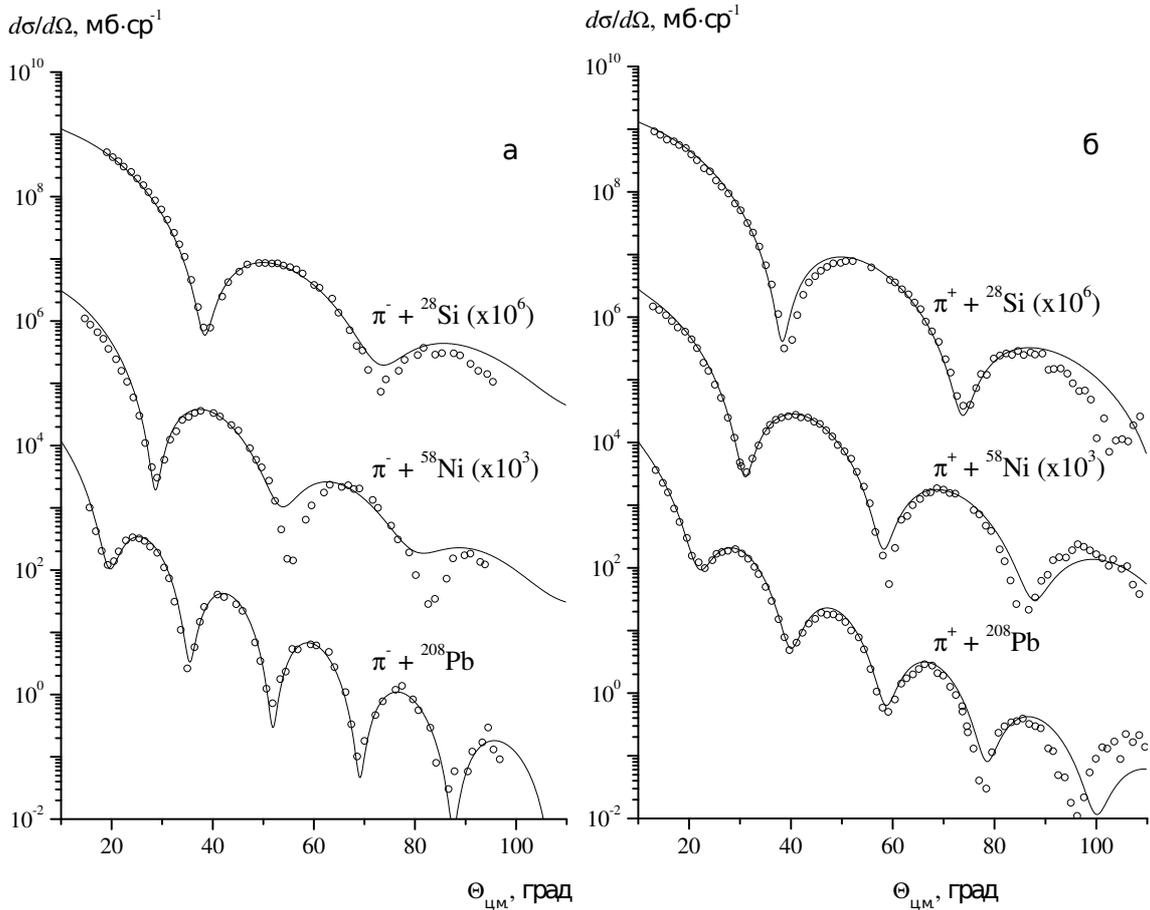


Рис. 2.5. То же, что на рис.2.3, но для $T^{lab} = 162$ МэВ. [2]. Экспериментальные данные — из [104]

ти измерений. При этом качество подгонки, оцениваемое значением χ^2/k , возрастает более чем в два раза.

Рассчитанные по редуцированному набору экспериментальных данных (с отбрасыванием измерений при углах $> 80^\circ$) параметры представлены в табл. 2.4. Сравнение данных в табл. 2.3 и табл. 2.4 показывает: в большинстве случаев отбрасывание измерений при больших углах приводит к улучшению согласия с остальными экспериментальными точками, и при этом, что важно, значения параметров подгонки меняются несущественно.

Поведение «in-medium» параметров πN -амплитуды. Для изучения влияния ядерной среды на амплитуду пион-нуклонного рассеяния нами проведено [2] сравнение усредненных («эффективных») значений параметров $X^{eff} = (X_{\pi^+} + X_{\pi^-})/2$ (где $X = \sigma, \alpha, \beta$), полученных из данных по рассеянию пионов в ядерной среде, с данными рассеяния на свободных нуклонах, представленными в работах [89, 103].

Эти результаты аккумулированы на рис. 2.8, где показано, как зависят от энергии усредненные указанным образом «in-medium» параметры

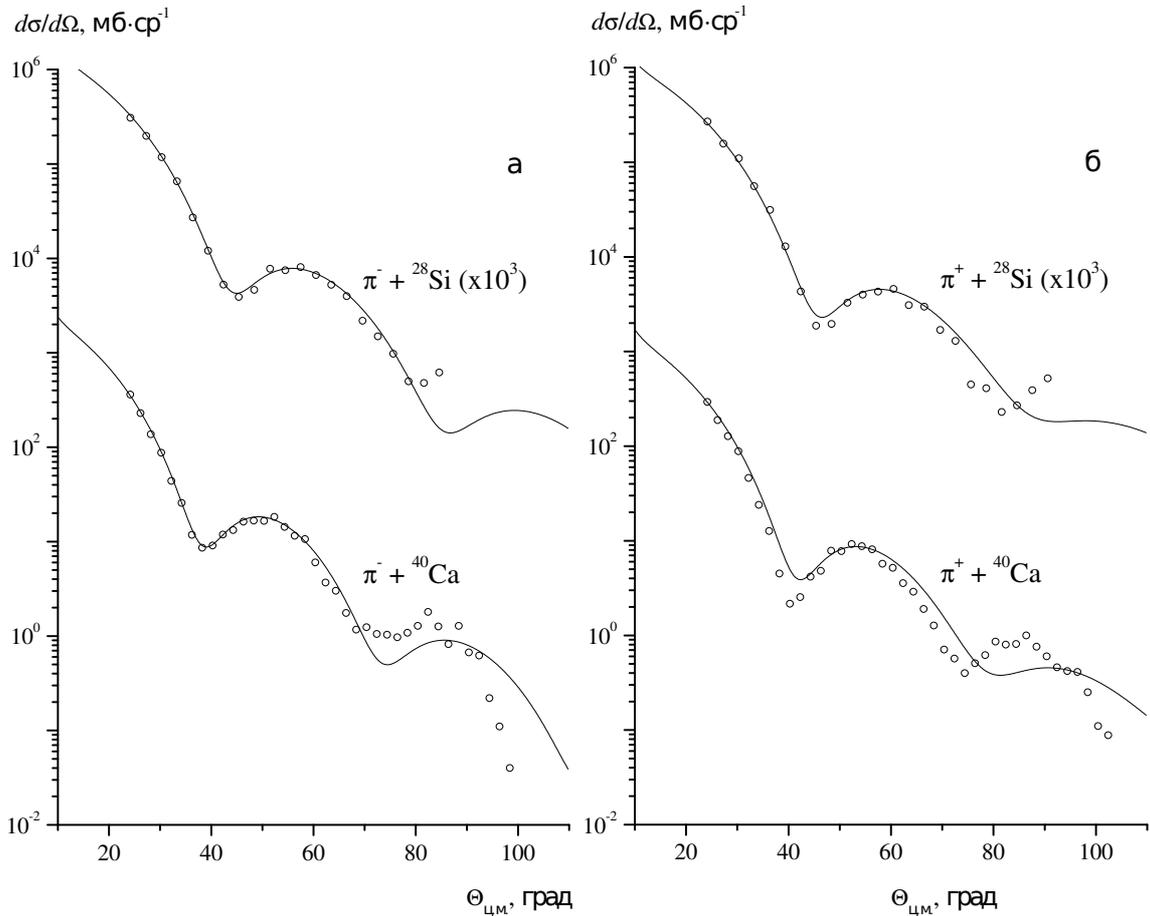


Рис. 2.6. То же, что рис.2.3, но для ядер ^{28}Si , ^{40}Ca при $T^{lab} = 130$ МэВ. [2]. Экспериментальные данные — из [106] и [95]

амплитуды πN -рассеяния. Во-первых, следует отметить, что полное сечение πN -рассеяния в среде σ^{eff} сохраняет общие черты резонансной зависимости в области (3,3)-резонанса, присущей рассеянию пионов на свободных нуклонах. В то же время, абсолютное значение сечения оказывается примерно на 30% ниже свободного, что может означать ослабление πN -взаимодействия в ядерной среде. Причиной тому может служить, в первую очередь, эффект Паули-блокировки нуклонов, находящихся на оболочках в потенциальной яме.

Что касается поведения параметра α , то видно, что по мере перехода энергии через точку максимума (3,3)-резонанса вещественная часть πN -амплитуды рассеяния в ядерной среде при 0° продолжает оставаться положительной, в то время как при рассеянии на свободных нуклонах эта часть амплитуды уходит в отрицательную область. На языке потенциального рассеяния это может означать, что в этой области пионы в ядерной среде испытывают преломление, которое ослабевает, начиная с энергии ~ 230 МэВ. В случае рассеяния на свободных нуклонах такая смена режи-

Таблица 2.3. Параметры подгонки «in-medium» πN -амплитуды для ряда пион-ядерных реакций при разных энергиях. k — число степеней свободы [2]

Реакция	T^{lab}	χ^2/k	σ	α	β
$\pi^- + {}^{28}\text{Si}$	130	2,4	$7,02 \pm 0,18$	$0,88 \pm 0,02$	$0,82 \pm 0,05$
$\pi^+ + {}^{28}\text{Si}$		5,5	$9,60 \pm 0,13$	$0,03 \pm 0,02$	$0,83 \pm 0,04$
$\pi^- + {}^{40}\text{Ca}$		4,0	$6,99 \pm 0,08$	$0,90 \pm 0,01$	$0,86 \pm 0,03$
$\pi^+ + {}^{40}\text{Ca}$		13,0	$8,62 \pm 0,09$	$0,09 \pm 0,01$	$0,74 \pm 0,02$
$\pi^- + {}^{28}\text{Si}$	162	4,0	$8,93 \pm 0,09$	$0,58 \pm 0,01$	$0,47 \pm 0,02$
$\pi^+ + {}^{28}\text{Si}$		4,01	$10,24 \pm 0,05$	$0,36 \pm 0,01$	$0,59 \pm 0,01$
$\pi^- + {}^{58}\text{Ni}$		10,12	$10,83 \pm 0,11$	$-0,16 \pm 0,01$	$1,11 \pm 0,02$
$\pi^+ + {}^{58}\text{Ni}$		6,5	$9,28 \pm 0,04$	$+0,44 \pm 0,01$	$0,75 \pm 0,01$
$\pi^- + {}^{208}\text{Pb}$		3,7	$9,66 \pm 0,10$	$0,35 \pm 0,01$	$1,03 \pm 0,02$
$\pi^+ + {}^{208}\text{Pb}$		10,5	$6,65 \pm 0,03$	$0,61 \pm 0,01$	$1,08 \pm 0,01$
$\pi^- + {}^{28}\text{Si}$	180	10,24	$10,11 \pm 0,05$	$0,33 \pm 0,01$	$0,24 \pm 0,01$
$\pi^+ + {}^{28}\text{Si}$		10,83	$10,36 \pm 0,06$	$0,29 \pm 0,01$	$0,31 \pm 0,01$
$\pi^- + {}^{40}\text{Ca}$		2,8	$9,65 \pm 0,10$	$0,23 \pm 0,02$	$0,28 \pm 0,01$
$\pi^+ + {}^{40}\text{Ca}$		3,35	$5,78 \pm 0,07$	$1,08 \pm 0,02$	$0,69 \pm 0,02$
$\pi^- + {}^{28}\text{Si}$	226	17,8	$9,48 \pm 0,06$	$-0,21 \pm 0,01$	$0,143 \pm 0,01$
$\pi^+ + {}^{28}\text{Si}$		26,9	$5,87 \pm 0,005$	$1,08 \pm 0,01$	$0,420 \pm 0,01$
$\pi^- + {}^{40}\text{Ca}$	230	5,98	$5,28 \pm 0,08$	$0,80 \pm 0,01$	$0,240 \pm 0,02$
$\pi^+ + {}^{40}\text{Ca}$		7,91	$9,14 \pm 0,19$	$-0,11 \pm 0,02$	$0,246 \pm 0,03$
$\pi^- + {}^{28}\text{Si}$	291	6,08	$5,14 \pm 0,11$	$-0,80 \pm 0,02$	$0,16 \pm 0,02$
$\pi^+ + {}^{28}\text{Si}$		4,9	$5,36 \pm 0,14$	$-0,79 \pm 0,02$	$0,368 \pm 0,02$
$\pi^- + {}^{58}\text{Ni}$		3,8	$4,79 \pm 0,09$	$-0,85 \pm 0,02$	$0,281 \pm 0,01$
$\pi^+ + {}^{58}\text{Ni}$		2,6	$5,56 \pm 0,09$	$-0,67 \pm 0,02$	$0,365 \pm 0,01$
$\pi^- + {}^{208}\text{Pb}$		4,1	$4,46 \pm 0,07$	$-1,07 \pm 0,02$	$0,673 \pm 0,02$
$\pi^+ + {}^{208}\text{Pb}$		2,7	$5,56 \pm 0,14$	$-0,46 \pm 0,01$	$0,588 \pm 0,02$

ма от положительных к отрицательным значениям α происходит в районе максимума (3,3)-резонанса при энергии 170 МэВ.

Интересно понять поведение β , параметра наклона πN -амплитуды, который, как видно из рис. 2.8, в ядерной среде оказывается примерно в два раза меньше, чем в случае рассеяния на свободных нуклонах. Если использовать для качественных оценок борновское приближение, то потенциал, порождающий амплитуду $\sim \exp(-\beta q^2/2)$, имел бы тоже гауссову форму вида $\sim \exp(-r^2/2\beta)$, что соответствует среднеквадратичному радиусу потенциала πN -взаимодействия $\sim \sqrt{2\beta}$ фм. Таким образом, уменьшение β в два раза означает уменьшение радиуса взаимодействия пиона с ядерным нуклоном примерно на 30%.

Таким образом, резюмируя полученные результаты, отметим следу-

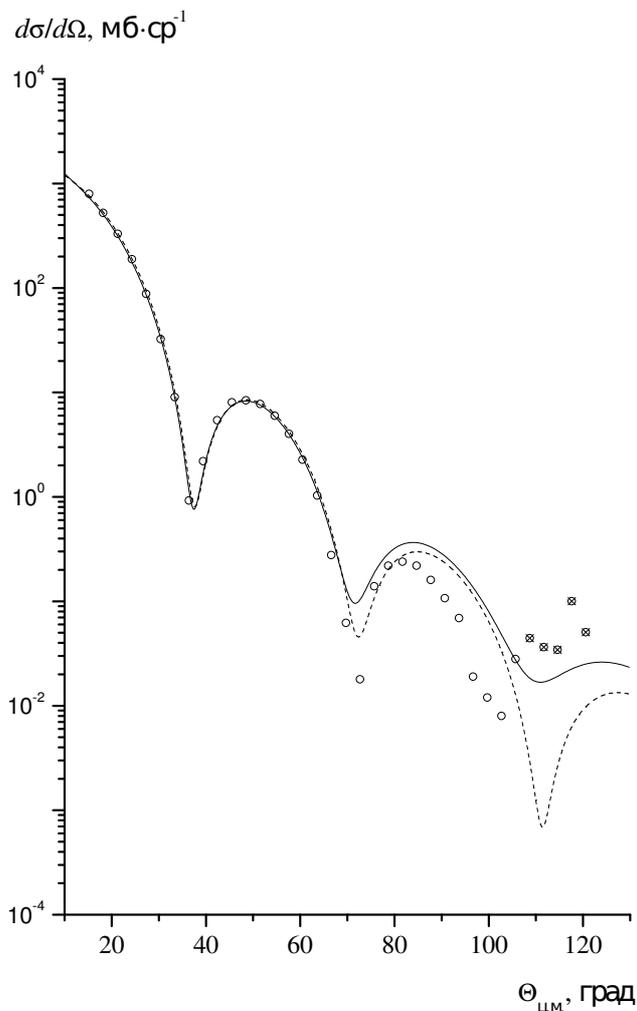


Рис. 2.7. Подгонка к экспериментальным данным $\pi^+ + {}^{28}\text{Si}$ при $T^{\text{lab}} = 180$ МэВ из [104] с учетом всех экспериментальных данных (сплошная кривая, параметры подгонки $\sigma = 10.36$, $\alpha = 0.29$, $\beta = 0.31$, $\chi^2/k = 10.83$) и с отбрасыванием последних пяти точек (штриховая кривая, параметры $\sigma = 7.54$, $\alpha = 0.771$, $\beta = 0.538$, $\chi^2/k = 5.1$) [2]

ющее:

1. Из сравнения с экспериментальными данными пион-ядерного рассеяния при большом наборе ядер-мишеней и при разных энергиях столкновения в области (3,3)-резонанса можно сделать вывод, что предложенная микроскопическая модель пион-ядерного оптического потенциала вполне себя оправдывает. В ее основу положена функция распределения нуклонов в ядре-мишени и трехпараметрическая пион-нуклонная амплитуда рассеяния с гауссовой формой зависимости от переданного импульса πN -амплитуды рассеяния.
2. Ограничение модели областью применения при углах рассеяния пионов $\theta \leq 80^\circ$, в которой обеспечиваются низкие значения χ^2/k -отклонений теоретических сечений рассеяния от экспериментальных, определяется в основном видом амплитуды πN -рассеяния, которая достаточно хорошо описывает данные по дифференциальным сечениям πN -рассеяния только в передней полусфере. Можно ожидать, что использование дру-

Таблица 2.4. Параметры подгонки «in-medium» πN -амплитуды для ряда пион-ядерных реакций при разных энергиях. Число степеней свободы — $k = n - l$. Расчет с отбрасыванием экспериментальных точек при углах более 80° .

Реакция	T^{lab}	χ^2/k	σ	α	β
$\pi^- + {}^{28}\text{Si}$	130	0,88	$7,22 \pm 0,18$	$0,84 \pm 0,02$	$0,78 \pm 0,04$
$\pi^+ + {}^{28}\text{Si}$		2,73	$9,59 \pm 0,11$	$0,11 \pm 0,02$	$0,83 \pm 0,03$
$\pi^- + {}^{40}\text{Ca}$		3,92	$6,86 \pm 0,09$	$0,90 \pm 0,01$	$0,88 \pm 0,03$
$\pi^+ + {}^{40}\text{Ca}$		14,42	$8,21 \pm 0,06$	$0,10 \pm 0,01$	$0,76 \pm 0,02$
$\pi^- + {}^{28}\text{Si}$	162	1,97	$11,63 \pm 0,13$	$0,10 \pm 0,01$	$0,26 \pm 0,02$
$\pi^+ + {}^{28}\text{Si}$		2,13	$10,46 \pm 0,07$	$0,39 \pm 0,01$	$0,51 \pm 0,02$
$\pi^- + {}^{58}\text{Ni}$		12,0	$10,95 \pm 0,13$	$-0,16 \pm 0,01$	$1,10 \pm 0,02$
$\pi^+ + {}^{58}\text{Ni}$		3,72	$8,86 \pm 0,03$	$0,44 \pm 0,01$	$0,81 \pm 0,01$
$\pi^- + {}^{208}\text{Pb}$		3,71	$9,69 \pm 0,10$	$0,34 \pm 0,01$	$1,02 \pm 0,02$
$\pi^+ + {}^{208}\text{Pb}$		1,84	$6,18 \pm 0,03$	$0,58 \pm 0,01$	$1,24 \pm 0,01$
„ $\pi^- + {}^{28}\text{Si}$	180	5,99	$9,33 \pm 0,07$	$0,43 \pm 0,01$	$0,28 \pm 0,01$
$\pi^+ + {}^{28}\text{Si}$		1,26	$7,75 \pm 0,06$	$0,76 \pm 0,01$	$0,49 \pm 0,01$
$\pi^- + {}^{40}\text{Ca}$		3,02	$9,75 \pm 0,12$	$0,21 \pm 0,02$	$0,27 \pm 0,02$
$\pi^+ + {}^{40}\text{Ca}$		2,87	$5,75 \pm 0,08$	$1,09 \pm 0,02$	$0,69 \pm 0,02$
$\pi^- + {}^{28}\text{Si}$	226	16,2	$7,43 \pm 0,06$	$0,60 \pm 0,01$	$0,167 \pm 0,01$
$\pi^+ + {}^{28}\text{Si}$		21,55	$9,99 \pm 0,08$	$-0,13 \pm 0,01$	$0,116 \pm 0,01$
$\pi^- + {}^{40}\text{Ca}$	230	6,37	$5,34 \pm 0,08$	$0,79 \pm 0,02$	$0,230 \pm 0,02$
$\pi^+ + {}^{40}\text{Ca}$		9,59	$9,17 \pm 0,19$	$-0,10 \pm 0,03$	$0,236 \pm 0,03$
$\pi^- + {}^{28}\text{Si}$	291	6,08	$5,14 \pm 0,11$	$-0,80 \pm 0,02$	$0,16 \pm 0,02$
$\pi^+ + {}^{28}\text{Si}$		4,72	$5,39 \pm 0,14$	$-0,79 \pm 0,02$	$0,36 \pm 0,01$
$\pi^- + {}^{58}\text{Ni}$		3,8	$4,79 \pm 0,08$	$-0,85 \pm 0,02$	$0,279 \pm 0,02$
$\pi^+ + {}^{58}\text{Ni}$		2,6	$5,58 \pm 0,13$	$-0,66 \pm 0,02$	$0,354 \pm 0,02$
$\pi^- + {}^{208}\text{Pb}$		4,09	$4,47 \pm 0,06$	$-1,07 \pm 0,02$	$0,672 \pm 0,02$
$\pi^+ + {}^{208}\text{Pb}$		2,7	$5,52 \pm 0,13$	$-0,46 \pm 0,01$	$0,581 \pm 0,02$

гой формы амплитуды с удачной аппроксимацией ее q^2 -асимптотики позволит получать хорошее согласие и при больших углах рассеяния. Также видно, что введение в микропотенциал дополнительных членов, корректирующих его поведение в области ядерной поверхности, может обеспечить более точное описание экспериментальных данных во всем угловом интервале измерений.

3. На базе довольно большой статистики подгонок теоретических сечений к экспериментальным получены усредненные значения параметров пион-нуклонной амплитуды в ядерной среде. Сравнение зависимостей полученных параметров от энергии столкновения с теми, которые имеют место для параметров амплитуды рассеяния на свободных нуклонах, позволяет увидеть основные закономерности в их отличиях друг

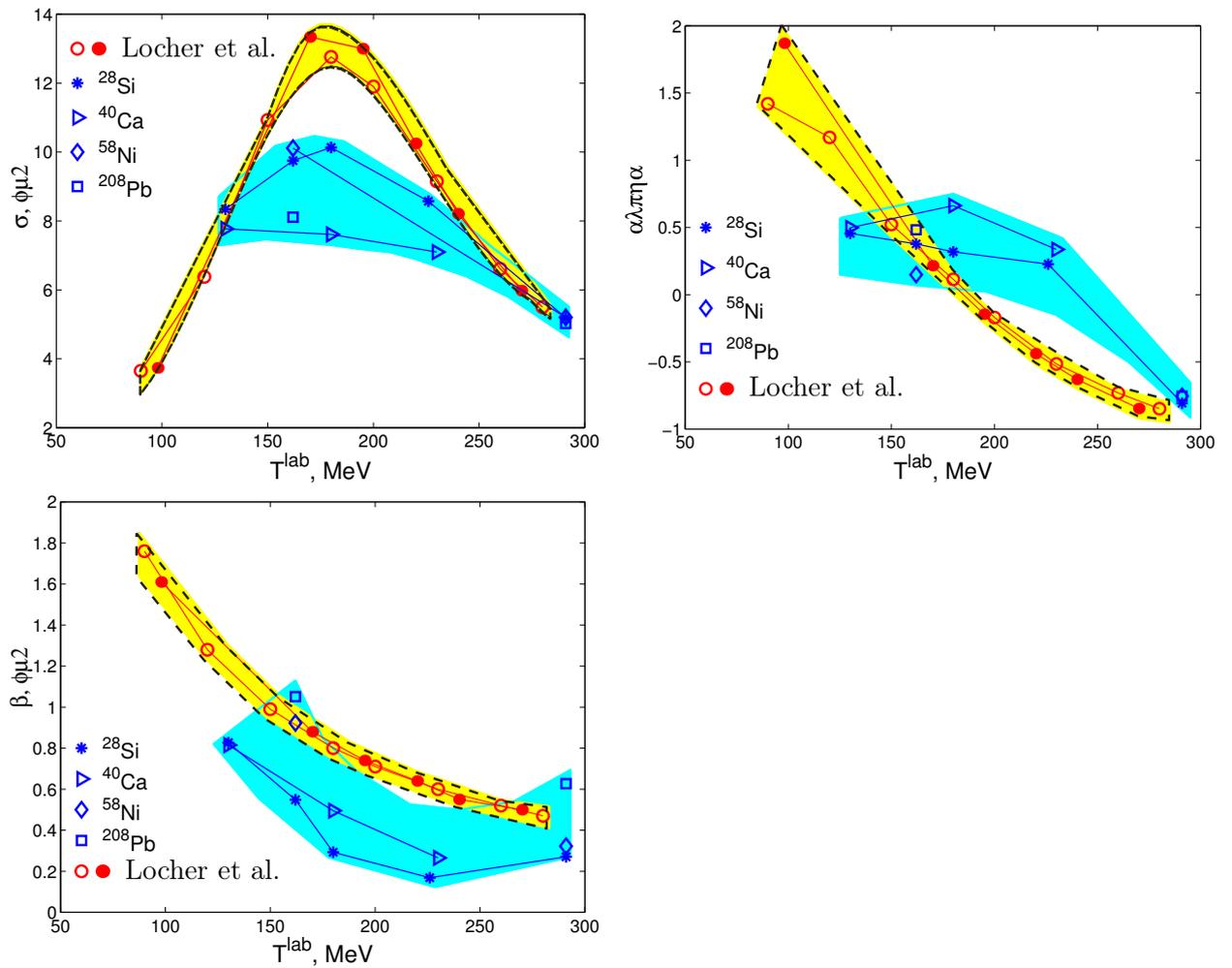


Рис. 2.8. Поведение усредненных параметров $X^{eff} = (X_{\pi^+} + X_{\pi^-})/2$, $X = \sigma, \alpha, \beta$ амплитуды рассеяния π -мезонов на внутриядерных нуклонах (область темного цвета) и аналогичных параметров X^{free} рассеяния на свободных нуклонах (область, обведенная пунктирной линией) [2]

от друга. По сравнению с рассеянием пионов на свободных нуклонах в ядерной среде полное сечение уменьшается, реальная часть амплитуды рассеяния вперед оказывается больше, а наклон кривых углового распределения в упругом рассеянии становится более пологим. Все это дает полезную информацию для последующих теоретических исследований поведения пионов в ядерной среде.

2.4. Результативность АДЭ по сравнению с другими методами оптимизации

Представлены оценки результативности методов АДЭ по сравнению с другими методами. На примере анализа сечения рассеяния π^+ на ядрах кремния ^{28}Si при энергии $T^{lab} = 162$ МэВ оптимизация проводилась мно-

гократно при помощи различных оптимизационных методов: метода симплексов (метод Нелдера — Мида) [53], квазиньютоновского метода (метод Дэвидона — Флетчера — Пауэлла) [30, 31, 37, 38, 55] (Root-реализация программы Minuit [46], методы SIMPLEX и MIGRAD) и различных стратегий методов АДЭ [6, 7, 9].

Параметры, с которыми проводился расчет. Расчеты проводились исходя из того, что допустимые значения параметров лежат в области Ω : $\sigma \in [0, 20]$ фм², $\alpha \in [-2, 2]$, $\beta \in [0, 2]$ фм⁻². Точки для начального приближения равномерно распределены в этой области. Для АДЭ-Р расчеты проводились с параметрами $F = 0,9$, $C_r = 0,9$ (это часто применяемая комбинация при расчетах с помощью ДЭ) и двумя значениями начальной популяции: $N_p^{\min}=10$ — это минимальный рекомендуемый размер популяции; и $N_p^{\min}=30$ — это чуть меньше, чем «оптимальный» размер популяции ДЭ для 10-мерной функции Розенброка (см. 1.1.3, табл. 1.1). Критерий рестарта (1.14) $\Delta x_j = 10^{-12}$. Требуемая точность локализации минимума (критерий выхода) (1.15) $\Delta f = 10^{-2}$. Настройки SIMPLEX и MIGRAD — по умолчанию. Расчет каждым из методов повторялся $N = 100$ раз.

Найденные решения. Глобальный минимум задачи найден в точке: $\theta^* = (\sigma^*, \alpha^*, \beta^*) = (10,20 \pm 0,05; 0,362 \pm 0,007; 0,60 \pm 0,01)$ (ошибки оценены исходя из условия $\Delta\chi^2 = 1$). Значение целевой функции в этой точке: $f^*(\theta^*) = 333.15$. Ближайший по глубине локальный минимум находится в точке $\theta^\dagger = (\sigma^\dagger, \alpha^\dagger, \beta^\dagger) = (8,48; 0,71; 0,708)$. Значение целевой функции в этой точке: $f^\dagger(\theta^\dagger) = 492.14$. Количество экспериментальных точек $n = 76$.

Сравнение результативности метода АДЭ, метода симплексов (Нелдера — Мида) и квазиньютоновского метода Дэвидона — Флетчера — Пауэлла. В табл. 2.5 приведены обобщенные результаты, полученные каждым из методов.

Таким образом, при решении данной задачи метод АДЭ предпочтительнее классических методов поиска по трем основным причинам:

- Во-первых, метод АДЭ-Р является робастным методом поиска. С вероятностью, близкой к единице, он завершает расчеты в одном из минимумов задачи, в то время как методы SIMPLEX и MIGRAD закончили поиск в одном из минимумов лишь с вероятностью $49 \pm 5\%$ и $7 \pm 2,6\%$ соответственно.

Таблица 2.5. Результаты минимизации для методов SIMPLEX, MIGRAD и АДЭ и общее количество попыток минимизации.

	MIG- RAD	SIMP- LEX	АДЭ-Р $N_p^{\min}=10$	АДЭ, АДЭ-Р, $N_p^{\min}=30$
Завершившиеся вблизи глобального минимума θ^* : $f < 334.2$	5 †	20	77	100
Завершившиеся в окрестности глобального минимума θ^* : $f \in (368, 503)$	3	0	0	0
Завершившиеся вблизи локального минимума θ^\dagger : $f \in (492.13, 493, 2)$	2 †	29	23	0
Завершившиеся в окрестности локального минимума θ^\dagger : $f \in (503, 741)$	8	0	0	0
Завершившиеся в других точках: $f > 760$	82	51	0	0
в том числе на границе Ω	18	18	0	0
Всего попыток	100	100	100	100

- Во-вторых, метод АДЭ-Р является методом поиска *глобального* минимума. Он с большой вероятностью обеспечивает нахождение самого глубокого из минимумов функции χ^2 в области допустимых значений параметров. В то время как вероятность обнаружения минимума квазиньютоновским методом с апдейтом гессиана по формуле Давидона – Флетчера – Пауэлла $p^{DFP} = 5 \pm 2,2$, а вероятность обнаружения минимума методом симплексов составила $p^{NM} = 20 \pm 4,0\%$, вероятность нахождения глобального минимума методом АДЭ-Р составила $p^{ADE-R} = 79 \pm 4\%$ в случае минимальной начальной популяции $N_p^{\min} = 10$. Эта вероятность зависит от величины исходной популяции и может быть приближена к единице путем увеличения популяции в три раза.
- С другой стороны, при необходимости для нахождения второстепенных минимумов можно использовать предварительные результаты, выдаваемые методом АДЭ в процессе расчетов.

В работе [1] на примере анализа дифференциальных сечений упругого рассеяния $\pi^- + {}^{28}\text{Si}$ при $T^{\text{lab}}=130$ МэВ было показано, что типичная картина профиля невязки χ^2 (2.14) с экспериментальными данными из [106] в зависимости от параметров σ и α при фиксированном значении $\beta = 0,9 \text{ фм}^{-2}$

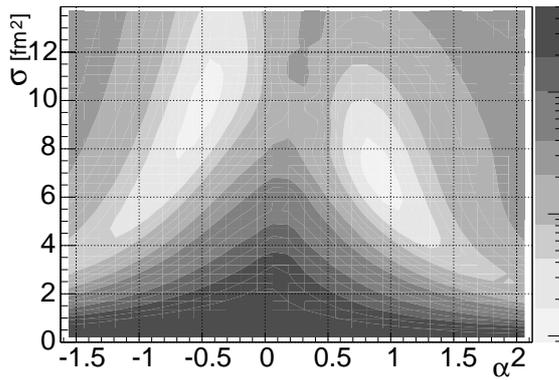


Рис. 2.9. Значение χ^2/k в зависимости от параметров α и σ при фиксированном $\beta = 0,9$ для рассеяния $\pi^- + {}^{28}\text{Si}$ при $T^{\text{lab}} = 130$ МэВ. Экспериментальные данные — из [106]

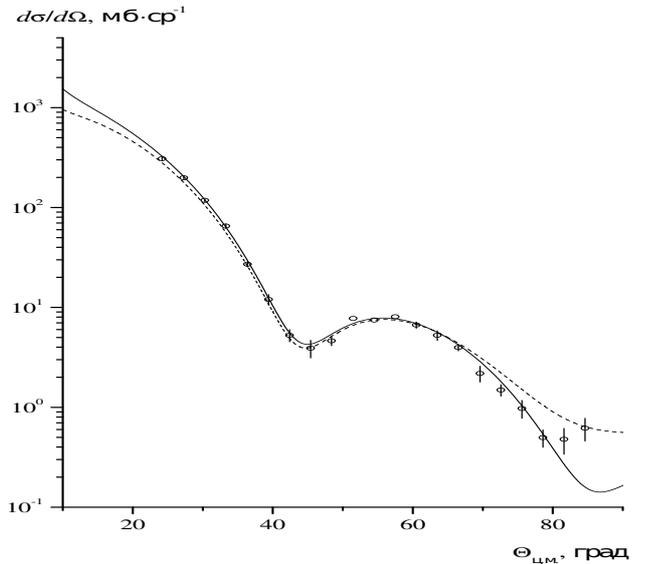


Рис. 2.10. Подгонка сечений рассеяния $\pi^- + {}^{28}\text{Si}$ при $T^{\text{lab}} = 130$ МэВ к экспериментальным данным из [106]. Сплошная кривая — главный минимум с параметрами $\sigma = 7.02$, $\alpha = 0.88$, $\beta = 0.82$, $\chi^2/k = 2.4$; штриховая линия — второстепенный минимум с параметрами $\sigma = 10.95$, $\alpha = -0.328$, $\beta = 0.598$, $\chi^2/k = 4.2$

имеет два минимума (рис. 2.9), для одного из которых параметр α лежит в области положительных, а для другого — в области отрицательных значений. Стандартные минимизаторы при наличии нескольких локальных минимумов в области допустимых значений параметров оставляют исследователю задачу выявления каждого из минимумов и детерминации наиболее глубокого из них. Используя традиционные методы фитирования, необходимо проводить массовые расчеты с разными начальными значениями параметров, чтобы найти все возможные локальные минимумы и затем выбрать один из них как оптимальный. В то же время метод АДЭ с большой вероятностью обеспечивает нахождение в конечном итоге самого глубокого минимума функции χ^2 в области допустимых значений параметров.

На рис. 2.10 представлены теоретические кривые дифференциальных сечений упругого рассеяния в сравнении с этими же экспериментальными данными, соответствующими двум минимумам χ^2 , когда подгонялись все три параметра πN -амплитуды. Видно, что сплошная и штриховая кривые сечений, соответствующие глобальному и второстепенному минимумам χ^2 ($\chi^2/k = 2,4$ и $\chi^2/k = 4,2$ ⁵), фактически совпадают во всей области изме-

⁵ $k = n - l$ — количество степеней свободы, n — количество экспериментальных точек, l — количество свободных подгоняемых параметров

рений, кроме области больших углов, где сплошная отклоняется от двух экспериментальных точек, а штриховая — от пяти. В то же время значения параметров подгонки существенно различаются, а полученные параметры α оказываются даже противоположными по знаку — соответственно 0,88 и $-0,328$.

Во всех приведенных ниже оценках метод АДЭ в качестве конечного результата расчетов самостоятельно определяет наиболее глубокий минимум с вероятностью, близкой к единице.

Отметим, однако, что в случаях, когда второй минимум имеет сопоставимое с главным значение χ^2 , а два соответствующих набора параметров обеспечивают неплохое согласие с экспериментальными данными, то для того чтобы сделать окончательные выводы, следует привлекать дополнительные данные. В качестве таковых могли бы выступать, например, экспериментальные данные о полных сечениях рассеяния и его поляризационных характеристиках.

Второй локальный минимум можно найти, искусственно ограничив область поиска окрестностью притяжения этого минимума либо задав область положения начальной популяции достаточно близко к предполагаемому положению искомого локального минимума.

Стандартный метод нахождения второго и следующих локальных минимумов — сканирование всей области поиска для определения приблизительного положения каждого из них. Однако в нашем случае для этого можно использовать предварительные результаты, выдаваемые АДЭ в процессе расчетов — и это также является преимуществом метода.

Для удобства отслеживания прогресса в расчетах исследователем (актуально в случае задач, требующих большого количества расчетов) реализована функция, позволяющая выдавать на стандартный output информацию о значении целевой функции в текущем лучшем из найденных векторов из пространстве параметров и координаты этого вектора. Алгоритм работает таким образом, что если изначальная область поиска захватывала области притяжения нескольких сравнимых локальных минимумов, то с большой вероятностью на некотором этапе расчетов, а именно, когда кончается этап дифференциации локальных минимумов и начинается этап конкуренции между ними, наилучший из текущих векторов популяции оказывается попеременно то в одном, то в другом из локальных минимумов. Отследив промежуточные результаты на этом этапе, часто можно доста-

точно точно локализовать положение других локальных минимумов целевой функции. Эта возможность особенно актуальна в многопараметрических задачах (см. раздел 3), в которых определение положения локального минимума прямым сканированием всей области поиска для определения локализации каждого минимума затруднено.

Основные результаты, представленные в главе 2

- На базе алгоритма глобальной минимизации АДЭ реализован программный комплекс для определения параметров «in-medium» амплитуды рассеяния π -мезонов на внутриядерных нуклонах.
- В диссертации проведен анализ данных по дифференциальным сечениям упругого рассеяния пионов с кинетическими энергиями от 130 до 290 МэВ, покрывающими область известного $\Delta(1232)P_{33}$ -резонанса пион-нуклонного рассеяния на ядрах ^{28}Si , ^{40}Ca , ^{54}Ni , ^{208}Pb . При этом используется микроскопический оптический потенциал, построенный на основе трехпараметрической πN -амплитуды явного вида. Само сечение пион-ядерного рассеяния вычисляется с помощью решения релятивистского волнового уравнения, что позволяет автоматически учесть как эффекты релятивизации, так и эффекты кулоновского и ядерного искажения во входном и выходном каналах рассеяния. Получен ряд физических результатов:
 - В рамках трехпараметрической модели МОП описаны экспериментальные данные о дифференциальных сечениях упругого рассеяния пионов на ядрах. При этом качество описания экспериментальных данных в рамках трехпараметрической модели сравнимо с результатами более сложных моделей с большим числом параметров.
 - Для ядер ^{28}Si , ^{40}Ca , ^{58}Ni , ^{208}Pb получен набор значений параметров «in-medium» амплитуды рассеяния π -мезонов на внутриядерных нуклонах при кинетических энергиях налетающих пионов 130, 162, 180, 226, 230, 291 МэВ. Полученный набор значений параметров «in-medium» амплитуды рассеяния π -мезонов на внутриядерных нуклонах позволил сделать ряд теоретических выводов о характере рассеяния пионов внутри ядра.
- Показана эффективность применения метода АДЭ для решения поставленной задачи.

- Показано, что в заданной области метод АДЭ находит минимум с существенно большей вероятностью, чем методы SIMPLEX и MIGRAD.
- Показано, что вероятность нахождения глобального минимума методом АДЭ может быть приближена к единице путем увеличения размера начальной популяции до 30.
- Реализована возможность определить статистическую точность оценки параметров по профилю минимизируемой функции.

По материалам главы опубликованы работы [1, 2, 6, 11, 15–18].

3. Анализ структуры однослойных везикул ДМФХ по данным МУРР в рамках модифицированной модели РФФ с использованием АДЭ

Информация о структуре и свойствах однослойных везикул играет важную роль в структурной биологии, биохимии и фармакологии. В диссертации метод разделенных формфакторов (РФФ) адаптирован для моделирования структуры полидисперсной популяции однослойных везикул димиристоилфосфатидилхолина (ДМФХ) по данным малоуглового рентгеновского рассеяния (МУРР) [3, 4]. Ранее метод РФФ, развитый в работах [112, 117, 123, 129], использовался для анализа структуры полидисперсных везикулярных систем по данным малоуглового рассеяния нейтронов (МУРН). Однако с помощью рентгеновского рассеяния существует возможность рассмотреть структуру везикул более контрастно. В диссертации в методе РФФ учтены флуктуации параметров модели бислоя. Из ряда рассмотренных моделей выбрана модель, позволяющая описать экспериментальные данные. Получены значения параметров везикулярной системы.

В начале главы приводится модель разделенных формфакторов, адаптированная для анализа структуры полидисперсной популяции однослойных везикул ДМФХ по данным МУРР. Параметры модели РФФ, определяющие структуру везикулярной системы, оптимизируются для наилучшего соответствия модели экспериментальным данным с использованием параллельного алгоритма асинхронной дифференциальной эволюции (АДЭ). Описание программного комплекса, оптимизирующего параметры модели, приводится во втором разделе.

Затем обсуждаются полученные результаты и их интерпретация. Сравниваются разные модели для внутренней структуры бислоя оболочки везикул. Показано, что разработанный на базе РФФ-АДЭ подход не только обеспечивает согласие рассчитанных параметров с предыдущими оценками на основе данных МУРН, но и дает возможность уточнить внутреннюю структуру липидного бислоя.

В последнем разделе производится оценка и сравнение вероятности нахождения верной оценки параметров методом АДЭ в сравнении с другими методами (SIMPLEX и MIGRAD). Описывается параллельная реа-

лизация комплекса программ и приводятся результаты оценки ускорения параллельных вычислений.

В конце сделаны выводы по результатам, изложенным в главе.

3.1. Постановка задачи, основные формулы

Макроскопическое когерентное сечение МУРР монодисперсной популяции везикул в рамках модели РФФ определяется следующей формулой [114]:

$$\frac{d\Sigma}{d\Omega_{mon}}(q) = n_x \cdot A^2(q) \cdot S(q), \quad (3.1)$$

где n_x — число везикул в единице объема, $A(q)$ — амплитуда рассеяния, $S(q)$ — структурный фактор (СФ), $q = \frac{4\pi \sin(\frac{\theta}{2})}{\lambda}$ — вектор рассеяния, θ — угол рассеяния, λ — длина волны гамма-кванта. Структурный фактор учитывает межвезикулярное взаимодействие. Для сферы СФ в дебаевском приближении записывается как [113, 114]:

$$S(qR) = 1 - \frac{8V_\nu}{\nu} \left(\frac{\sin(2qR)}{2qR} \right). \quad (3.2)$$

Здесь $V_\nu = \frac{4}{3}\pi R^3$ — объем везикулы, $\nu = 1/n_x$ — объем растворителя, приходящийся на одну везикулу. В дальнейших расчетах структурный фактор пока считается равным единице: $S(q) \approx 1$.

Амплитуда рассеяния $A(q)$ в сферически-симметричном случае равна [114]:

$$A(q) = 4\pi \int \rho(r) \frac{\sin(qr)}{qr} r^2 dr. \quad (3.3)$$

Здесь $\rho(r) = \rho_C(r) - \rho_{suc}(r)$ — разность плотности длины рассеяния липидного бислоя везикулы $\rho_C(r)$ и раствора сахарозы $\rho_{suc}(r)$, называемая контрастом.

Формулу (3.3) можно переписать в виде

$$A(q) = 4\pi \int_{-d/2}^{d/2} \rho(x) \frac{\sin[(R+x)q]}{(R+x)q} \cdot (R+x)^2 dx. \quad (3.4)$$

Здесь R — радиус везикулы, d — толщина мембраны. Интегрирование выражения (3.4) в предположении $R \gg \frac{d}{2}$, $R+x \approx R$ приводит к следующей

формуле:

$$A_{\text{SFF}}(q) = 4\pi R^2 \frac{\sin(qR)}{qR} \int_{-d/2}^{d/2} \rho(x) \cos(qx) dx. \quad (3.5)$$

Таким образом, макроскопическое сечение монодисперсной популяции везикул (3.1) может быть записано в виде

$$\frac{d\Sigma}{d\Omega_{\text{mon}}} (q) = n_x F_s(q; R) F_b(q; \Theta_b), \quad (3.6)$$

где $F_s(q; R)$ — формфактор бесконечно тонкой сферы радиуса R [123]:

$$F_s(q; R) = \left(4\pi R^2 \frac{\sin(qR)}{qR} \right)^2, \quad (3.7)$$

и $F_b(q; \Theta_b)$ — формфактор симметричного липидного бислоя:

$$F_b(q; \Theta_b) = \left(\int_{-d/2}^{d/2} \rho(x; \Theta_b) \cos(qx) dx \right)^2, \quad (3.8)$$

Θ_b — набор параметров, характеризующих распределение плотности длины рассеяния внутри липидного бислоя. Формулы (3.6), (3.7) и (3.8) описывают модель разделенных формфакторов для однослойных везикул большого радиуса R [119]. Ее достоинством является возможность получить явное выражение для макроскопического сечения в случае, когда $\rho(x)$ — интегрируемая функция. Например, в случае (рис.3.1, модель «а»), когда контраст описывается одним параметром $\Theta_b = \{d\}$:

$$\rho(x; \Theta_b) = \rho(x; d) = \begin{cases} \Delta\rho = \text{const}, & -\frac{d}{2} < x < \frac{d}{2}, \\ 0, & x < -\frac{d}{2}; \quad x > \frac{d}{2}, \end{cases} \quad (3.9)$$

формфактор симметричного липидного бислоя равен

$$F_b(q; d) = \left(\Delta\rho \cdot d \cdot \frac{\sin(qd/2)}{qd/2} \right)^2. \quad (3.10)$$

Соответственно, макроскопическое сечение

$$\frac{d\Sigma}{d\Omega_{\text{mon}}} (q; R, d) = n_x \left(4\pi R^2 \Delta\rho d \cdot \frac{\sin(qR)}{qR} \frac{\sin(qd/2)}{qd/2} \right)^2. \quad (3.11)$$

Несмотря на то что однородная плотность в виде (3.9) не является реалистичной, она довольно часто использовалась в расчетах (см., например, [116, 120, 124, 126]). Однородная плотность позволяет оценить значения среднего радиуса везикул и их полидисперсность, при этом, как показыва-

ют расчеты, толщина бислоя недооценивается по сравнению с моделями, учитывающими внутреннюю структуру бислоя.

Учет полидисперсности везикул. Полидисперсность (флуктуации радиуса) везикул описывается несимметричным распределением Шульца [115, 120, 127]:

$$G(\tilde{R}, \langle R \rangle) = \frac{\tilde{R}^m}{m!} \left(\frac{m+1}{\langle R \rangle} \right)^{m+1} \exp \left[-\frac{(m+1)\tilde{R}}{\langle R \rangle} \right], \quad (3.12)$$

где $\langle R \rangle$ — средний радиус везикулы, m — коэффициент полидисперсности¹. Тогда

$$F_s^{\text{poly}}(q; \langle R \rangle, m) = \int_{R_{\min}}^{R_{\max}} F_s(q, \tilde{R}) G(\tilde{R}, \langle R \rangle) d\tilde{R} \Big/ \int_{R_{\min}}^{R_{\max}} G(\tilde{R}, \langle R \rangle) d\tilde{R}. \quad (3.13)$$

В расчетах, приведенных ниже, $R_{\min} = 10$ нм, $R_{\max} = 100$ нм.

Учет флуктуаций параметров внутренней структуры бислоя. Как показали расчеты в [21] и [3], для адекватного воспроизведения МУРР-спектров при больших значениях вектора рассеяния q требуется также учет флуктуаций (разброса) параметров внутренней структуры бислоя. Для учета разброса в значениях параметров внутренней структуры бислоя мы будем использовать усеченное нормальное распределение² $f(\xi)$. Тогда

$$F_b^{\text{fluct}}(q, \langle \Theta_b \rangle) = \int_{-K}^K F_b(q, \tilde{\Theta}_b(\langle \Theta_b \rangle, k_{\text{fl}}, \xi)) f(\xi) d\xi \Big/ \int_{-K}^K f(\xi) d\xi. \quad (3.14)$$

Допустим, функция контраста $\rho(x; \Theta_b)$ описывается параметрами внутренней структуры: $\Theta_b = \{d, D, L, \rho_H\}$. В расчетах, представленных ниже, предполагается, что разброс в значениях параметров внутренней структуры обусловлен флуктуацией толщины мембраны³. Тогда компоненты $\tilde{\Theta}_b$

¹Среднеквадратичное относительное отклонение, характеризующее полидисперсность, связано с коэффициентом полидисперсности m как $\sigma = \sqrt{\frac{1}{m+1}}$.

²Функция распределения и плотность вероятности усеченного с двух сторон нормального распределения: $P(\xi) = \begin{cases} 0, & \xi < -K, \\ \frac{\Phi(\xi) - \Phi(-K)}{1 - \Phi(-K) - \Phi(K)}, & -K < \xi < K, \\ 1, & \xi > K \end{cases}$ и $f(\xi) = \begin{cases} \frac{\phi(\xi)}{1 - \Phi(-K) - \Phi(K)}, & -K < \xi < K, \\ 0, & \xi < -K, \xi > K. \end{cases}$

Здесь $\Phi(\xi)$ — функция нормального распределения $\phi(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}}$.

³Именно этот вариант дал наилучшее согласие с экспериментом при однопараметрическом описании величины разброса параметров.

под интегралом в (3.14) равны соответственно:

$$\tilde{d}=\langle d \rangle(1+k_{\text{fl}} \xi), \tilde{D}=\langle D \rangle\left(1+\frac{\langle d \rangle}{\langle D \rangle} k_{\text{fl}} \xi\right), \tilde{L}=\langle L \rangle\left(1-\frac{\langle d \rangle}{\langle L \rangle} k_{\text{fl}} \xi\right), \tilde{\rho}_H \equiv\langle \rho_H \rangle \equiv \rho_H,$$

где $\langle d \rangle$, $\langle D \rangle$ и $\langle L \rangle$ — средние значения параметров внутренней структуры бислоя везикулы; k_{fl} — коэффициент флуктуаций параметров внутренней структуры бислоя везикулы. В расчетах, приведенных ниже, пределы интегрирования выбраны равными $K=3$. С учетом флуктуаций параметров внутренней структуры бислоя макроскопическое сечение полидисперсной везикулярной системы имеет следующий вид:

$$\frac{d\Sigma}{d\Omega}(q)=I_m(q)=n_x F_s^{\text{poly}}(q ; \langle R \rangle, m) F_b^{\text{fluct}}(q ; \langle \Theta_b \rangle, k_{\text{fl}}). \quad (3.15)$$

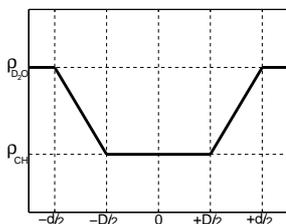
Учет функции разрешения спектрометра. Измеряемое значение макроскопического сечения отличается от его теоретического значения (3.15), так как спектрометр имеет конечное разрешение. Считаем, что функция разрешения спектрометра является симметричной функцией, близкой к гауссовой [125]. Тогда измеряемое в эксперименте сечение $I(q)$ может быть записано как

$$I(q)=I_m(q)+\frac{1}{2} \Delta^2 \frac{d^2 I_m(q)}{dq^2}+I_B, \quad (3.16)$$

где $\Delta^2=3,6 \cdot 10^{-7} \text{ см}^2$ — второй момент функции разрешения спектрометра, I_B — величина некогерентного фона.

Моделирование плотности длины рассеяния липидного бислоя везикулы. В работе [117] для анализа спектров МУРН была предложена простая двухпараметрическая гидрофильно-гидрофобная (ГГ)⁴ модель плотности длины рассеяния поперек мембраны. ГГ-приближение с линейным распределением воды внутри мембраны является эффективным для моделирования липидных систем с большой величиной гидратации или с малым размером полярной головы молекул, поскольку для таких везикулярных систем вклад в распределение от воды внутри бислоя в общем

⁴ ГГ-модель плотности длины рассеяния (Hydrophilic-hydrophobic (HH-model)) для моделирования везикул ДМФХ в растворе сахарозы в тяжелой воде D_2 согласно модели РФФ-ГГ по данным МУРН [117] и значения параметров везикул ДМФХ в растворе сахарозы в тяжелой воде D_2O согласно модели РФФ-ГГ по данным МУРН [117]



sucrose, %	R , nm	d , nm	D , nm	m	σ	$\frac{\chi^2}{k}$
0	$30,8 \pm 0,5$	$5,0 \pm 0,1$	$1,7 \pm 0,2$	$9,8 \pm 0,8$	33%	1,1
5	$27,1 \pm 0,5$	$4,8 \pm 0,2$	$2,0 \pm 0,3$	$9,7 \pm 0,9$	34%	6,9
10	$25,1 \pm 0,3$	$4,6 \pm 0,2$	$2,3 \pm 0,3$	$12,8 \pm 0,8$	29%	6,6
20	$21,3 \pm 0,2$	$4,5 \pm 0,2$	$2,5 \pm 0,3$	$13,8 \pm 0,5$	26%	9,5

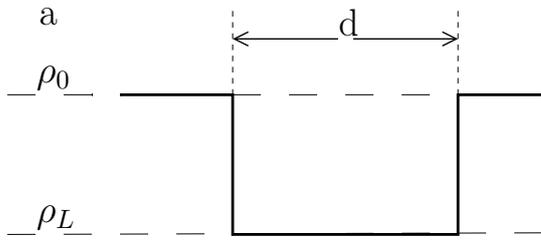


Рис. 3.1. а-модель. $\rho_0 = 10.94 \cdot 10^{12} \text{ см}^{-1}$; $\rho_L = 7.95 \cdot 10^{12} \text{ см}^{-1}$

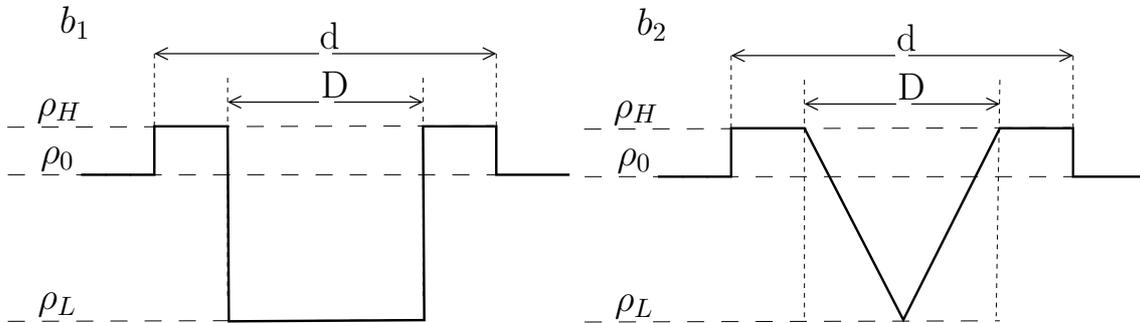


Рис. 3.2. b-модели. $\rho_0 = 10.94 \cdot 10^{12} \text{ см}^{-1}$; $\rho_L = 4.71 \cdot 10^{12} \text{ см}^{-1}$; ρ_H — варьируемый параметр

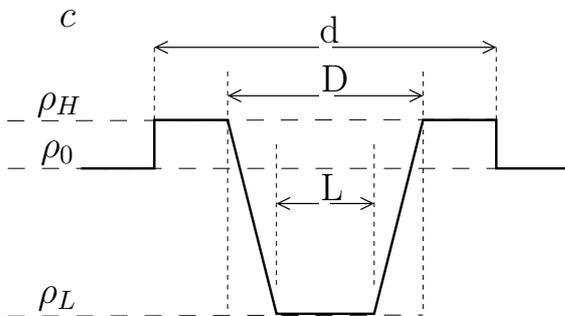


Рис. 3.3. c-модель. $\rho_0 = 10.94 \cdot 10^{12} \text{ см}^{-1}$; $\rho_L = 4.71 \cdot 10^{12} \text{ см}^{-1}$; ρ_H — варьируемый параметр

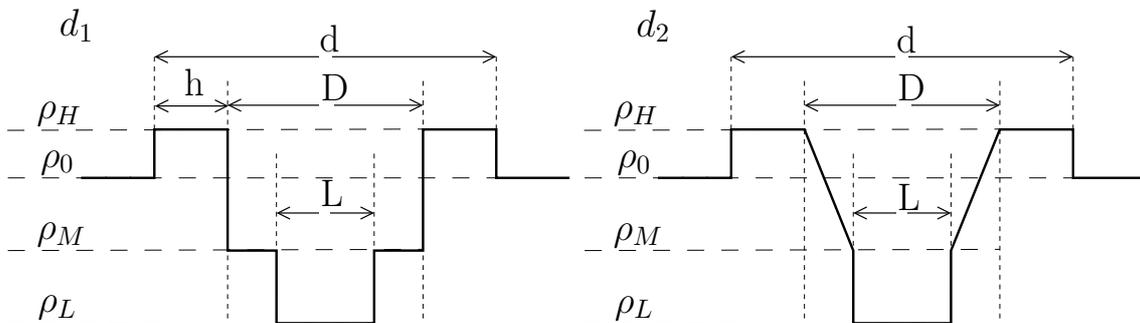


Рис. 3.4. d-модели. $\rho_0 = \rho_{\text{a0S}} = 10.94 \cdot 10^{12} \text{ см}^{-1}$; $\rho_L = \rho_{\text{CH}_3} = 4.71 \cdot 10^{12} \text{ см}^{-1}$; $\rho_M = \rho_{\text{CH}_2} = 7.95 \cdot 10^{12} \text{ см}^{-1}$; ρ_H — варьируемый параметр

распределение плотности длины рассеяния ρ доминирует по сравнению со вкладом от полярной части мембраны [118]. В рамках ГГ-модели для моделирования функции ρ необходимо только два свободных параметра: толщина мембраны d и толщина гидрофобной области D , что существенно упрощает процесс подгонки по сравнению с более сложными моделями для ρ . И все же, такой упрощенный подход во многих случаях практически

Таблица 3.1. Модели плотности длины рассеяния липидного бислоя. Зафиксированные ($\rho_0 = 10,94 \times 10^{12} \text{ см}^{-1}$, $\rho_{L_a} = 7,95 \times 10^{12} \text{ см}^{-1}$, $\rho_M = 7,95 \times 10^{12} \text{ см}^{-1}$ и $\rho_L = 4,71 \times 10^{12} \text{ см}^{-1}$) и варьируемые (Θ) параметры моделей. Θ_s — параметры, описывающие размер, концентрацию везикул и некогерентный фон. Θ_b — параметры, относящиеся к внутренней структуре бислоя

Модель	Фиксир. параметры	Варьируемые параметры Θ		
		Θ_s	Θ_b	k_{fl}
a	ρ_0, ρ_{L_a}	R, m, n_x, I_B	$d,$	k_{fl}
b_1	ρ_0, ρ_L	R, m, n_x, I_B	d, D, ρ_H	k_{fl}
b_2	ρ_0, ρ_L	R, m, n_x, I_B	d, D, ρ_H	k_{fl}
c	ρ_0, ρ_L	R, m, n_x, I_B	d, D, L, ρ_H	k_{fl}
d_1	ρ_0, ρ_L, ρ_M	R, m, n_x, I_B	d, D, L, ρ_H	k_{fl}
d_2	ρ_0, ρ_L, ρ_M	R, m, n_x, I_B	d, D, L, ρ_H	k_{fl}

не влияет на качество подгонки для данных МУРН-спектров и обеспечивает возможность определить основные параметры везикулярных систем без привлечения дополнительных экспериментальных методов (рассеяния света, дифракции и т.д.). В работах [117, 129] ГГ-модель подтвердила свою эффективность для МУРН-анализа структуры везикулярных систем в избытке жидкости.

Однако в силу физических особенностей МУРР ГГ-модель плотности не применима для анализа МУРР-спектров. Поэтому в диссертации проведено исследование с использованием целого ряда моделей $\rho(x, \Theta_b)$ для того, чтобы определить наиболее простую из моделей, которая позволила бы описать экспериментальные результаты. Исследованные модели представлены на рис. 3.1 – 3.4 и упорядочены в табл. 3.1 по возрастанию числа варьируемых параметров. Простейшая модель — «а» — описывается одним параметром. Модели «b» и «с» — двухступенчатые, они описываются соответственно тремя и четырьмя ($\Theta_b = \{d, D, L, \rho_H\}$) свободными параметрами. Трехступенчатые модели «d» описываются четырьмя свободными параметрами.

Формулы (3.17)–(3.22) описывают соответствующие формфакторы

симметричного бислоя для разных моделей:

$$F_b^a = \left(\frac{-2(\rho_0 - \rho_L)}{q} \sin\left(\frac{qd}{2}\right) \right)^2, \quad (3.17)$$

$$F_b^{b1} = \left(\frac{2(\rho_H - \rho_0)}{q} \sin\left(\frac{qd}{2}\right) - \frac{2(\rho_H - \rho_L)}{q} \sin\left(\frac{qD}{2}\right) \right)^2, \quad (3.18)$$

$$F_b^{b2} = \left(\frac{2(\rho_H - \rho_0)}{q} \sin\left(\frac{qd}{2}\right) + \frac{2(\rho_H - \rho_L)}{\frac{Dq^2}{2}} \left(\cos\left(q\frac{d-D}{2}\right) - 1 \right) \right)^2, \quad (3.19)$$

$$F_b^c = \left(\frac{2(\rho_H - \rho_0)}{q} \sin\left(\frac{qd}{2}\right) + \frac{2(\rho_H - \rho_L)}{\left(\frac{D}{2} - \frac{L}{2}\right)q^2} \left(\cos\left(\frac{qD}{2}\right) - \cos\left(\frac{qL}{2}\right) \right) \right)^2, \quad (3.20)$$

$$F_b^{d1} = \left(\frac{2(\rho_H - \rho_0)}{q} \sin\left(\frac{qd}{2}\right) - \frac{2(\rho_H - \rho_M)}{q} \sin\left(\frac{qD}{2}\right) - \frac{2(\rho_M - \rho_L)}{q} \sin\left(\frac{qL}{2}\right) \right)^2, \quad (3.21)$$

$$F_b^{d2} = \left(\frac{2(\rho_H - \rho_0)}{q} \sin\left(\frac{qd}{2}\right) + \frac{2(\rho_H - \rho_M)}{\left(\frac{D}{2} - \frac{L}{2}\right)q^2} \left(\cos\left(\frac{qD}{2}\right) - \cos\left(\frac{qL}{2}\right) \right) - \frac{2(\rho_M - \rho_L)}{q} \sin\left(\frac{qL}{2}\right) \right)^2. \quad (3.22)$$

Подгонка параметров моделей бислоя осуществляется минимизацией взвешенной суммы квадратов отклонения теоретических сечений $\hat{I}_i(q_i, \Theta)$ от соответствующих экспериментальных данных $I_i(q_i)$:

$$\Delta^2 = f(\Theta) = \sum_i \frac{\left(I_i(q_i) - \hat{I}_i(q_i, \Theta) \right)^2}{\sigma_i^2}, \quad (3.23)$$

где Θ — вектор варьируемых параметров (см. столбцы 4 и 5 табл. 3.1), а веса выбраны как⁵ $w_i = \frac{1}{\sigma_i^2} = \frac{1}{0.01 \times I_i(q_i)^2}$.

Число варьируемых параметров в МУРР-расчетах существенно возрастает по сравнению с МУРН, что приводит к снижению устойчивости итерационного процесса на основе взвешенного метода наименьших квадратов, применявшегося для обработки МУРН-спектров в [117, 129]. В работе значения параметров Θ , обеспечивающие минимум Δ^2 -функционала, определяются с помощью параллельного алгоритма асинхронной дифференциальной эволюции (АДЭ) [1, 6, 7, 9, 10, 122].

3.2. Описание комплекса программ

Комплекс программ, осуществляющий нахождение параметров липидного бислоя везикул ДМФХ, состоит из следующих компонент:

⁵Такой выбор весов соответствует следующей интерпретации полученного значения Δ^2 : если $\Delta^2/k < 1$, где число степеней свободы $k = N - l$, N — число экспериментальных точек, l — количество фитируемых параметров — это означает, что теоретическая кривая совпадает с экспериментальными данными в пределах 10% погрешности.

- Модуль, написанный на языке ФОРТРАН, предназначенный для расчета теоретической кривой и оценки Δ^2 при сравнении модели с экспериментальными данными (`dmprc_saxsJ_calc.f` и `dmprc_saxsJ_calc.h`)⁶.
- Библиотека АДЭ (MPI-вариант). Заголовочные файлы использованных частей библиотеки: «`ade_mpi_optimizer.h`», «`ade_settings.h`», «`parameter_list.h`», «`problem_definition.h`».
- Модуль (программа) (`dmprc_mpi_master.cpp`), которая
 - загружает исходные данные: экспериментальные точки и фиксированные параметры модели;
 - задает начальные параметры алгоритма АДЭ и границы допустимых значений варьируемых параметров;
 - получает из алгоритма очередную точку (вектор параметров Θ), для которого необходимо провести расчет;
 - вызывает функцию `CALCPART()` (получение $f(\Theta)$);
 - передает результат обратно в алгоритм (и так до окончания расчетов);
 - распечатывает окончательный результат вычислений, оценки ошибок, код завершения алгоритма и др. параметры завершения.

3.3. Обсуждение численных результатов

Эксперимент и приготовление образцов. Измерение спектров МУРР проводилось на установке А2 синхротронного источника Doris III (DESY, Германия). В эксперименте детекторы располагались на расстоянии 1 м и 3 м (D_1 и D_3) от образца. Спектры измерялись при температуре 30°C. Однослойные везикулы ДМФХ изготавливались экструзией 15 ммоль (1% по весу) суспензии многослойных везикул ДМФХ в водном растворе сахарозы с концентрацией 40% через фильтры с диаметром пор 50 нм.

Результаты оптимизации МУРР без учета флуктуаций толщины мембраны по полному и по усеченному спектрам. Сначала приведем результаты, получаемые при оптимизации целевого функционала

⁶Исходный код был перекомпилирован таким образом, чтобы выделить функцию, которую можно было бы использовать в написанном на C++ модуле, запускающем алгоритм АДЭ: `FUNCTION CALCPART(inpned, rinpA, eeq, eei, eer, ees)`.

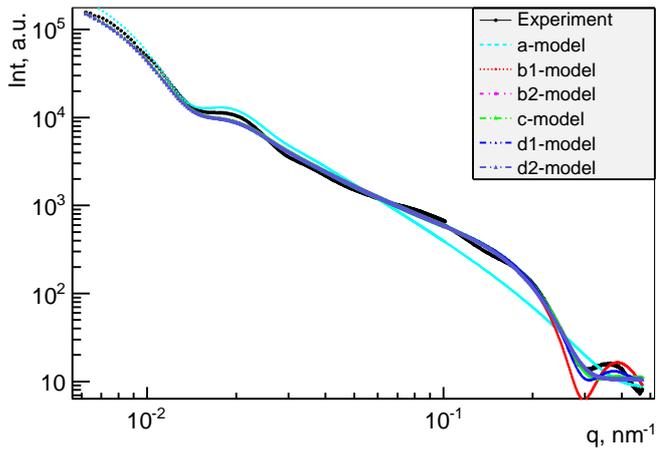


Рис. 3.5. Экспериментальный и теоретические спектры, соответствующие оптимальным параметрам без учета флуктуаций Θ_b (фитирование по полному спектру)

результаты⁷ представлены на рис. 3.5 и в табл. 3.2.

Все модели дают приблизительно одинаковую оценку радиуса R везикул (см. табл. 3.2 и 3.3), однако недостаточно хорошо описывают поведение спектра при больших значениях вектора рассеяния q (см. рис. 3.5).

Упрощенная модель, не учитывающая флуктуаций толщины мембраны, не может описать бóльшую глубину последнего минимума.

Оптимизация целевого функционала (3.23) для РФФ-модели, не учитывающей флуктуаций толщины мембраны по усеченному набору данных (при $q < 0.4 \text{ nm}^{-1}$, 820 из 894 точек), приводит к результатам, представленным в табл. 3.3 и на рис. 3.6.

Оценки для толщины мембраны d при этом для всех моделей сбли-

⁷Поскольку экспериментальный спектр изначально не был нормирован, оценка параметра n_x здесь и далее не показывает реальное количество везикул в единице объема, т.к. включает в себя нормировочный фактор.

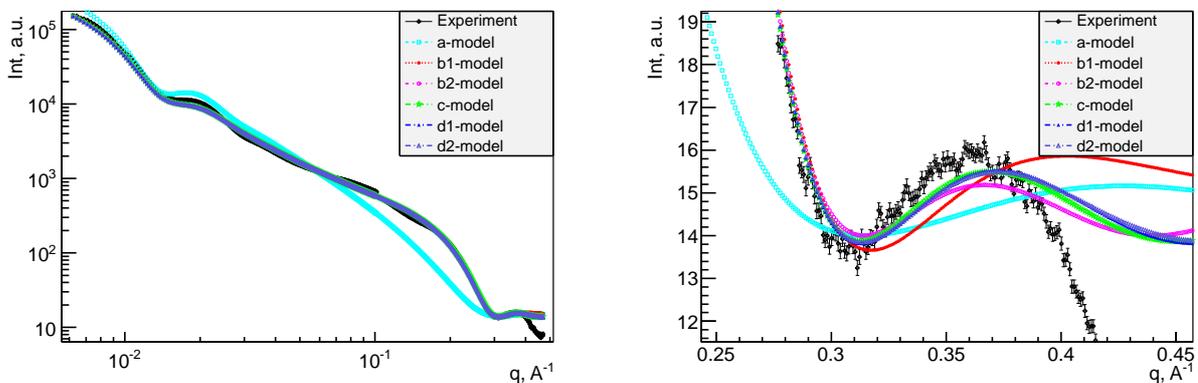


Рис. 3.6. То же, что на рис. 3.5, но при фитировании по усеченному спектру ($q < 3.8 \text{ nm}^{-1}$)

(3.23) для РФФ-модели, не учитывающей флуктуаций толщины мембраны. Это ситуация, когда в формуле (3.15) для макроскопического сечения полидисперсной популяции везикулярной системы вместо $F_b^{\text{fluct}}(q, \langle \Theta_b \rangle, k_{\text{fl}})$, рассчитываемого по формуле (3.14), используется $F_b(q, \Theta_b)$, рассчитываемый по формуле (3.8). Полученные в этом случае

Таблица 3.2. Результаты оценки параметров везикулярных систем по данным МУРР без учета флуктуаций Θ_b . Количество экспериментальных точек $n = 894$; количество варьируемых переменных l ; количество степеней свободы $n - l$. Результаты округлены до последней значащей цифры согласно статистическим ошибкам приближения [3]

Модель	$\langle R \rangle$, нм	(σ)	d , нм	D , нм	L , нм	ρ_H , $\times 10^{10} \text{см}^{-1}$	n_x , $\times 10^{18}$ см^{-3}	I_B , а.у.	l	$\frac{\Delta^2}{n-l}$
a	20,13	0,230	—	1,31	—	$\rho_L = 7,95$	8,7	8,9	6	7,64
b_1	18,73	0,274	5,34	1,53	—	11,7	2,38	9,7	7	2,00
b_2	19,70	0,257	3,91	3,91	—	14,3	3,32	10,8	7	1,92
c	18,95	0,269	3,88	3,88	3,16	15,6	2,24	10,4	7	1,78
d_1	18,97	0,267	3,83	2,57	1,61	11,0	2,74	9,9	8	1,68
d_1^*	18,9	0,269	4,12	3,0	0,52	12,7	3,6	10,4	8	1,80
d_2	18,95	0,268	3,97	3,97	1,37	15,1	2,06	10,5	8	1,87

Таблица 3.3. То же, что в таблице 3.2, но по усеченному спектру: фитирование данных только при $q < 0,38 \text{ нм}^{-1}$, $n = 820$ [3]

Модель	$\langle R \rangle$, нм	σ	d , нм	D , нм	L , нм	ρ_H , $\times 10^{10} \text{см}^{-1}$	n_x , $\times 10^{18}$ см^{-3}	I_B , а.у.	l	$\frac{\Delta^2}{n-l}$
a	20,1	0,230	—	1,31	—	$\rho_L = 7,95$	8,7	7,7	5	7,8
b_1	18,9	0,268	4,54	1,88	—	12,4	1,8	13,7	7	0,762
b_2	19,0	0,267	4,00	4,00	—	14,3	3,3	14,0	7	0,678
c	19,6	0,249	4,59	2,82	0,60	12,7	2,6	14,0	7	0,897
d_1	19,0	0,267	4,18	3,78	0,80	12,8	2,7	13,7	8	0,738
d_2	19,0	0,267	4,04	3,97	1,37	14,7	2,0	13,8	8	0,738

жаются. Оценка для некогерентного фона повышается. На графике амплитуда колебаний теоретического спектра в области больших q для всех моделей сближается и уменьшается.

Однако на экспериментальном спектре прослеживается более глубокий минимум при больших значениях вектора рассеяния q .

Результаты оптимизации МУРР с учетом флуктуаций толщины мембраны по полному и по усеченному спектрам. При учете флуктуаций толщины мембраны по формуле (3.14) фитирование целевого функционала (3.23) приводит к результатам, представленным в табл. 3.4 и на рис. 3.7 и 3.8.

Фитирование целевого функционала (3.23) по усеченному набору данных (при $q < 0,4 \text{ нм}^{-1}$, 820 из 894 точек) приводит к результатам, представленным в табл. 3.5.

Таблица 3.4. Результаты МУРР с учетом флуктуаций Θ_b (толщины мембраны). Количество экспериментальных точек $n = 894$; l — количество варьируемых переменных; $n - l$ — количество степеней свободы; Результаты округлены до последней значащей цифры (согласно статистическим ошибкам приближения) [3]

Модель	$\langle R \rangle$, нм	σ	$\langle d \rangle$, нм	$\langle D \rangle$, нм	$\langle L \rangle$, нм	k_{fl}	ρ_H , $\times 10^{10}$ см $^{-1}$	n_x , $\times 10^{18}$ см $^{-3}$	I_B , а.у.	l	$\frac{\Delta^2}{n-l}$
a	20,1	0,230	—	1,31	—	0,0	$\rho_{L=}$ =7,95	8,7	8,9	6	7,6
b_1	19,0	0,267	3,53	2,38	—	0,18	17,0	1,6	0,0	8	0,715
b_2	18,7	0,275	3,85	3,83	—	0,21	15,7	2,7	9,0	8	1,84
c	18,5	0,253	4,59	2,41	1,27	0,0	13,2	2,0	9,5	8	1,80
d_1	19,1	0,265	3,48	2,72	0,39	0,10	16,3	3,8	6,0	9	0,688
d_2	18,9	0,268	3,97	3,97	1,41	0,0	15,1	2,1	10,5	9	1,87

Таблица 3.5. То же, что в табл. 3.4, но по усеченному спектру: фитирование данных только при $q < 0,38$ нм $^{-1}$, $n = 820$ [3]

Модель	$\langle R \rangle$, нм	σ	$\langle d \rangle$, нм	$\langle D \rangle$, нм	$\langle L \rangle$, нм	k_{fl}	ρ_H , $\times 10^{10}$ см $^{-1}$	n_x , $\times 10^{18}$ см $^{-3}$	I_B , а.у.	l	$\frac{\Delta^2}{n-l}$
a	20,1	0,230	—	1,28	—	0,00	$\rho_{L=}$ =7,95	9,0	7,7	7	8,1
b_1	19,0	0,266	3,64	2,29	—	0,17	15,7	1,6	2,5	8	0,727
b_2	19,2	0,236	4,06	3,66	—	0,00	17,5	3,3	14,2	8	0,748
c	19,5	0,249	4,59	2,82	0,60	0,05	12,7	2,6	14,0	8	0,897
d_1	19,0	0,265	3,47	2,87	0,78	0,10	12,7	2,6	9,0	9	0,734
d_2	18,9	0,268	3,97	3,97	1,41	0,00	15,1	2,1	10,5	9	0,7...

Выбор модели для бислоя Результаты фитирования для разных моделей плотности длины рассеяния показывают, что все модели дают близкие оценки для среднего радиуса везикул R и его относительного отклонения σ . Соответственно, отличия в описании спектра рассеяния при малых q несущественны (рис. 3.7). Различия в моделях, связанные с описанием внутренней структуры бислоя оболочки везикул, сказываются в области больших значений q .

На рис. 3.8 представлены, в сравнении с экспериментальными данными, расчетные кривые для разных моделей плотностей.

Одноступенчатый вариант функции плотности a (см. рис. 3.1) не моделирует поведение спектра в правой его части и дает сильно заниженные оценки для параметра толщины бислоя d (рис. 3.8, слева).

Варианты b_2 , c и d_2 (рис. 3.8, слева) обеспечивают качественное воспроизведение экспериментальной кривой, однако не могут описать всю амплитуду колебаний интенсивности в области больших значений q .

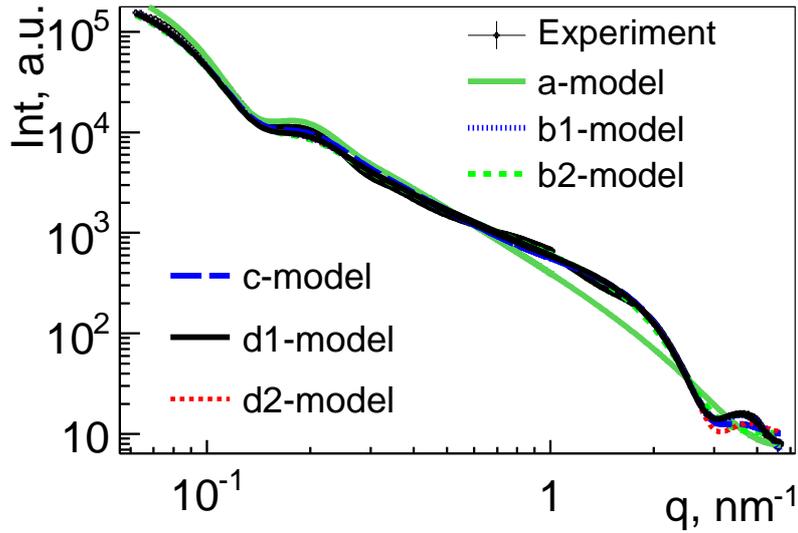


Рис. 3.7. Экспериментальный и теоретические спектры МУРР, соответствующие оптимальным параметрам модели РФФ с использованием различных функций плотности $\rho_C(x, \Theta_b)$. с учетом флуктуаций Θ_b (толщины мембраны). Количество экспериментальных точек $n = 894$ [3]

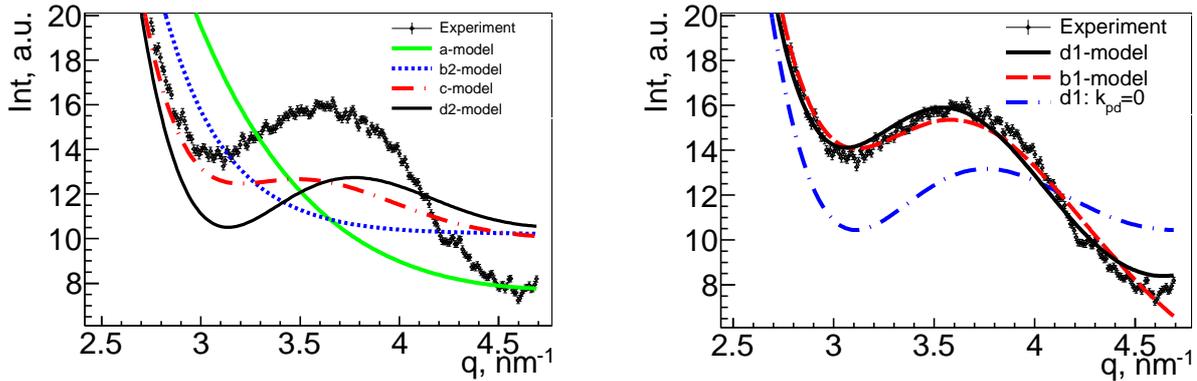


Рис. 3.8. Правая часть спектра, приведенного на рис. 3.7: $q > 2,5 \text{ nm}^{-1}$; слева: модели a , b_2 , c и d_2 ; справа: модели b_1 , d_1 и d_1 с $k_{\text{fl}} = 0$ [3]

Модели b_1 и d_1 (рис. 3.8, справа) имеют наилучшее согласие с экспериментальными данными, о чем можно судить как визуально, так и по значениям невязки $\Delta^2/(n - l)$ в оптимуме. При этом модель d_1 описывает также и намечающуюся тенденцию к возрастанию интенсивности при $q > 4,5 \text{ nm}^{-1}$, в то время как модель b_1 ее не описывает.

Для плотности d_1 , оптимальной с точки зрения соотношения «число варьируемых параметров — согласие с экспериментальными данными», расчеты сделаны с учетом и без учета полидисперсности бислоя ($k_{\text{fl}} = 0$). Оптимизация модели d_1 с $k_{\text{fl}} = 0$ (штрихпунктирная кривая на рис. 3.8, справа) показывает, что без введения флуктуаций внутренних параметров бислоя (3.14) описать различия в глубине минимумов на экспериментальной кривой в области больших значений q не удастся. Именно

учет флуктуаций толщины липидного бислоя обеспечивает воспроизведение моделью экспериментальных данных в правой части спектра.

Интерпретация полученных оценок и сравнение с результатами МУРН-экспериментов. Толщина бислоя и средний радиус везикул в наших расчетах получились ниже по сравнению с оценками, сделанными в [117] на базе МУРН⁸ для везикул ДМФХ в D₂O. С другой стороны, в работе [119] показано, что значения этих параметров должны уменьшаться с ростом концентрации сахарозы в растворе. Поэтому в целом можно считать, что полученные нами параметры везикулярной системы ДМФХ в 40% растворе сахарозы не противоречат результатам предыдущих работ. Для уточнения оценки параметров, отвечающих за структуру везикул, необходимы систематические расчеты по МУРР- и МУРН-спектрам, измеренным с разной концентрацией сахарозы.

3.4. Результативность АДЭ по сравнению с другими методами оптимизации

Вид и особенности профиля целевой функции. На рисунке 3.9 в качестве примера показано, что типичная картина профиля невязки Δ^2 в зависимости от параметров d и D (L — зафиксировано) для d_2 -модели даже при фиксированных значениях остальных параметров имеет несколько минимумов, часто овражного типа. Во всех приведенных ниже оценках метод

⁸см. сноску на с. 91.

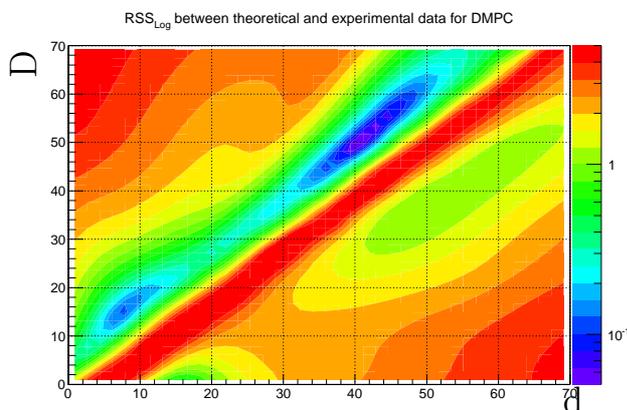


Рис. 3.9. Модель d_2 . Невязка Δ^2 с весами, пропорциональными квадратам ошибок экспериментальных данных. Двумерный (D и d — варьируемые параметры) профиль 9-мерной целевой функции (3.23). Остальные параметры зафиксированы: $R = 192,0$, $m = 13,79$, $L = 5,26$, $\rho_H = 14,9 \cdot 10^{10}$, $n_x = 2,4 \cdot 10^{18}$, $I_{\mathbf{B}} = 11,48$ $k_H = 0$

Таблица 3.6. Границы области поиска и координаты минимума для модели d_1 с учетом флуктуаций Θ_b . Вместо D варьировался параметр $2h = d - D$

Мо- дель	R, нм	(m)	d, нм	2h, нм	L, нм	k_{fluct}	$\rho_H,$ $\times 10^{10}$ см $^{-1}$	$n_x,$ $\times 10^{18}$ см $^{-3}$	$I_B,$ a.u.
Θ_i^{\min}	18,0	11.0	3,00	10^{-5}	10^{-5}	0,1	9,0	10^{-4}	-10,0
Θ_i^{\max}	40,0	55.0	6,50	2,50	2,20	0,5	25,0	80,0	17,0
$d_1: \Theta^*$	19,1 $\pm 0,2$	13,2 $\pm 0,8$	3,42 $\pm 0,08$	0,65 $\pm 0,14$	0,41 $\pm 0,03$	0,1001 $\pm 0,001$	17,4 $\pm 1,4$	3,76 $\pm 0,07$	5,3 $\pm 0,3$

АДЭ с вероятностью, близкой к единице, определяет наиболее глубокий минимум для заданной модели.

Отметим здесь, что в случаях, когда второй и, возможно, следующие, минимумы имеют сопоставимое с главным значение Δ^2 и два соответствующих набора параметров обеспечивают неплохое согласие с экспериментальными данными, необходимы дополнительные критерии выбора параметров (например, из других экспериментов известно, что $d - D = 2h$ имеет порядок $2 \cdot 0,9 = 1,8$ нм).

Поэтому значения параметров Θ , обеспечивающие минимум Δ^2 -функционала (3.23), определяются с помощью алгоритма АДЭ [1, 7], описанного в главе 1.

Параметры, с которыми проводилось сравнение вероятностей сходимости методов АДЭ, SIMPLEX и MIGRAD. На примере анализа d_1 -модели с учетом флуктуаций толщины мембраны оптимизация проводилась многократно при помощи различных оптимизационных методов: метода симплексов (метод Нелдера – Мида) [53], квазиньютоновского метода Дэвидона – Флетчера – Пауэлла [30, 31, 37, 38, 55] (методы SIMPLEX и MIGRAD из Root-реализация программы Minuit [46]) и различных стратегий методов АДЭ [6, 7, 9] и АДЭ с самоадаптирующимся кроссовером — АДЭ-АКМ [10].

Расчеты проводились с учетом того, что допустимые значения параметров лежат в области $\Theta_i \in [\Theta_i^{\min}, \Theta_i^{\max}]$. Значения Θ_i^{\min} и Θ_i^{\max} представлены в строках (2) и (3) табл. 3.6. Точки для начального приближения равномерно распределены в этой области. Метод АДЭ использовался в двух вариантах. В варианте, предложенном в [6, 7, 9] со стратегией rand/rand/1/bin/. В этом случае были установлены значения для масшта-

бирующего фактора $F = 0,9$ и для кроссовера $C_r = 0,9$. (это часто применяемая комбинация при расчетах с помощью ДЭ). Использовались два значения начальной популяции: $N_p^{\min}=10$ — это минимальный рекомендуемый размер популяции; и $N_p^{\min}=40$ — это «оптимальный» размер популяции ДЭ для 10-мерной функции Розенброка (см.табл. 1.1). Вторым вариантом — с адаптивным кроссовером [10], со стратегиями `rand/rand/1/asm` и `linworst/current-to-pbest/1/asm`. Критерий рестарта (1.14): $\Delta x_j = 10^{-12}$. Требуемая точность локализации минимума (критерий выхода) (1.15): $\Delta f = 10^{-2}$. В настройках SIMPLEX и MIGRAD увеличено (по сравнению с настройками по умолчанию) максимально допустимое количество вычислений функций ($N_{FE}^{\max}=100\,000$) — этого более чем достаточно для того, чтобы эти программы (SIMPLEX, MIGRAD) диагностировали в рассматриваемой задаче окончание расчетов до достижения N_{FE}^{\max} . Расчет каждым из методов повторялся $N = 100$ раз.

Найденные решения. Глобальный минимум задачи найден в точке Θ^* , координаты которой представлены в строке (4) табл. 3.6. (ошибки оценены исходя из условия $((\Delta^2)^* + 1)$). Значение целевой функции в этой точке: $f^*(\Theta^*) = 608,46$. Количество экспериментальных точек $n = 894$. Таким образом, стандартному отклонению будет соответствовать отклонение значения целевой функции от глобального минимума на величину около $f^* - f_{min}^* = 0,7$.

Сравнение результативности метода АДЭ, метода симплексов (Нелдера – Мида) и квазиньютоновского метода Дэвидона – Флетчера – Пауэлла). На рис. 3.10 показано количество попыток, завершившихся со значением целевой функции (см. (3.23)), превышающим значение целевой функции в глобальном минимуме на величину, не большую, чем отложенное по горизонтальной оси значение для различных использованных методов. В табл. 2.5 приведены обобщенные результаты, полученные каждым из методов. Вероятность нахождения решения методом SIMPLEX оказалась равной 2-3% (см. столбец 3 табл. 3.7). Чаще всего этот метод диагностировал сходимость в точках, далеких от глобального минимума и прекращал дальнейшие расчеты (о предупреждении преждевременной сходимости в методе АДЭ см. раздел 1.2). Вероятность нахождения верных оценок для параметров решаемой задачи методом MIGRAD оказалась близкой к 13%. При этом метод MIGRAD заканчивал расчеты

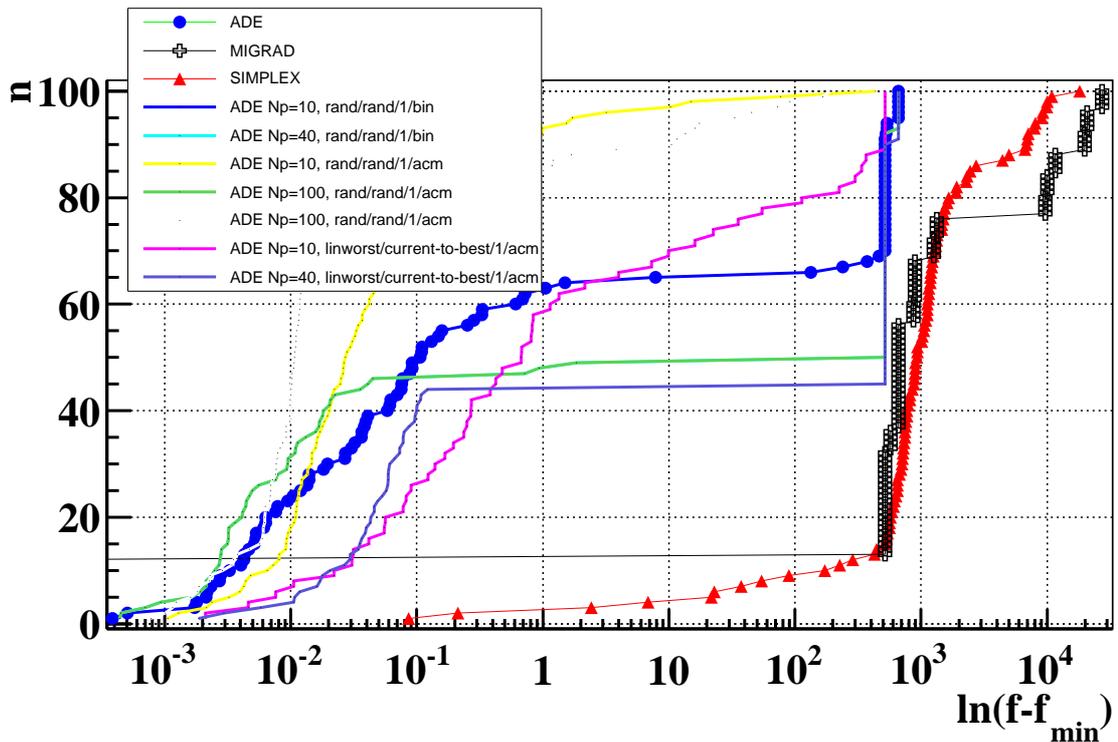


Рис. 3.10. Процент попыток, в которых значение целевой функции в найденном решении (f_i^*) превышает значение целевой функции в глобальном минимуме (f_{min}^*) на величину, меньшую, чем значение, отложенное по горизонтальной оси

Таблица 3.7. Процент попыток с соответствующим результатом для трех методов оптимизации. † — MIGRAD TERMINATED WITHOUT CONVERGENCE во всех этих попытках. * — среди несошедшихся попыток 4(четыре) достигли установленного предела $\max N_{FE} = 300000$. ** — обе несошедшихся попытки достигли установленного предела $\max N_{FE} = 400000$. Среди сошедшихся 9 (девять) попыток закончили с $N_{FE} > 300000$. *** — все 5 (пять) несошедшихся попыток достигли установленного предела $\max N_{FE} = 400000$. Цифры в скобках — число сошедшихся попыток в пределах $\max N_{FE} = 300000$. В этом случае также все 17 несошедшихся попыток достигли установленного предела $\max N_{FE} = 300000$ [3]

$f^* - f_{min}^*$, не бо- лее	MIG- RAD	SIMP- LEX	АДЭ		АДЭ-АКМ				
			rand/rand/ 1/bin		rand/rand/1/acm			linworst/cur- rent-to-pbest /1/acm	
			$N_p^i=10$	$N_p^i=40$	$N_p^i=10$	$N_p^i=40$	$N_p^i=100$	$N_p^i=10$	$N_p^i=40$
$\sigma_f =$ $= 0,7$	13	2	61	46	90*	98**	95 (83)***	52	44
1	13	2	62	48	93*	98**	90 (85)***	58	44
$3\sigma_f =$ $= 2,1$	13	3	64	49	95*	99**	96 (86)***	66	44
$\langle N_{FE} \rangle$	2980 ± 200	2045 ± 60	94300 ± 7800	106900 ± 7500	97230 ± 8700	171000 ± 11000	263200 ± 3200	22400 ± 1200	29200 ± 2300
Всего попы- ток	100 †	100	100	100	100	100	100	100	100

как близко, так и далеко от минимума, диагностировав «окончание расчетов без сходимости». В то же время все протестированные стратегии АДЭ сходились к глобальному минимуму с вероятностью большей 50%, а наиболее удачные (rand/rand/1/acm) — с вероятностью больше 90%. При этом эта вероятность могла быть еще увеличена путем увеличения максимально допустимого количества вычислений функции.

Использование адаптивного кроссовера позволяет существенно увеличить вероятность нахождения глобального минимума. Но несмотря на то, что использование «ускоренной» стратегии «linworst/current-to-pbest» сокращает среднее количество вычислений функции приблизительно в три раза, эта стратегия имеет бóльшую вероятность закончить вычисления в одном из локальных минимумов.

Таким образом, при решении данной задачи метод АДЭ предпочтительнее классических методов поиска по трем основным причинам:

- Во-первых, метод АДЭ-Р является робастным методом поиска *глобального* минимума. Он с большой вероятностью завершает расчеты в одном из минимумов задачи, в то время как методы SIMPLEX и MIGRAD заканчивали поиск в одном из минимумов лишь с вероятностью $2 \pm 1,4\%$ и $13 \pm 3,7\%$ соответственно. Вероятность нахождения глобального минимума методом АДЭ-Р в зависимости от выбранной стратегии составила от $p_{lca40} = 44 \pm 5\%$ (для стратегии DE/linworst/current-to-pbest/1/acm, $N_p = 40$) до $p_{rra10} = 98 \pm 2\%$ (для стратегии DE/rand/rand/acm, $N_p = 10$).
- С другой стороны, при необходимости для нахождения второстепенных минимумов можно использовать предварительные результаты, выдаваемые методом АДЭ в процессе расчетов.
- Являясь прямым (без использования производных) методом поиска, АДЭ продолжает расчеты и находит решения в ситуациях, когда квази-ньютонский метод Давидона – Флетчера – Пауэлла диагностирует невозможность дальнейших поисков.

3.5. Оценка ускорения АДЭ при параллельных вычислениях

Комплекс программ (описанный в 3.2) подготовлен для компиляции и запуска на параллельном кластере Центрального информационно-вычислительного комплекса (ЦИВК) ОИЯИ. Для распараллеливания мо-

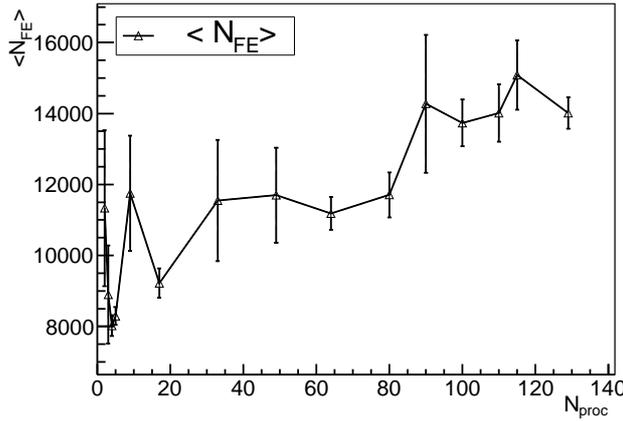


Рис. 3.11. Среднее количество вычислений функции, понадобившееся для достижения заданной точности при запуске на разном количестве параллельных процессоров: $\langle N_{FE} \rangle (N_{proc})$

для (программы) `dmpc_mpi_master.cpp` использована технология MPI (см. раздел 1.5.2). Эффективность параллельной MPI-реализации алгоритма АДЭ тестировалась на LINUX-кластере ЛИТ ОИЯИ.

Расчет d_1 -модели, имеющей максимальное число варьируемых параметров $l = 9$, запускался многократно с одними и теми же начальными параметрами алгоритма и границами начальных и допустимых значений на разном количестве вычислительных узлов: $N_{proc} \in [2, 134]$. В представленной статистике каждая комбинация запускалась $N = 10$ или 20 раз.

Так как алгоритм АДЭ недетерминированный, количество вычислений функции, необходимое для достижения определенной точности, есть случайная величина. На рис. 3.11 представлено среднее количество вычислений функции, понадобившееся для достижения заданной точности при запуске на разном количестве параллельных процессоров.

На рис. 3.12 показано общее среднее время, затраченное на вычисления. Видно, что оно падает обратно пропорционально количеству задействованных вычислительных узлов.

Рис. 3.13 демонстрирует близкое к линейному ускорение параллельных вычислений (speed-up) вплоть до $N_{proc} = 128$ параллельно задействованных вычислительных узлов. Сплошной линией показано результирующее ускорение. Ускорение вычислений с учетом возрастания необходимого количества вычислений функции при $N_{proc} = 128$ составило 98 ± 21 раз. Пунктирной линией показано ускорение вычислений в расчете на одно вычисление целевой функции (проведена коррекция на рост необходимого числа вычислений функции для достижения критерия останова).

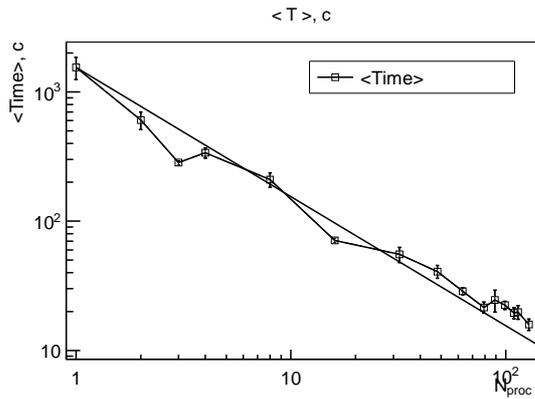


Рис. 3.12. Среднее время, затраченное на вычисления в зависимости от числа задействованных процессоров

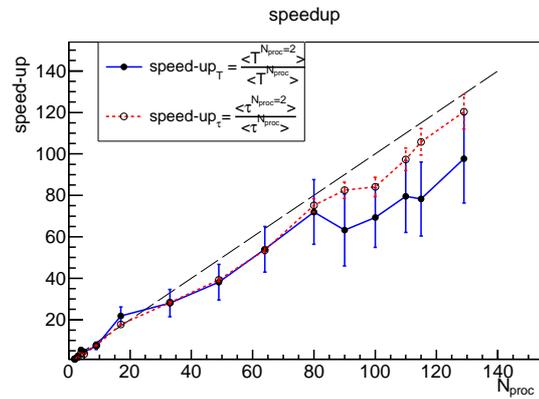


Рис. 3.13. Ускорение параллельных вычислений (speed-up) в зависимости от числа задействованных процессоров. Сплошная — результирующее ускорение. Пунктир — ускорение в расчете на одно вычисление целевой функции (коррекция на рост необходимого числа вычислений функции для достижения критерия останова $\varepsilon_f = 10^{-5}$)

Таким образом, можно сформулировать следующие **основания** для того, чтобы отдать предпочтение методу АДЭ по сравнению с классическими методами минимизации:

- Решаемая задача имеет размерность от 6 до 9 варьируемых параметров.
- Целевая функция — невязка — является **многомодальной**. Стандартные алгоритмы локального поиска (SIMPLEX, MIGRAD) требуют хорошего начального приближения, в противном случае не находят решения.
- Расчет целевой функции нетривиален, требует значительных временных затрат. При планируемом увеличении детализации модели число варьируемых параметров и время расчета невязки может значительно возрасти. Следовательно, распараллеливание расчетов сократит время ожидания результата.

Основные результаты, представленные в главе 3

В этой главе представлены следующие результаты.

- Реализован комплекс программ для нахождения параметров полидисперсных везикулярных систем по данным малоуглового синхротронного рассеяния.
- Модифицирована модель разделенных формфакторов: учтены флуктуации липидного бислоя. Смоделированный разброс значений параметров за счет флуктуаций толщины мембраны позволил улучшить

согласие рассчитанного спектра с экспериментальными данными в области значений вектора рассеяния от 3 до 4,5 нм⁻¹. Полученные значения оценки качества подгонки $\frac{\Delta^2}{n-k}$ улучшились для модели d_1 более чем в два раза (0,688 в случае учета флуктуаций против 1,68 в случае, когда флуктуации не учитываются).

- С использованием АДЭ проведено численное исследование полидисперсных везикулярных систем, состоящих из однослойных везикул ДМФХ.
 - Исследован ряд моделей для описания плотности длины рассеяния внутри мембраны. Получено, что
 - * Однопараметрическая одноступенчатая модель (а-модель) обеспечила согласующуюся с другими моделями оценку для радиуса R везикулы и параметра m полидисперсности, однако не позволила корректно оценить толщину липидного бислоя.
 - * Двухпараметрические (b_1 - и b_2 -модели) и трехпараметрическая (c -модель) двухступенчатые модели (соответственно трех- и четырехпараметрические в случае учета флуктуаций толщины мембраны) позволили, помимо R и m , оценить толщину мембраны d . Среди этих моделей наилучшее согласие с экспериментальными данными демонстрирует модель b_1 .
 - * Из трехпараметрических (четырёхпараметрических в случае учета флуктуаций толщины мембраны) трехступенчатых моделей (d_1 - и d_2 -модели) наилучшей является модель d_1 с учетом флуктуаций толщины мембраны. Она позволила описать экспериментальный спектр с относительной ошибкой не более 8,3%.
 - * Из сравнения с экспериментальными данными рассеяния гамма-излучения на растворе везикул ДМФХ в сахарозе можно сделать вывод о реалистичности предложенных линейных моделей описания плотности длины рассеяния в липидном бислое. Линейность моделей позволила получить выражения для фактора липидного бислоя в явном виде.
 - Получены оценки параметров везикулярных систем, которые качественно не противоречат результатам других экспериментов.
- Реализована параллельная версия комплекса программ для расчета параметров моделей полидисперсных везикулярных систем по данным

МУРР, которая в настоящее время используется для дальнейших исследований [121, 122, 128]

- Эффективность MPI-реализации методов АДЭ и АДЭ-Р протестирована на LINUX-кластере ОИЯИ на примере расчетов структуры везикул ДМФХ. Ускорение близко к N_{proc} для $N_{\text{proc}} \sim 5N_p \sim 80$. При дальнейшем увеличении числа параллельно задействованных процессоров стало сказываться возрастание необходимого количества вычислений целевой функции для достижения требуемой точности. При этом ускорение расчетов (с учетом возрастания необходимого количества вычислений функции для достижения требуемой точности) составило 98 ± 21 при $N_{\text{proc}} = 128$. Скорость вычислений (в расчете на одно вычисление целевой функции) возросла в 118 ± 23 раза, то есть ускорение вычислений оказалось равным $0,91N_{\text{proc}}$.
- Использование АДЭ-минимизатора позволило эффективно распараллелить вычисления, требующие глобальной минимизации многоэкстремальной функции с 9 частично коррелированными переменными и значительных затрат компьютерного времени, более чем в 50 раз уменьшив время расчетов, увеличив при этом вероятность определения глобального минимума с 15% (с использованием минимизатора на основе квазиньютоновского метода) до 90% [19].

По материалам главы автором опубликованы работы [3–5, 14, 19–21].

Заключение

В диссертации разработаны, обоснованы, реализованы в виде программного комплекса и численно исследованы методы асинхронной дифференциальной эволюции (АДЭ) и АДЭ с кроссовером, задаваемым адаптивной корреляционной матрицей. С использованием перечисленных методов проведено численное исследование двух многопараметрических систем.

1. Разработаны: (i) метод АДЭ [7], (ii) метод АДЭ с рестартом [9, 11] и адаптивной корреляционной матрицей [10], а также проблемно-ориентированные комплексы программ.
2. Реализованный на базе параллельного минимизатора АДЭ комплекс программ для определения параметров пион-нуклонной амплитуды упругого рассеяния заряженных пионов на нуклонах ядерной среды позволил описать экспериментально измеренные дифференциальные сечения упругого рассеяния пионов на ядрах в рамках трехпараметрической модели микроскопического оптического потенциала и релятивистского волнового уравнения Шредингера. Полученный набор значений параметров «in-medium» амплитуды рассеяния π -мезонов на внутриядерных нуклонах позволил сделать ряд теоретических выводов о характере рассеяния пионов внутри ядра, он используется для моделирования неупругого пион-ядерного рассеяния.
3. Реализованный на базе параллельного минимизатора АДЭ комплекс программ для расчета параметров полидисперсных везикулярных систем позволил получить структурные параметры везикулярной системы ДМФХ в 40% растворе сахарозы и в совокупности с данными нейтронного рассеяния сделать выводы о влиянии сахарозы на структуру везикулярной системы.
4. Аналитическое и численное исследование эффективности методов АДЭ и основанных на них комплексов программ выявило основные преимущества методов АДЭ по сравнению с другими методами минимизации. Эти преимущества состоят (i) в увеличении скорости и вероятности сходимости по сравнению с ДЭ; (ii) в улучшении показателей вероятности сходимости по сравнению с другими методами минимизации, особенно

для многоэкстремальных задач с частично разделяющимися переменными и (iii) в возможности эффективного применения разработанных методов в режиме параллельных вычислений.

Все это говорит о перспективности использования методов АДЭ и АДЭ-АКМ для расчетов как в последовательном режиме, так и на параллельных вычислительных системах для решения широкого круга задач, в том числе задач, выходящих за рамки данной работы.

Список публикаций по теме диссертации

1. В российских рецензируемых изданиях

1. **Жабицкая, Е. И., Жабицкий, М. В., Земляная, Е. В., Лукьянов, К. В.** Расчет параметров микроскопического оптического потенциала упругого рассеяния π -мезонов на ядрах с применением алгоритма асинхронной дифференциальной эволюции // Компьютерные исследования и моделирование. — 2012. — Т. 4, № 3. — С. 585–595.
2. **Лукьянов, В. К., Земляная, Е. В., Лукьянов, К. В., Жабицкая, Е. И., Жабицкий, М. В.** Моделирование пион-ядерного микроскопического оптического потенциала при энергиях (3,3)-резонанса и влияние ядерной среды на пион-нуклонную амплитуду // Ядерная физика. — 2014. — Т. 77, № 1. — С. 103–112 // Препринт ОИЯИ Р4-2012-105. — 2012. — Дубна, ОИЯИ // ENGL: *Lukyanov, V. K., Zemlyanaya, E. V., Lukyanov, K. V., Zhabitskaya, E. I. and Zhabitsky, M. V.* Modeling of a Microscopic Optical Pion–Nucleon Potential at Energies in the (3, 3)-Resonance Region and Nuclear-Matter Effect on the Pion–Nucleon Amplitude // Physics of Atomic Nuclei. — 2014. — V. 77, No 1. — Pp. 100–109.
3. **Жабицкая, Е. И., Жабицкий, М. В., Киселев М. А.** Исследование структуры однослойных везикул DMPC с использованием параллельного алгоритма асинхронной дифференциальной эволюции // Вестник Российского университета дружбы народов. Серия: Математика. Информатика. Физика. — 2014. — № 2. — С. 253–259.
4. **Киселев, М. А., Земляная, Е. В., Жабицкая, Е. И., Аксенов, В. Л.** Исследование однослойных везикул ДМФХ в водных растворах сахарозы методами малоуглового рассеяния нейтронов и рентгеновских лучей // Кристаллография. — 2015. — Т. 60. — С. 140–145 // Препринт ОИЯИ Р3-2014-34. — 2014. — Дубна, ОИЯИ // ENGL: *Kiselev, M. A., Zemlyanaya, E. V., Zhabitskaya, E. I. and Aksenov, V. L.* Investigation of the Structure of Unilamellar Dimyristoylphosphatidylcholine Vesicles in Aqueous Sucrose Solutions by Small Angle Neutron and X Ray Scattering // Crystallography Reports. — 2015. — V. 60, No 1. — Pp. 143–147.
5. **Жабицкая, Е. И., Земляная, Е. В., Киселев, М. А.** Численный ана-

лиз экспериментальных данных по синхротронному рассеянию на полидисперсных везикулярных системах с использованием метода асинхронной дифференциальной эволюции // Математическое моделирование. — 2015. — Т. 27, № 7. — С. 58–64.

6. **Жабицкая, Е.И., Жабицкий, М.В.** Решение оптимизационных задач на распределенных вычислительных системах с помощью алгоритма асинхронной дифференциальной эволюции // Математическое моделирование. — 2012. — Т. 24, № 12. — С. 33–37.

В иностранных рецензируемых изданиях

7. **Zhabitskaya, E.I., Zhabitsky, M.V.** Asynchronous Differential Evolution // Mathematical Modeling and Computational Science (MMCP-2011). — Lecture Notes in Computer Science, Springer, Berlin, Heidelberg. — 2012. — 7125. — Pp. 328–333.
8. **Zhabitskaya, E.I.** Constraints on Control Parameters of Asynchronous Differential Evolution // Mathematical Modeling and Computational Science (MMCP-2011). — Lecture Notes in Computer Science, Springer, Berlin, Heidelberg. — 2012. — 7125. — Pp. 322–327.
9. **Zhabitskaya, E., Zhabitsky M.** Asynchronous Differential Evolution with Restart // In I. Dimov, I. Farag, and L. Vulkov (Eds.): Numerical Analysis and its Applications (NAA'2012). — Lecture Notes in Computer Science, Springer, Berlin, Heidelberg. — 2013. — 8236. — Pp. 555–561.
10. **Zhabitskaya, E.I., Zhabitsky M.V.** Asynchronous Differential Evolution with Adaptive Correlation Matrix // Proceeding of the 15th Annual Conference on Genetic and Evolutionary Computation, USA, New York. — 2013. — Pp. 455–462.

Другие публикации, материалы и тезисы конференций

11. **Жабицкая, Е.И., Жабицкий, М.В.** Алгоритм асинхронной дифференциальной эволюции для решения оптимизационных задач // Шестнадцатая научная конференция молодых ученых и специалистов. Труды конференции. г. Дубна, 6–11 февраля 2012 г. Объединение молодых учёных и специалистов ОИЯИ (ОМУС XVI). — 2012. — 286 с. — С. 50–53.
12. **Жабицкий, М.В., Жабицкая, Е.И.** Решение оптимизационных задач на распределенных вычислительных системах с помощью алгоритма асинхронной дифференциальной эволюции // Современные про-

- блемы прикладной математики и информатики (МРАМС'2012): Тезисы докладов международной молодежной конференции-школы (Дубна, 22–17 августа 2012 г.), 230 с. — Дубна. — 2012. — С. 107–109.
13. **Жабицкая, Е. И., Жабицкий, М. В.** Эффективные стратегии параллельной реализации метода асинхронной дифференциальной эволюции с рестартом // Современные проблемы прикладной математики и информатики (МРАМС'2012): Тезисы докладов международной молодежной конференции-школы (Дубна, 22-27 августа 2012 г.) — Дубна, ОИЯИ. — 2012. — 230 с. — С. 110–112
 14. **Жабицкая, Е. И., Жабицкий М. В.** Развитие алгоритма асинхронной дифференциальной эволюции // Научный отчет за 2012-2013 гг. Лаборатории информационных технологий // ОИЯИ 2014-26, Дубна. — 2014. — 265 с. — С. 144–147.
 15. **Жабицкая, Е.И.** Применение асинхронной дифференциальной эволюции для моделирования микроскопического оптического потенциала упругого рассеяния пи-мезонов на ядрах // Шестнадцатая научная конференция молодых ученых и специалистов. Труды конференции. г. Дубна, 6–11 февраля 2012 г. Объединение молодых учёных и специалистов ОИЯИ (ОМУС XVI), — Дубна, ОИЯИ. — 2012. — 286 с. — С. 46–49.
 16. **Жабицкая, Е. И., Жабицкий, М. В., Земляная, Е. В., Лукьянов, К. В.** Применение метода асинхронной дифференциальной эволюции для оценки параметров микроскопического оптического потенциала упругого рассеяния π^+ на ядрах // Тезисы XIX международной конференции «Математика. Компьютер. Образование» (МКО-2012). — Дубна, ОИЯИ. — 2012. — С.139.
 URL:www.mce.su/rus/archive/abstracts/mce19/authors/person21631/doc151854/ ;
 Calculation of the parameters of microscopic optical potential by asynchronous differential evolution algorithm. — Там же, с. ;31.
 17. *Lukyanov, V. K., Zemlyanaya, E. V., Zhabitskaya, E. I., Lukyanov, K. V., Zhabitsky, M. V.* Study of In-Medium Effect on the Pion-Nucleon Amplitude from Analysis of Pion-Nucleus Data within the Microscopic Optical Potentia // XXI International Baldin Seminar on High Energy Physics Problems, JINR, Dubna, Russia //Proceedings of Science, SISSA
 URL:http://pos.sissa.it/archive/conferences/173/021/Baldin\%20ISHEPP\%20XXI_021.pdf
 18. *Zemlyanaya, E. V., Lukyanov, V. K., Lukyanov, K. V., Zhabitskaya, E. I.,*

- Zhabitsky, M. V.* Pion-Nucleus Microscopic Optical Potential at Intermediate Energies and In-Medium Effect on the Elementary πN Scattering Amplitude // 31st International Workshop on Nuclear Theory, INRNE BAS, Sofia, Bulgaria, Rila Mountains, Bulgaria. — Heron press Ltd, Sofia, Nuclear Thoery. — 2012. — V. 31. — Pp. 175-184. — URL:<http://arxiv.org/abs/1210.1069>
19. ***Жаблицкая, Е.И.*** Анализ сложных везикулярных систем с использованием метода асинхронной дифференциальной эволюции // Современные проблемы прикладной математики и информатики (МРАМС-2014): Тезисы докладов международной молодежной конференции-школы (Дубна, 25–29 августа 2014 г.) — Дубна, ОИЯИ. — 2014.— 173 с. — С. 71–74.
20. ***Жаблицкая, Е.И., Земляная, Е.В., Киселев, М.А.*** Применение метода асинхронной дифференциальной эволюции при исследовании структуры однослойных везикул DMPC // Тезисы XXI международной конференции "Математика. Компьютер. Образование"(МКО-2014). — Дубна, ОИЯИ. — www.mce.su/archive/doc186896/rus.pdf
21. ***Zhabitskaya, E., Zemlyanaya, E., Kiselev, M.*** Unilamellar Vesicles Structure Analysis Using Parallel Asynchronous Differential Evolution // Mathematical Modeling and Computational Physics (ММСП'2013): Book of Abstracts of the International Conference (Dubna, July 8–12, 2013). — Dubna, JINR. — 2013.

Список цитируемой литературы

1. **Методы оптимизации**
22. Городецкий, С.Ю., Гришагин, В.А. Учебный курс «Модели и методы конечномерной оптимизации». Часть 2. Нелинейное программирование и многоэкстремальная оптимизация // Нижний Новгород. — 2003.
23. Карпенко, А.П. Методы оптимизации // МГТУ им. Н.Э. Баумана, учебный пакет. — 2007. — URL:<http://bigor.bmstu.ru/?pkg/OSORROXOZUMMRTWMA500>.
24. Нестеров, Ю.Е. Введение в выпуклую оптимизацию // МЦНМО. — 2010. — 280 с.
25. Трифонов, А.Г. Optimization Toolbox 2.2. Руководство пользователя // URL:http://matlab.exponenta.ru/optimiz/book_1/index.php.
26. Bäck, T., Fogel, D.B. and Michalewicz, Z. (Eds), Handbook of Evolutionary Computation // Oxford University Press. — 1997.
27. Broyden, C.G. Quasi-Newton Methods and their Application to Function Minimization // Math. Comput. — 1967. — 21. — P. 368.
28. Broyden, C.G. The Convergence of a Class of Double-Rank Minimization Algorithms // J. Inst. Maths. Applics. — 1970. — 6. — Pp. 76–79.
29. Colorni, A., Dorigo, M. et Maniezzo, V. Distributed Optimization by Ant Colonies, Actes De La Première Conférence Européenne Sur La Vie Artificielle // Paris, France, Elsevier Publishing. — 1991. — Pp. 134–142.
30. Davidon, W.C. Variable Metric Methods for Minimization // A.E.C. Res. and Develop. Report ANL-5990. Argonne National Laboratory. Argonne; Illinois. — 1959. — P. 21.
31. Davidon, W.C. Variance Algorithm for Minimization // Comput. J. — 1968. — 10. — P. 406.
32. Davidon, W.C. Variable Metric Method for Minimization // SIAM Journal on Optimization. — 1991. — 1. — Pp. 1–17.
33. Dixon, L.C.W. Non-Linear Optimization // English Universities Press, London. — 1972.
34. Dorigo, M. Optimization, Learning and Natural Algorithms // PhD thesis, Politecnico di Milano, Italie. — 1992.
35. Eiben, A.E. and Smith, J.E. Introduction to Evolutionary Computing //

- Springer. — 2003.
36. Fletcher, R. A New Approach to Variable Metric Algorithms // *Comput. J.* — 1970. — 13. — P. 317.
 37. Fletcher, R. and Powell, M.J.D. A Rapidly Converging Descent Method for Minimization // *Comput. J.* — 1963. — 6. — Pp. 163–168.
 38. Fletcher, R. and Reeves, C.M. Function Minimization by Conjugate Gradients // *Comput. J.* — 1964. — 7. — P. 149.
 39. Fogel, D.B. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence* // IEEE Press, Piscataway, NJ. — 1995.
 40. Goldfarb, D. A Family of Variable Metric Methods Derived by Variational Means // *Mathematics of Computation.* — 1970. — No 24. — Pp. 23–26.
 41. Hansen, N., Finck, S., Ros, R. and Auger, A. Real-Parameter Black-Box Optimization Benchmarking 2012: Experimental Setup // In Technical report, INRIA, 2012.
 42. Holland, J.H. *Adaptation in Natural and Artificial Systems* // University of Michigan Press, Ann Arbor. — 1975.
 43. Hooke, R. and Jeeves, T.A. Direct Search Solution of Numerical and Statistical Problems // *J. Assoc. Comput. Mach.* — 1961. — 8. — P. 212.
 44. James, F. Function Minimization // *Proceedings of the 1972 CERN Computing and Data Processing School, CERN.* — 1972. — Pp. 72–21
 45. James, F., Roos, M. Minuit — a System for Function Minimization and Analysis of the Parameter Errors and Correlations // *Computer Physics Communications.* — 1975. — V. 10, Iss. 6. — Pp. 343–446.
 46. James, F. Winkler, M. MINUIT User's Guide // CERN, Geneva. — June 16. — 2004.
 47. De Jong, K.A. *Evolutionary Computation: A Unified Approach* // MIT Press, Cambridge MA. — 2006.
 48. Kowalik, J., Osborne, M.R. *Methods for Unconstrained Optimization Problems* // American Elsevier Publishing Co., Inc., New York. — 1968.
 49. Kennedy, J., Eberhart, R. and Shi, Y. *Swarm Intelligence* // Morgan Kaufmann, US. — 2001.
 50. Kennedy, J. and Eberhart, R.C. Particle Swarm Optimization // *Proceedings of IEEE International Conference on Neural Networks IV.* — 1995. — Pp. 1942–1948.
 51. Lera, D., Sergeyev, Ya.D. Lipschitz and Holder Global Optimization Using

- Space-Filling Curves // Applied Numerical Mathematics. — 2010. — 60(1-2). — Pp. 115–129.
52. Metropolis, N., Ulam, S. The Monte Carlo Method // Journal of the American Statistical Association. — 1949. — 44. № 247. — Pp. 335–341.
 53. Nelder, J.A., Mead, R. A Simplex Method for Function Minimization // Computer J. — 1965. — 7. — Pp. 308–313.
 54. Powell, M.J.D. Rank One Methods for Unconstrained Optimization, Appearing in Integer and Non-Linear Programming // North-Holland Publ. Co., Amsterdam. — 1970.
 55. Powell, M.J.D. An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives // Comput. J. — 1964. — 7. — P. 155.
 56. Poli, R. An Analysis of Publications on Particle Swarm Optimisation Applications // Technical Report CSM-469 (Department of Computer Science, University of Essex, UK). — 2007.
 57. Poli, R. Analysis of the Publications on the Applications of Particle Swarm Optimisation // Journal of Artificial Evolution and Applications. — 2008. — P. 1.
 58. Price, W.L. A Controlled Random Search Procedure for Global Optimization // Computer J. — 1977. — 20. — Pp. 367–370.
 59. Rosenbrock. H.H. An Automatic Method for Finding the Greatest or Least Value of a Function // Comput. J. — 1960. — 3. — P. 175.
 60. Shanno, D.F. Conditioning of Quasi-Newton Methods for Function Minimization // Mathematics of Computing. — 1970. — V. 24. — Pp. 647–656.
 61. Shi, Y., and Eberhart, R.C., A Modified Particle Swarm Optimizer // Proceedings of IEEE International Conference on Evolutionary Computation. — 1998. — Pp. 69–73.
 62. Sean Luke, Essentials of Metaheuristics // Lulu, 2009, available for free at <http://cs.gmu.edu/~sean/book/metaheuristics/>. // Люк Ш. Основы метаэвристик. — 2009. — URL: <http://qai.narod.ru/GA/metaheuristics.html>
 63. Whitley, Darrell and Kauth, Joan A Different Genetic Algorithm // GENITOR: Technical Report CS-88-101, Colorado State University. — 1988.

64. Beyer, H.G. On the Dynamic of EAs without Selection // W.Banzaf, C.Reeves (eds.), *Foundation of Genetic Algorithms*, Morgan Kaufmann. — 1999. — Pp. 5–26.
65. Das, S., Suganthan, P.N. Differential Evolution: A Survey of the State-of-the-Art // *IEEE Trans. Evol. Comput.* — 2011. — 15. — Pp. 4–31.
66. Lampinen, J. A Bibliography of Differential Evolution Algorithms // Technical Report. Lappeenranta University of Technology, Department of Information Technology, Laboratory of Information Processing, Oct. 16, 1999 (updated 2002). <http://www2.lut.fi/~jlampine/debiblio.htm>
67. Lampinen, J. Differential Evolution — New Naturally Parallel Approach for Engineering Design Optimization, // Barry, H.V. Topping (ed.) *Development in Computational Mechanics with High Performance Computing*, Civil-Comp Press, Edinburg. — 1999. — Pp. 217–228.
68. Milani, A., Santucci, V. Asynchronous Differential Evolution // *Proc. 2010 IEEE Congr. Evol. Comput.* — 2010. — Pp. 1210–1216.
69. Pošik, P. and Klemš, V. Jade, an Adaptive Differential Evolution Algorithm, Benchmarked on the Bbob Noiseless Testbed // In *Proceedings of the 14-th International Conference on Genetic and Evolutionary Computation Conference Companion, GECCO*. ACM. — 2012. — Pp. 197–204.
70. Price, K.V., Storn, R.M., Lampinen, J.A. *Differential Evolution: A Practical Approach to Global Optimization* // Springer-Verlag, Berlin Heidelberg. — 2005.
71. Price, K.V. Differential Evolution vs. the Functions of the 2nd ICEO // *Proc. IEEE Int. Conf. Evol. Comput.* — Apr. 1997. — Pp. 153–157.
72. Price, K.V. and Storn, R. Differential Evolution: A Simple Evolution Strategy for Fast Optimization // *Dr. Dobb's J.* — 1997. — V. 22, No 4. — Pp. 18–24.
73. Price, K.V., Storn, R.V. Differential Evolution — A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces // *J. of Global Optimization.* — 1997. — 11. — P. 341–359.
74. Qing, A. Dynamic Differential Evolution Strategy and Applications in Electromagnetic Inverse Scattering Problems // *IEEE Trans. Geosci. Remote Sensing.* — 2006. — 44. — P. 116–125.
75. Storn, R., Price, K.V. Differential Evolution — A Simple and Efficient

- Adaptive Scheme for Global Optimization over Continuous Spaces // ICSI, USA, Technical Report TR-95-012. — March 1995. — URL:<http://icsi.berkeley.edu/~storn/litera.html>.
76. Storn, R. and Price, K.V. Minimizing the Real Functions of the ICEC 1996 Contest by Differential Evolution // Proc. IEEE Int. Conf. Evol. Comput. — 1996. — Pp. 842–844.
 77. Storn, R. On the Usage of Differential Evolution for Function Optimization // Proc. North Am. Fuzzy Inform. Process. Soc. — 1996. — Pp. 519–523.
 78. Suganthan, P.N. et al. Problem Definitions and Evaluation Criteria for the CEC05 Special Session on Real-Parameter Optimization // Tech. rep., Nanyang Technological University, Singapore. — 2005. — URL:http://www.ntu.edu.sg/home/epnsugan/index_files/CEC-05/CEC05.htm.
 79. Zaharie, D. Critical Values for the Control Parameters of Differential Evolution Algorithms // Proc. 8th Int. Mendel Conf. Soft Comput. — 2002. — Pp. 62–67.
 80. Zhang, J. and Sanderson, A. JADE: Adaptive Differential Evolution With Optional External Archive // IEEE Trans. Evol. Comput. — 2009. — 13(5). — Pp. 945–958.

Рассеяние адронов на нуклонах и микроскопический оптический потенциал

81. Грейпеос, М.Е., Кутрулос, Х.Г., Лукьянов, В.К., Шебеко, А.В. Properties of Fermi and Symmetrized Fermi Functions and Applications in Nuclear Physics // Физика элементарных частиц и атомного ядра. — 2001. — **32**, вып. 6. — С. 1494–1565.
82. Лукьянов, В.К., Земляная, Е.В., Лукьянов, К.В. Ядро-ядерное рассеяние и оптический потенциал фолдинга // Ядерная физика. — 2006. — **69**. — С. 262–275.
83. Лукьянов, В.К., Земляная, Е. В., Лукьянов, К.В., Ибрагим Абдулмагед, Али-Эль-Лити, Словинский, Б. Анализ упругого рассеяния пимезонов ядрами в рамках микроскопического оптического потенциала // Направлено в “Известия РАН, серия физическая”.
84. Лукьянов, В.К., Земляная, Е.В., Лукьянов, К.В., Ханна, К.М. Микроскопический K^+ -ядерный оптический потенциал и соответствующие расчеты дифференциальных сечений упругого рассеяния и полных

- сечений реакции // Ядерная физика. — 2010. — **73**. — С.1489–1496.
85. Лукьянов, В.К., Земляная, Е.В., Словинский, Б. Полные сечения ядро-ядерных реакций в подходе Глаубера-Ситенко для реалистических распределений ядерной материи // Ядерная физика. — 2004. — **67**. — Рр. 1282–1306.
 86. Лукьянов, В.К., Словинский, Б., Земляная, Е.В. // Ядерная Физика. — 2001. — **64**, №7. — С.1349–1357
 87. Эрикссон, Т., Вайзе, В. Пионы и ядра // Пер. с англ./ Под ред. И.С. Шапиро. М.: Наука. Гл. ред. физ.-мат. лит. — 1991. —512с.
 88. El-Azab Farid, M., Satchler, G.R. A Density-Dependent Interaction in the Folding Model for Heavy-Ion Potentials+ // Nucl. Phys. A. — 1985. — **438**, — Рр.525–535.
 89. Bano, N. and Ahmed, I. Medium-Energy Pion Scattering and the α -Cluster Description of ^{12}C // J. Phys. — 1979. — **G 5**. — P.39.
 90. Czyz, W. and Maximon, L.C. // Ann. Phys. — 1969. — **52**. — P. 59.
 91. Donnachie, A., Kirsopp, R.G., Lovelace, C. // Phys. Lett. B. — 1968. — **26**. — P. 161–165.
 92. Tag, Eldint, I.M.A.T., Esmaelj, E.H., Hassan, M.Y.M. and Comsans, M.N.H. Study of the Optical Potential for Hadron-Nucleus Scattering Using Glauber's theory //J. Phys. G: Nucl. Part. Phys.. — 1990. — V.16. — Рр. 1051–1076.
 93. Geesaman, D.F. et al, // Phys. Rev. — 1981. — **C 23**, no.6. — P. 2635.
 94. Glauber, R.J. Lectures in Theoretical Physics // New York, Interscience. — 1959. — P. 315.
 95. Gretillat, P. et al, // Nucl. Phys. — 1981. — **A 364**.— P. 270.
 96. Igo, G.I. // Rev. Mod. Phys. — 1978. — 50. — P. 522.
 97. Khoa, Dao, T., Satchler, G.R. // Nucl. Phys.A. — 2000. — **668**. — Рр. 2–41.
 98. Kisslinger, L.S., Scattering of Mesons by Light Nuclei // Phys. Rev. — 1955. — **98**. — Рр. 761–765.
 99. Krell, M., Ericson, T.E.O. Energy Levels and Wave Functions of Pionic Atoms // Nucl. Phys. — 1969. — **B 11**. — p. 521–550.
 100. Kunz, P.D. and Rost, E. // Comp.Nuc. Ph., Springer-Verlag. — 1993. — **2**. — Рр. 88–107. // Kunz, P.D. URL:<http://spot.colorado.edu/~kunz/Home.html>
 101. Lesnyak, H. and Lesnyak, L. // Nucl. Phys. — 1971. — **B 38**. — P. 221.
 102. Lukyanov, V.K., Zemlyanaya, E.V., Lukyanov, K.V., Abdul-

- Magead, I.A.M. Analysis of Inelastic Pion-Nucleus Scattering Within the Microscopic Optical Potential and the In-Medium Effect on pN Amplitude in Nuclei // Proceedings of XXII International Baldin Seminar on High Energy Physics Problems "Relativistic Nuclear Physics and Quantum Chromodynamics"(JINR, Dubna, Russia), PoS SISSA. — 2015. — P.124(1–10)
103. Locher, M.P., Steinmann, O., Straumann, N. Why Is the $\Delta(1232)$ -Resonance Shifted in Nuclei? // Nucl. Phys. B. — 1971. — **27**. — Pp.598–804.
104. Olmer, C. et al. Elastic and Inelastic Scattering of 162 MeV Pions by ^{28}Si , ^{58}Ni , and ^{208}Pb // Phys. Rev. C. 1980. — **21**, No 1. — Pp.254–272.
105. Patterson, J.D., Peterson, R.J. Empirical Distributions of Protons Within Nuclei // Nucl. Phys. A. — 2003. — **717**, 3. — Pp.235–246.
106. Preedom, B.M. et al. A Systematic Study of π^+ and π^- Inelastic Scattering From ^{28}Si in the Region of the $\pi N(3, 3)$ Resonance // Nuc. Phys. A. — 1979. — **326**. — Pp.385–400.
107. Ray, L. // Phys. Rev. — 1979. — C 19 1855, C 20 1867
108. Roper, L.D., Wright, R.M. // Phys. Rev. B. — 1965. — **138**, 48. — Pp.921–932.
109. Shalaby, A.S., Hassan, M.Y.M., El-Gogary, M.M.H. // Brazilian J. Phys. — 2007. — **37**, No 2A. — P. 388.
110. Sitenko, A.G. // Ukr. Fiz. J. — 1959. — **4**. — p.152. // Ситенко, А.Г. // Укр. Физ. Журн.— 1959. — Т.4.— С.152.
111. Shuka, P. // Phys. Rev. C. — 2003. — V. 67. — P. 054607.

Исследование липидных мембран и везикул

112. Земляная, Е.В., Киселев, М.И. Определение структуры однослойных везикул Димиристоилфосфатидилхолина по данным малоуглового рассеяния нейтронов в рамках модели разделенных формфакторов // Направлено на XVII Собрание по использованию рассеяния нейтронов в исследованиях конденсированного состояния “РИНКС-2002”, Гатчина, Октябрь 2002. // Препринт РЗ-2002-163. Дубна. ОИЯИ. — 2002.
113. Киселев, М.А., Ломбардо, Д., Киселев, А.М., Лези, П., Аксенов, В.Л. Исследование структурного фактора однослойных везикул димиристоилфосфатидилхолина методом малоуглового рассеяния рентгеновских лучей // Поверхность. Рентгеновские, синхротронные и нейтрон-

- ные исследования. — 2003. — №11. — С. 20–24.
114. Свергун, Д.И., Фейгин, Л.А. Рентгеновское и нейтронное малоугловое рассеяние // М. Наука. — 1986. — 276 с. // L.A. Feigin, D.I. Svergun, Structure Analysis by Small-Angle X-Ray and Neutron Scattering // Plenum Publishing Corporation, New York. — 1987.
 115. Hallet, F.R., Watton, J., Krygsman, P. // Biophys. J. — 1991. — 59. — P. 357.
 116. Jin, A.J., Huster, D., Gawrisch, K., Nossal, R. Light Scattering Characterization of Extruded Lipid Vesicles // Eur. Biophys. J. — 1999. — 28. — Pp. 187–199.
 117. Kiselev, M.A., Zemlyanaya, E.V., Aswal, V.K., Neubert, R.H.H. What Can We Learn About the Lipid Vesicle Structure From the Small-Angle Neutron Scattering Experiment? // European Biophysics Journal. — 2006. — V. 35, No. 6. — Pp. 477–493.
 118. Kiselev, M.A., Zemlyanaya, E.V., Ryabova, N.Y., Hauss, T., Dante, S., Lombardo, D. Water Distribution Function Across the Curved Lipid Bilayer: SANS Study // Chemical Physics. — 2008. — V. 345. — Pp. 185–190.
 119. Kiselev, M.A., Lesieur, P., Kisselev, A.M., Lombardo, D., Aksenov, V.L., // J. JINR Preprint E3-2001-133, Dubna. — 2001; // J. Appl. Phys. A. — 2002. — 74. — P.S1654.
 120. Kiselev, M.A., Lesieur, P., Kisselev, A.M., Lombardo, D., Killany, M., Lesieur, S. Sucrose Solutions as Prospective Medium to Study the Vesicle Structure: SAXS and SANS Study // J Alloys Compd. — 2001. — 328. — Pp. 71–76
 121. Kiselev, M.A., Zemlyanaya, E.V., Ipatova, O.M., Gruzinov, A.Yu., Ermakova, E.V., Zabelin, A.V., **Zhabitskaya, E.I.**, Druzhilovskaya, O.S., Aksenov, V.L. Application of Small-Angle X-Ray Scattering to the Characterization and Quantification of the Drug Transport Nanosystem Based on the Soybean Phosphatidylcholine // Journal of Pharmaceutical and Biomedical Analysis. — 2015. — 114. — Pp. 288–291.
 122. **Evgeniya Zhabitskaya**, Elena Zemlyanaya, Mikhail Kiselev and Andrey Gruzinov. The Parallel Asynchronous Differential Evolution Method as a Tool to Analyze Synchrotron Scattering Experimental Data from Vesicular Systems // EPJ Web of Conferences Vol. — 2016 — 108. — P.02047–p.1. // Mathematical Modeling and Computational Physics (MMCP 2015) Published online: 09 February 2016 DOI: <http://dx.doi.org/10.1051/>

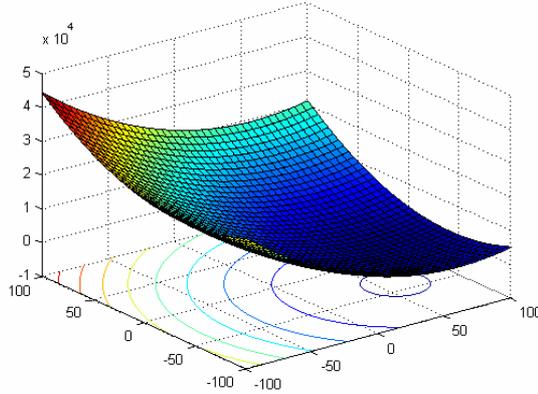
- epjconf/201610802047.
123. Lesieur, P., Kiselev, M.A., Barsukov, L.I., Lombardo, D., // J. Appl. Crystallogr. — 2000. — 33 — P. 623.
 124. Mason, P.C. Gaulin, B.D., Eppard, R.M., Wignall, G.D. Lin, J.S. // Phys. Rev. E. — 1999. — 59. — Pp. 3361–3367.
 125. Ostanovich, Y.M., // Makromol. Chem. Macromol. Symp. 15. — 1988. — Pp. 91–103.
 126. Pencer, J, Hallet, R Small-Angle Neutron Scattering From Large Unilamellar Vesicles: An Improved Method for Membrane Thickness Determination // Phys Rev E. — 2000. — 61. — Pp. 3003–3008.
 127. Schmiedel, H., Joerchel, P., Kiselev, M., Klose, G. Determination of Structural Parameters and Hydration of Unilamellar POPC/C12E4 Vesicles at High Water Excess From Neutron Scattering Curves Using a Novel Method of Evaluation // J. Phys. Chem. B. — 2001. — 105. — P. 111.
 128. E.V. Zemlyanaya, ... **E.I. Zhabitskaya**, et al. SFF analysis of the small angle scattering data for investigation of a vesicle systems structure // Journal of Physics: Conf. Series. — 2016. — 724 012056. — URL:<http://iopscience.iop.org/1742-6596/724/1/012056>
 129. Zemlyanaya, E.V., et al. Structure of Unilamellar Vesicles: Numerical Analysis Based on Small-Angle Neutron Scattering Data // Crystallography reports. — 2006. — v. 51, Suppl. 1. — Pp. S22–S26.

Список используемых сокращений

АДЭ — (метод) асинхронной дифференциальной эволюции
 АДЭ-Р — (метод) АДЭ с рестартом
 АДЭ-АКМ — (метод) АДЭ с адаптивной корреляционной матрицей
 ВС — везикулярные системы
 ГМ — глобальный минимум
 ДМФХ — Димиристоилфосфатидилхолин
 ДЭ — (метод) дифференциальной эволюции
 КП — комплекс программ
 ЛМ — локальный минимум
 ММПЯУР — микроскопическая модель пион-ядерного рассеяния
 МОП — микроскопический оптический потенциал
 МУРН — малоугловое рассеяние нейтронов
 МУРР — малоугловое рентгеновское рассеяние
 РВУ — релятивистское волновое уравнение
 РФФ — (метод) разделенных формфакторов
 ЦФ — целевая функция
 ЭА — эволюционный(е) алгоритм(ы)

Приложения

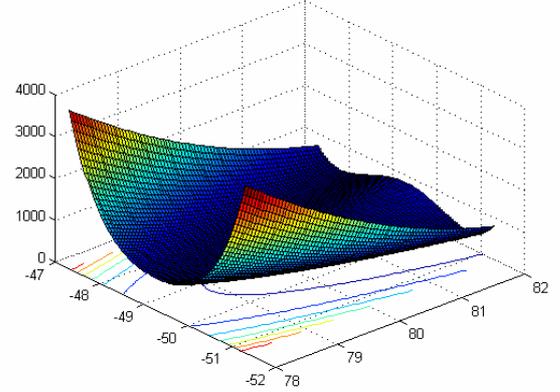
П. I. Используемые в работе тестовые функции из СЕС-2005 [78]



Сдвинутая гиперсфера ($D = 10$).

$$f_1(\mathbf{x}) = \sum_{j=1}^D z_j^2 + f_{\text{bias}},$$

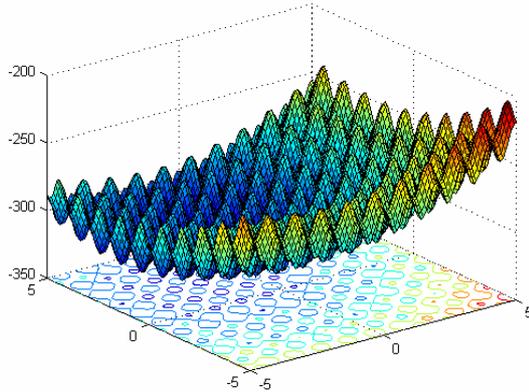
$$\mathbf{z} = \mathbf{x} - \mathbf{o}, \quad \mathbf{x} \in [-100, 100]^D$$



Сдвинутая функция Розенброка ($D = 10$)

$$f_6(\mathbf{x}) = \sum_{j=1}^{D-1} (100(z_j^2 - z_{j+1}^2) + (z_j - 1)^2) + f_{\text{bias}},$$

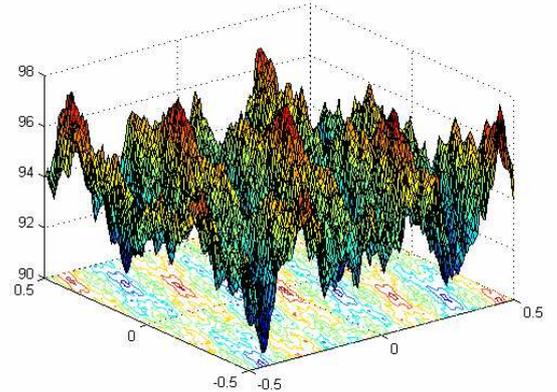
$$\mathbf{z} = \mathbf{x} - \mathbf{o} + 1, \quad \mathbf{x} \in [-100, 100]^D$$



Сдвинутая функция Растригина ($D = 10$)

$$f_9(\mathbf{x}) = \sum_{j=1}^D (z_j^2 - 10 \cos(2\pi z_j) + 10) + f_{\text{bias}},$$

$$\mathbf{z} = \mathbf{x} - \mathbf{o}, \quad \mathbf{x} \in [-5, 5]^D$$



Приближение сдвинутой повернутой функции Вейерштрасса ($D = 10$).

$$f_{11}(\mathbf{x}) = \sum_{j=1}^{D-1} \sum_{k=0}^{k_{\text{max}}} a^k \cos(2\pi b^k(z_j + 0.5)) + f_{\text{bias}},$$

$$\mathbf{z} = \mathbf{x} - \mathbf{o}, \quad \mathbf{x} \in [-0.5, 0.5]^D$$

f_1	Гиперсфера	один минимум, с разделяемыми переменными
f_6	функция Розенброка	один минимум, с неразделяемыми переменными
f_9	функция Растригина	много минимумов, с разделяемыми переменными
f_{11}	функция Вейерштрасса	много минимумов, с неразделяемыми переменными

II. II. Доказательства теорем 1.2.1 — 1.2.3 из раздела 1.2 диссертации

Доказательство теоремы 1.2.1 (rand/1/bin) (a)

Каждый элемент v_l популяции P_v равен

$$v_l = \theta_r + F(\theta_p - \theta_q), \text{ где } l, r, p \text{ и } q - \text{различны.}$$

Нетрудно показать, что при фиксированном l^9 :

$$E(\theta_r | r \neq l) = E(\theta_p | p \neq l) = E(\theta_q | q \neq l) = \frac{N_p}{N_p - 1} \left(\bar{\theta} - \frac{1}{N_p} \theta_l \right), \quad (\text{II.1})$$

$$\begin{aligned} E(\theta_r^T \theta_r | r \neq l) &= E(\theta_p^T \theta_p | p \neq l) = E(\theta_q^T \theta_q | q \neq l) \\ &= \frac{N_p}{N_p - 1} \left(\overline{\theta^T \theta} - \frac{1}{N_p} \theta_l^T \theta_l \right), \end{aligned} \quad (\text{II.2})$$

$$E(\theta_r \theta_p | r \neq l, p \neq l, r \neq p) = \frac{N_p(N_p \bar{\theta}^T \bar{\theta} - \overline{\theta^T \theta}) - 2(N_p \bar{\theta}^T \theta_l - \theta_l^T \theta_l)}{(N_p - 1)(N_p - 2)}. \quad (\text{II.3})$$

С другой стороны, очевидно, что для случая случайного масштабирующего фактора, определенного в (1.5) $E(\tilde{\mathbf{F}}_l) = 0$, а

$$\text{Var}(\tilde{\mathbf{F}}_l) = E(\tilde{\mathbf{F}}_l^T \tilde{\mathbf{F}}_l) = \mathbf{F}_l^T \mathbf{F}_l. \quad (\text{II.4})$$

Случайная переменная $\tilde{\mathbf{F}}_l$ не зависит от случайных переменных θ_i для любых i .

В рассматриваемых вариантах алгоритма операции мутации и кроссовера действуют независимо на каждую из компонент. Поэтому в доказательстве можно ограничиться покомпонентным анализом. Для краткости записи будем опускать знак вектора и индекс компоненты. Тогда в покомпонентном виде уравнения (II.1)-(II.4) примут вид:¹⁰

$$E(\theta_r | r \neq l) \equiv E_l(\theta_r) = E_l(\theta_p) = E_l(\theta_q) = \frac{N_p}{N_p - 1} \left(\bar{\theta} - \frac{1}{N_p} \theta_l \right), \quad (\text{II.5})$$

$$E(\theta_r^2 | r \neq l) \equiv E_l(\theta_r^2) = E_l(\theta_p^2) = E_l(\theta_q^2) = \frac{N_p}{N_p - 1} \left(\overline{\theta^2} - \frac{1}{N_p} \theta_l^2 \right), \quad (\text{II.6})$$

$$E(\theta_r \theta_p | r \neq l, p \neq l, r \neq p) \equiv E_l(\theta_r \theta_p) = \frac{N_p(N_p \bar{\theta}^2 - \overline{\theta^2}) - 2(N_p \bar{\theta} \theta_l - \theta_l^2)}{(N_p - 1)(N_p - 2)}, \quad (\text{II.7})$$

$$\text{Var}(\tilde{F}_l) = F^2. \quad (\text{II.8})$$

Введенные здесь обозначения $E_l(\cdot)$ означают условное математическое ожидание.

Рассчитаем $E(v_l | r \neq l, p \neq l, q \neq l, r \neq p, r \neq q, p \neq q) \equiv E_l(v_l)^{11}$:

$$E_l(v_l) = E_l(\theta_{r_l}) + 0 = \frac{N_p}{N_p - 1} \left(\bar{\theta} - \frac{1}{N_p} \theta_l \right). \quad (\text{II.9})$$

Рассчитаем $E(v_l^2 | r \neq l, p \neq l, q \neq l, r \neq p, r \neq q, p \neq q) \equiv E_l(v_l^2)^{12}$:

$$E_l(v_l^2) = (1 + 2F^2) \frac{N_p}{N_p - 1} \left(\overline{\theta^2} - \frac{1}{N_p} \theta_l^2 \right) - 2F^2 \frac{N_p(N_p \bar{\theta}^2 - \overline{\theta^2}) - 2(N_p \bar{\theta} \theta_l - \theta_l^2)}{(N_p - 1)(N_p - 2)}. \quad (\text{II.10})$$

Также нетрудно показать, что¹³:

$$E_{l,k}(v_l v_k) = \frac{N_p^2}{(N_p - 1)^2} \left(\bar{\theta} - \frac{1}{N_p} \theta_l \right) \left(\bar{\theta} - \frac{1}{N_p} \theta_k \right). \quad (\text{II.11})$$

⁹Если не учитывать $r, p, q \neq l$, то $E(\theta_r) = \bar{\theta}$ и $E(\theta_r^2) = \overline{\theta^T \theta}$, а $E(\theta_r^T \theta_p | r \neq p) = \frac{N_p \bar{\theta}^T \bar{\theta} - \overline{\theta^T \theta}}{N_p - 1}$.

¹⁰Если не учитывать $r, p, q \neq l$, то $E(\theta_r) = \bar{\theta}$ и $E(\theta_r^2) = \overline{\theta^2}$, а $E(\theta_r \theta_p | r \neq p) = \frac{N_p \bar{\theta}^2 - \overline{\theta^2}}{N_p - 1}$.

¹¹Если не учитывать $r, p, q \neq l$, то $E(v_l) = \bar{\theta}$.

¹²Если не учитывать $r, p, q \neq l$, то $E(v_l^2) = \overline{\theta^2} - \frac{2N_p F^2}{N_p - 1} \left(\overline{\theta^2} - \bar{\theta}^2 \right)$.

¹³Если не учитывать $r, p, q \neq l, r, p, q \neq k$, то $E(v_l v_k) = \bar{\theta}^2$.

Теперь рассчитаем $E\left(\overline{P_v^2}\right)$. Для стратегии $rand/1/bin$ ¹⁴:

$$E\left(\overline{P_v^2}\right) = \frac{1}{N_p} \sum_{l=0}^{l=N_p-1} E_l(v_l^2) = \overline{\theta^2} + \frac{2N_p F^2}{N_p - 1} (\overline{\theta^2} - \overline{\theta}^2). \quad (\text{II.12})$$

Следующее, что нужно рассчитать, это $E\left(\overline{P_v^2}\right)$. Для расчета нам понадобятся соотношения:

$$\sum_l \sum_{k \neq l} (\theta_l + \theta_k) = 2N_p(N_p - 1)\overline{\theta}, \quad (\text{II.13})$$

$$\sum_l \sum_{k \neq l} (\theta_l \theta_k) = N_p^2 \overline{\theta^2} - N_p \overline{\theta}^2, \quad (\text{II.14})$$

а также

$$\begin{aligned} \frac{1}{N_p^2} E\left(\sum_{l=0}^{l=N_p-1} \sum_{k=0, k \neq l}^{k=N_p-1} v_l v_k\right) &= \\ &= \frac{1}{N_p^2 (N_p - 1)^2} \sum_{l=0}^{l=N_p-1} \sum_{k=0, k \neq l}^{k=N_p-1} \left(\overline{\theta} - \frac{1}{N_p} \theta_l\right) \left(\overline{\theta} - \frac{1}{N_p} \theta_k\right) = \\ &= \frac{1}{(N_p - 1)^2} \sum_{l=0}^{l=N_p-1} \sum_{k=0, k \neq l}^{k=N_p-1} \left(\overline{\theta^2} - \frac{1}{N_p} \overline{\theta} (\theta_l + \theta_k) + \frac{1}{N_p^2} \theta_l \theta_k\right) = \\ &= \frac{1}{(N_p - 1)^2} \left[N_p(N_p - 1)\overline{\theta^2} - \frac{\overline{\theta}}{N_p} 2N_p(N_p - 1)\overline{\theta} + \frac{N_p^2 \overline{\theta^2} - N_p \overline{\theta}^2}{N_p^2} \right] = \\ &= \frac{N_p - 1}{N_p} \overline{\theta^2} - \frac{\overline{\theta^2} - \overline{\theta}^2}{N_p(N_p - 1)^2}. \quad (\text{II.15}) \end{aligned}$$

Таким образом¹⁵,

$$\begin{aligned} E\left(\overline{P_v^2}\right) &= \frac{1}{N_p^2} E\left(\left[\sum_{l=0}^{l=N_p-1} v_l\right]^2\right) = \frac{1}{N_p} E\left(\overline{P_v^2}\right) + \frac{1}{N_p^2} E\left(\sum_{l=0}^{l=N_p-1} \sum_{k=0, k \neq l}^{k=N_p-1} v_l v_k\right) = \\ &= \frac{1}{N_p} E\left(\overline{P_v^2}\right) + \frac{N_p - 1}{N_p} \overline{\theta^2} - \frac{\overline{\theta^2} - \overline{\theta}^2}{N_p(N_p - 1)^2}. \quad (\text{II.16}) \end{aligned}$$

В результате для каждой компоненты¹⁶:

$$\begin{aligned} E(\text{Var}(P_v))_i &= E\left(\overline{P_v^2}\right) - E\left(\overline{P_v}\right)^2 = \\ &= \left(1 - \frac{1}{N_p}\right) \left(\overline{\theta^2} + \frac{2N_p F^2}{N_p - 1} (\overline{\theta^2} - \overline{\theta}^2)\right) - \left(\frac{N_p - 1}{N_p} \overline{\theta^2} - \frac{\overline{\theta^2} - \overline{\theta}^2}{N_p(N_p - 1)^2}\right) = \\ &= \left(1 + 2F_i^2 - \frac{1}{N_p} + \frac{1}{N_p(N_p - 1)^2}\right) \text{Var}(P_\theta)_i. \quad (\text{II.17}) \end{aligned}$$

Таким образом, мы доказали утверждение 1.6.

(б) Каждый элемент популяции P_u — это случайная величина со структурой:

$$u_{l,j} = \begin{cases} v_{l,j} & \text{с вероятностью } C_r; \\ \theta_{l,j} & \text{с вероятностью } (1 - C_r). \end{cases} \quad (\text{II.18})$$

¹⁴Если не учитывать $r, p, q \neq l$, то результат будет тот же.

¹⁵Если не учитывать $r, p, q \neq l$ и $r, p, q \neq k$, то $E\left(\overline{P_v^2}\right) = \frac{1}{N_p} E\left(\overline{P_v^2}\right) + \frac{N_p - 1}{N_p} \overline{\theta^2}$.

¹⁶Если не учитывать $r, p, q \neq l$ и $r, p, q \neq k$, то $E(\text{Var}(P_v))_i = \left(1 + 2F_i^2 - \frac{1}{N_p}\right) \text{Var}(P_\theta)_i$.

Для того, чтобы определить $E(\text{Var}(P_u))$, необходимо вычислить $E(\overline{P_u^2})$ и $E(\overline{P_u}^2)$. Используя свойства средних для случайных величин, получаем:

$$\begin{aligned} E(\overline{P_u^2}) &= \frac{1}{N_p} \sum_{l=0}^{l=N_p-1} E(u_l^2) = \frac{1}{N_p} \sum_{l=0}^{l=N_p-1} (CrE(v_l^2) + (1-Cr)\theta_l^2) = \\ &= C_r E(\overline{P_v^2}) + (1-C_r)\overline{\theta^2} = (\text{для стратегии } rand/1/bin) \\ &= C_r \left(\overline{\theta^2} + \frac{2N_p F^2}{N_p - 1} (\overline{\theta^2} - \overline{\theta}^2) \right) + (1-C_r)\overline{\theta^2} = \overline{\theta^2} + \frac{2N_p F^2 C_r}{N_p - 1} (\overline{\theta^2} - \overline{\theta}^2). \end{aligned} \quad (\text{II.19})$$

С другой стороны, для $E(\overline{P_u}^2)$ имеем

$$\begin{aligned} E(\overline{P_u}^2) &= \frac{1}{N_p^2} E \left\{ \left[\sum_{l=0}^{l=N_p-1} u_l \right]^2 \right\} = \frac{1}{N_p^2} E \left[\sum_{l=0}^{l=N_p-1} \sum_{k=0, k \neq l}^{k=N_p-1} u_l u_k \right] + \frac{1}{N_p} E(\overline{P_u^2}) = \\ &= \frac{1}{N_p^2} E \left\{ \sum_{l=0}^{l=N_p-1} \sum_{k=0}^{k=N_p-1} (C_r v_l + (1-C_r)\theta_l) (C_r v_k + (1-C_r)\theta_k) \right\} - \\ &\quad - \frac{1}{N_p^2} E \left\{ \sum_{l=0}^{l=N_p-1} (C_r v_l + (1-C_r)\theta_l)^2 \right\} + \frac{1}{N_p} E(\overline{P_u^2}) = \\ &= \frac{1}{N_p^2} E \left\{ \sum_{l=0}^{l=N_p-1} \sum_{k=0}^{k=N_p-1} (C_r^2 v_k v_l + C_r(1-C_r)(v_l \theta_k + v_k \theta_l) + (1-C_r)^2 \theta_k \theta_l) \right\} - \\ &\quad - \frac{1}{N_p^2} E \left\{ \sum_{l=0}^{l=N_p-1} (C_r^2 v_l^2 + 2C_r(1-C_r)v_l \theta_l + (1-C_r)^2 \theta_l^2) \right\} + \frac{1}{N_p} E(\overline{P_u^2}). \end{aligned}$$

Учтем, что

$$\begin{aligned} \frac{1}{N_p^2} E \left(\sum_{l=0}^{l=N_p-1} \sum_{k=0}^{k=N_p-1} v_l v_k \right) &= E(\overline{P_v^2}) = \\ &= \frac{1}{N_p} E(\overline{P_v^2}) + \frac{N_p-1}{N_p} \overline{\theta}^2 - \frac{\overline{\theta^2} - \overline{\theta}^2}{N_p(N_p-1)^2}; \end{aligned} \quad (\text{II.20})$$

$$\begin{aligned} \frac{1}{N_p^2} E \left(\sum_{l=0}^{l=N_p-1} \sum_{k=0}^{k=N_p-1} v_l \theta_k \right) &= \frac{1}{N_p^2} \sum_{l=0}^{l=N_p-1} \sum_{k=0}^{k=N_p-1} E(\theta_{r_l} + F(\theta_{r_l} - \theta_{r_l})) \theta_k = \\ &= \frac{1}{N_p^2} \sum_{l=0}^{l=N_p-1} \sum_{k=0}^{k=N_p-1} \frac{N_p}{N_p-1} \left(\overline{\theta} - \frac{1}{N_p} \theta_l \right) \theta_k + 0 = \overline{\theta}^2; \end{aligned} \quad (\text{II.21})$$

$$\frac{1}{N_p^2} E \left(\sum_{l=0}^{l=N_p-1} \sum_{k=0}^{k=N_p-1} \theta_l \theta_k \right) = \overline{\theta}^2; \quad (\text{II.22})$$

$$\frac{1}{N_p^2} E \left(\sum_{l=0}^{l=N_p-1} v_l^2 \right) = \frac{1}{N_p} E(\overline{P_v^2}); \quad (\text{II.23})$$

$$\frac{1}{N_p^2} E \left(\sum_{l=0}^{l=N_p-1} \theta_l^2 \right) = \frac{1}{N_p} \overline{\theta^2}; \quad (\text{II.24})$$

$$\frac{1}{N_p^2} E \left(\sum_{l=0}^{l=N_p-1} v_l \theta_l \right) = \frac{1}{N_p^2} \sum_{l=0}^{l=N_p-1} \frac{N_p}{N_p-1} \left(\overline{\theta} - \frac{1}{N_p} \theta_l \right) \theta_l = \frac{N_p \overline{\theta^2} - \overline{\theta}^2}{(N_p-1)N_p}. \quad (\text{II.25})$$

Откуда

$$\begin{aligned} E(\overline{P_u}^2) &= \frac{1}{N_p} E(\overline{P_u^2}) + \\ &+ C_r^2 \left[\frac{1}{N_p} E(\overline{P_v^2}) + \frac{N_p-1}{N_p} \overline{\theta}^2 - \frac{\overline{\theta^2} - \overline{\theta}^2}{N_p(N_p-1)^2} \right] + 2C_r(1-C_r)\overline{\theta^2} + (1-C_r)^2 \overline{\theta}^2 - \\ &- \left[C_r^2 \frac{1}{N_p} E(\overline{P_v^2}) + 2C_r(1-C_r) \frac{1}{N_p-1} \left(\overline{\theta^2} - \frac{1}{N_p} \overline{\theta^2} \right) + (1-C_r)^2 \frac{1}{N_p} \overline{\theta^2} \right] = \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{N_p} E \left(\overline{P_u^2} \right) + \frac{N_p - 1}{N_p} \overline{\theta^2} - C_r^2 \frac{\overline{\theta^2} - \overline{\theta}^2}{N_p(N_p - 1)^2} + 2C_r(1 - C_r) \frac{\overline{\theta^2} - \overline{\theta}^2}{N_p(N_p - 1)} - (1 - C_r)^2 \frac{\overline{\theta^2} - \overline{\theta}^2}{N_p} = \\
&= \frac{1}{N_p} E \left(\overline{P_u^2} \right) + \frac{N_p - 1}{N_p} \overline{\theta^2} - \frac{\overline{\theta^2} - \overline{\theta}^2}{N_p} \left(1 - \frac{N_p}{N_p - 1} C_r \right)^2. \quad (\text{II.26})
\end{aligned}$$

Таким образом¹⁷,

$$\begin{aligned}
E(\text{Var}(P_u))_i &= E \left(\overline{P_u^2} \right) - E \left(\overline{P_u}^2 \right) = \\
&= \left(1 - \frac{1}{N_p} \right) \left(\overline{\theta^2} + \frac{2N_p F^2 C_r}{N_p - 1} (\overline{\theta^2} - \overline{\theta}^2) \right) - \frac{N_p - 1}{N_p} \overline{\theta^2} + \frac{\overline{\theta^2} - \overline{\theta}^2}{N_p} \left(1 - \frac{N_p}{N_p - 1} C_r \right)^2 = \\
&= \left(1 - \frac{1}{N_p} + 2F^2 C_r + \frac{1}{N_p} \left(1 - \frac{N_p}{N_p - 1} C_r \right)^2 \right) (\overline{\theta^2} - \overline{\theta}^2) = \\
&= \left(1 + 2F^2 C_r - \frac{C_r}{N_p - 1} \left(2 - \frac{N_p}{N_p - 1} C_r \right) \right) \text{Var}(P_\theta). \quad (\text{II.27})
\end{aligned}$$

Таким образом, мы доказали утверждение 1.7.

ч. и т. д. ■

Доказательство теоремы 1.2.2 — для стратегии "best/1/bin"(a)

Каждый элемент v_l популяции P_v равен

$$v_l = \theta_0 + F(\theta_p - \theta_q),$$

где l, r, p и q — различны; θ_0 — лучший ("best") вектор в популяции.

Нетрудно показать, что при фиксированном l покомпонентно¹⁸:

$$E(\theta_p | p \neq l) = E(\theta_q | q \neq l) = \frac{N}{N-1} (\overline{\theta} - \frac{1}{N} \theta_l), \quad (\text{II.28})$$

$$E(\theta_p^2 | p \neq l) = E(\theta_q^2 | q \neq l) = \frac{N}{N-1} (\overline{\theta^2} - \frac{1}{N} \theta_l^2), \quad (\text{II.29})$$

$$E(\theta_r \theta_p | q \neq l, p \neq l, q \neq p) = \frac{N(N\overline{\theta^2} - \overline{\theta^2}) - 2(N\overline{\theta} \theta_l - \theta_l^2)}{(N-1)(N-2)}, \quad (\text{II.30})$$

Случайная переменная $E(\tilde{F}_l)$ независима от случайных переменных θ_i для любых i ; $E(\tilde{F}_l) = 0$ и $\text{Var}(\tilde{F}_l) = F^2$.

Рассчитаем $E_l(v_l)$, $E_l(v_l^2)$ и $E_{l,k}(v_l v_k)$ ¹⁹:

$$E_l(v_l) = E_l(\theta_{r_l}) + 0 = \frac{N}{N-1} (\overline{\theta} - \frac{1}{N} \theta_l), \quad (\text{II.31})$$

$$E_l(v_l^2) = \theta_0 + 2F^2 \frac{N\overline{\theta^2} - \theta_l^2}{N-1} - 2F^2 \frac{N(N\overline{\theta^2} - \overline{\theta^2}) - 2(N\overline{\theta} \theta_l - \theta_l^2)}{(N-1)(N-2)}, \quad (\text{II.32})$$

$$E_{l,k}(v_l v_k) = \theta_0^2. \quad (\text{II.33})$$

Теперь рассчитаем $E(\overline{P_v^2})$. Для стратегии *best/1/bin*²⁰:

$$E(\overline{P_v^2}) = \frac{1}{N} \sum_{l=0}^{l=N-1} E_l(v_l^2) = \theta_0^2 + \frac{2NF^2}{N-1} (\overline{\theta^2} - \overline{\theta}^2). \quad (\text{II.34})$$

Следующее, что нужно рассчитать, это $E(\overline{P_v^2})$. Учитывая то, что для стратегии

¹⁷Если не учитывать "П1", то $E(\text{Var}(P_u))_i = \left(1 + 2F^2 C_r - \frac{C_r}{N_p} (2 - C_r) \right) \text{Var}(P_\theta)$.

¹⁸Если не учитывать $r, p, q \neq l$, то $E(\theta_r) = \overline{\theta}$ и $E(\theta_r^2) = \overline{\theta^2}$, а $E(\theta_p \theta_q | q \neq p) = \frac{N\overline{\theta^2} - \overline{\theta^2}}{N-1}$.

¹⁹Если не учитывать $r, p, q \neq l$, то $E(v_l) = \overline{\theta}$, $E(v_l^2) = \overline{\theta^2} - \frac{2NF^2}{N-1} (\overline{\theta^2} - \overline{\theta}^2)$, и $E(v_l v_k) = \overline{\theta^2}$.

²⁰Если не учитывать $r, p, q \neq l$, то результат будет тот же.

best/1/bin $\frac{1}{N^2} E \left(\sum_{l=0}^{l=N-1} \sum_{k=0, k \neq l}^{k=N-1} v_l v_k \right) = \frac{N-1}{N} \theta_0^2$, получаем²¹:

$$\begin{aligned} E \left(\overline{P_v^2} \right) &= \frac{1}{N^2} E \left(\left[\sum_{l=0}^{l=N-1} v_l \right]^2 \right) = \\ &= \frac{1}{N} E \left(\overline{P_v^2} \right) + \frac{1}{N^2} E \left(\sum_{l=0}^{l=N-1} \sum_{k=0, k \neq l}^{k=N-1} v_l v_k \right) = \frac{1}{N} E \left(\overline{P_v^2} \right) + \frac{N-1}{N} \theta_0^2. \end{aligned} \quad (\text{II.35})$$

В результате для каждой компоненты вариации²²:

$$\begin{aligned} E(\text{Var}(P_v))_j &= E \left(\overline{P_v^2} \right) - E \left(\overline{P_v} \right)^2 = \\ &= \left(1 - \frac{1}{N} \right) \left(\theta_0^2 + \frac{2NF^2}{N-1} \left(\overline{\theta^2} - \overline{\theta}^2 \right) \right) - \frac{N-1}{N} \theta_0^2 = 2F_j^2 \text{Var}(P_\theta)_j, \end{aligned} \quad (\text{II.36})$$

и мы доказали утверждение 1.8.

(6) Каждый элемент популяции P_u — это случайная величина со структурой:

$$u_{l,j} = \begin{cases} v_{l,j} & \text{с вероятностью } C_r; \\ \theta_{l,j} & \text{с вероятностью } (1 - C_r). \end{cases} \quad (\text{II.37})$$

Для того, чтобы определить $E(\text{Var}(P_u))$, необходимо вычислить $E \left(\overline{P_u^2} \right)$ и $E \left(\overline{P_u} \right)^2$. Используя свойства средних для случайных величин, получаем:

$$E \left(\overline{P_u^2} \right) = \frac{1}{N} \sum_{l=0}^{l=N-1} E(u_l^2) = \frac{1}{N} \sum_{l=0}^{l=N-1} (C_r E(v_l^2) + (1 - C_r) \theta_l^2) = C_r E(P_v^2) + (1 - C_r) \overline{\theta^2} =$$

для стратегии *best/1/bin*

$$\begin{aligned} E \left(\overline{P_u^2} \right) &= C_r \left(\theta_0^2 + \frac{2NF^2}{N-1} \left(\overline{\theta^2} - \overline{\theta}^2 \right) \right) + (1 - C_r) \overline{\theta^2} = \\ &= \overline{\theta^2} + C_r \left(\theta_0^2 - \overline{\theta^2} \right) + \frac{2NF^2 C_r}{N-1} \left(\overline{\theta^2} - \overline{\theta}^2 \right). \end{aligned} \quad (\text{II.38})$$

С другой стороны, для $E \left(\overline{P_u} \right)^2$ имеем²³:

$$\begin{aligned} E \left(\overline{P_u} \right)^2 &= \frac{1}{N^2} E \left\{ \left[\sum_{l=0}^{l=N-1} u_l \right]^2 \right\} = \frac{1}{N^2} E \left[\sum_{l=0}^{l=N-1} \sum_{k=0, k \neq l}^{k=N-1} u_l u_k \right] + \frac{1}{N} E \left(\overline{P_u^2} \right) = \\ &= \frac{1}{N^2} E \left\{ \sum_{l=0}^{l=N-1} \sum_{k=0}^{k=N-1} (C_r^2 v_k v_l + C_r (1 - C_r) (v_l \theta_k + v_k \theta_l) + (1 - C_r)^2 \theta_k \theta_l) \right\} - \\ &\quad - \frac{1}{N^2} E \left\{ \sum_{l=0}^{l=N-1} (C_r^2 v_l^2 + 2C_r (1 - C_r) v_l \theta_l + (1 - C_r)^2 \theta_l^2) \right\} + \frac{1}{N} E \left(\overline{P_u^2} \right). \end{aligned}$$

Учтем, что

²¹Если не учитывать $r, p, q \neq l$ и $r, p, q \neq k$, то $E \left(\overline{P_v^2} \right) = \frac{1}{N} E \left(\overline{P_v^2} \right) + \frac{N-1}{N} \overline{\theta^2}$.

²²Если не учитывать $r, p, q \neq l$ и $r, p, q \neq k$, то результаты те же.

²³Если не учитывать $r, p, q \neq l$ и $r, p, q \neq k$, то $E \left(\overline{P_v^2} \right) = \frac{1}{N} E \left(\overline{P_v^2} \right) + \frac{N-1}{N} \overline{\theta^2}$.

$$\frac{1}{N^2} E \left(\sum_{l=0}^{l=N-1} \sum_{k=0}^{k=N-1} v_l v_k \right) = E \left(\overline{P_v^2} \right) = \frac{1}{N} E \left(\overline{P_v^2} \right) + \frac{N-1}{N} \theta_0^2; \quad (\text{II.39})$$

$$\begin{aligned} \frac{1}{N^2} E \left(\sum_{l=0}^{l=N-1} \sum_{k=0}^{k=N-1} v_l \theta_k \right) &= \frac{1}{N^2} \sum_{l=0}^{l=N-1} \sum_{k=0}^{k=N-1} E \left(\theta_0 + F(\theta_{r_l} - \theta_{r_k}) \right) \theta_k = \\ &= \frac{1}{N^2} \sum_{l=0}^{l=N-1} \sum_{k=0}^{k=N-1} \theta_0 \theta_k + 0 = \overline{\theta} \overline{\theta}_0; \end{aligned} \quad (\text{II.40})$$

$$\frac{1}{N^2} E \left(\sum_{l=0}^{l=N-1} \sum_{k=0}^{k=N-1} \theta_l \theta_k \right) = \overline{\theta}^2; \quad (\text{II.41})$$

$$\frac{1}{N^2} E \left(\sum_{l=0}^{l=N-1} v_l^2 \right) \equiv \frac{1}{N} E \left(\overline{P_v^2} \right); \quad (\text{II.42})$$

$$\frac{1}{N^2} E \left(\sum_{l=0}^{l=N-1} \theta_l^2 \right) = \frac{1}{N} \overline{\theta}^2; \quad (\text{II.43})$$

$$\frac{1}{N^2} E \left(\sum_{l=0}^{l=N-1} v_l \theta_l \right) = \frac{1}{N^2} \sum_{l=0}^{l=N-1} \theta_0 \theta_l = \frac{\theta_0 \overline{\theta}}{N}. \quad (\text{II.44})$$

Откуда

$$\begin{aligned} E \left(\overline{P_u^2} \right) &= \frac{1}{N} E \left(\overline{P_u^2} \right) + C_r^2 \left[\frac{1}{N} E \left(\overline{P_v^2} \right) + \frac{N-1}{N} \theta_0^2 \right] + 2C_r(1-C_r) \overline{\theta} \theta_0 + (1-C_r)^2 \overline{\theta}^2 - \\ &\quad - \left[C_r^2 \frac{1}{N} E \left(\overline{P_v^2} \right) + 2C_r(1-C_r) \frac{\theta_0 \overline{\theta}}{N} + (1-C_r)^2 \frac{\overline{\theta}^2}{N} \right] = \\ &= \frac{1}{N} E \left(\overline{P_u^2} \right) + \frac{N-1}{N} \left[C_r^2 \theta_0^2 + 2C_r(1-C_r) \theta_0 \overline{\theta} + (1-C_r)^2 \overline{\theta}^2 \right] - \frac{1}{N} (1-C_r)^2 \left(\overline{\theta}^2 - \overline{\theta}^2 \right). \end{aligned} \quad (\text{II.45})$$

Таким образом²⁴,

$$\begin{aligned} E(\text{Var}(P_u))_j &= E \left(\overline{P_u^2} \right) - E \left(\overline{P_u} \right)^2 = \\ &= \left(1 - \frac{1}{N} \right) \left(\overline{\theta}^2 + C_r \left(\theta_0^2 - \overline{\theta}^2 \right) + \frac{2NF^2C_r}{N-1} \left(\overline{\theta}^2 - \overline{\theta}^2 \right) \right) - \\ &\quad - \frac{N-1}{N} \left[C_r^2 \theta_0^2 + 2C_r(1-C_r) \theta_0 \overline{\theta} + (1-C_r)^2 \overline{\theta}^2 \right] + \frac{1}{N} (1-C_r)^2 \left(\overline{\theta}^2 - \overline{\theta}^2 \right) = \\ &= 2F^2C_r \left(\overline{\theta}^2 - \overline{\theta}^2 \right) + \frac{1}{N} (1-C_r)^2 \left(\overline{\theta}^2 - \overline{\theta}^2 \right) + \\ &\quad + \frac{N-1}{N} \left[\overline{\theta}^2 - \overline{\theta}^2 + C_r \left(\theta_0^2 - \overline{\theta}^2 - 2\theta_0 \overline{\theta} + 2\overline{\theta}^2 \right) - C_r^2 \left(\theta_0^2 - 2\theta_0 \overline{\theta} + \overline{\theta}^2 \right) \right] = \\ &= \left(2F^2C_r + \frac{N-1}{N} (1-C_r) + \frac{1}{N} (1-C_r)^2 \right) \left(\overline{\theta}^2 - \overline{\theta}^2 \right) + \frac{N-1}{N} C_r (1-C_r) (\theta_0 - \overline{\theta})^2 = \\ &= \left(2F^2C_r + (1-C_r) \left(1 - \frac{C_r}{N} \right) \right) \left(\overline{\theta}^2 - \overline{\theta}^2 \right) + \frac{N-1}{N} C_r (1-C_r) (\theta_0 - \overline{\theta})^2 = \\ &= \left(2F^2C_r + (1-C_r) \left(1 - \frac{C_r}{N} + C_r \frac{N-1}{N} \frac{(\theta_0 - \overline{\theta})^2}{\overline{\theta}^2 - \overline{\theta}^2} \right) \right) \text{Var}(P_\theta)_j, \end{aligned} \quad (\text{II.46})$$

и мы доказали утверждение 1.7. \blacksquare

Доказательство теоремы 1.2.3 (для rand/rand/1/bin) (a)

В популяции P_v один из элементов v_l равен

$$v_l = \theta_r + F(\theta_p - \theta_q), \text{ где } l, r, p \text{ и } q - \text{ различны,}$$

а все остальные элементы совпадают с соответствующими элементами популяции P_θ .

В рассматриваемых вариантах алгоритма операции мутации и кроссовера действуют независимо на каждую из компонент. Поэтому в доказательстве можно ограничиться покомпонентным анализом. Для краткости далее в записи будем опускать знак

²⁴Если не учитывать "У1", то $E(\text{Var}(P_u))_i = (1 + 2F^2C_r - \frac{C_r}{N} (2 - C_r)) \text{Var}(P_\theta)$.

вектора и индекс компоненты. Тогда покомпонентно²⁵:

$$E(\theta_r | r \neq l) = E(\theta_p | p \neq l) = E(\theta_q | q \neq l) = \frac{N}{N-1} \left(\bar{\theta} - \frac{1}{N} \theta_l \right), \quad (\text{II.47})$$

$$E(\theta_r^2 | r \neq l) = E(\theta_p^2 | p \neq l) = E(\theta_q^2 | q \neq l) = \frac{N}{N-1} \left(\bar{\theta}^2 - \frac{1}{N} \theta_l^2 \right), \quad (\text{II.48})$$

$$E(\theta_r \theta_p | r \neq l, p \neq l, r \neq p) = \frac{N(N\bar{\theta}^2 - \bar{\theta}^2) - 2(N\bar{\theta}\theta_l - \theta_l^2)}{(N-1)(N-2)}. \quad (\text{II.49})$$

Случайная переменная $E(\tilde{F}_l)$ независима от случайных переменных θ_i для любых i и $E(\tilde{F}) = 0$, $\text{Var}(\tilde{F}_l) = F^2$.

Рассчитаем $E(v_l | r \neq l, p \neq l, q \neq l, r \neq p, r \neq q, p \neq q) \equiv E_l(v_l)$ ²⁶:

$$E_l(v_l) = E_l(\theta_{r_l}) + 0 = \frac{N}{N-1} \left(\bar{\theta} - \frac{1}{N} \theta_l \right). \quad (\text{II.50})$$

Для стратегии "rand/rand/1/bin" индекс l выбирается случайно. Тогда

$$E(v_l) = E_{\text{rand}}(E_l(v_l)) = \frac{1}{N} \sum_{l=0}^{l=N-1} E_l(v_l) = \bar{\theta}. \quad (\text{II.51})$$

Рассчитаем $E(v_l^2 | r \neq l, p \neq l, q \neq l, r \neq p, r \neq q, p \neq q) \equiv E_l(v_l^2)$ ²⁷:

$$E_l(v_l^2) = (1 + 2F^2) \frac{N}{N-1} \left(\bar{\theta}^2 - \frac{1}{N} \theta_l^2 \right) - 2F^2 \frac{N(N\bar{\theta}^2 - \bar{\theta}^2) - 2(N\bar{\theta}\theta_l - \theta_l^2)}{(N-1)(N-2)}. \quad (\text{II.52})$$

Для стратегии "rand/rand/1/bin"

$$E(v_l^2) = \frac{1}{N} \sum_{l=0}^{l=N-1} E_l(v_l^2) = \bar{\theta}^2 + 2F^2 \frac{N}{N-1} (\bar{\theta}^2 - \bar{\theta}^2). \quad (\text{II.53})$$

Теперь рассчитаем $E(\overline{P_v^2})$. Для стратегии "rand/rand/1/bin"

$$\begin{aligned} E(\overline{P_v^2}) &= \frac{1}{N} \left[\sum_{k=0}^{k=N-1} \theta_k^2 - E_l(\theta_l^2) + E_l(v_l^2) \right] = \\ &= \bar{\theta}^2 + \frac{1}{N} \left[\bar{\theta}^2 + 2F^2 \frac{N}{N-1} (\bar{\theta}^2 - \bar{\theta}^2) - \bar{\theta}^2 \right] = \bar{\theta}^2 + \frac{2F^2}{N-1} (\bar{\theta}^2 - \bar{\theta}^2). \end{aligned} \quad (\text{II.54})$$

Следующее, что нужно оценить, это $E(\overline{P_v^2})$. Для расчета нам понадобятся соотношения²⁸:

$$E(v_l - \theta_l) = E\left(E_l\left(\frac{N}{N-1}(\bar{\theta} - \theta_l) - \theta_l\right)\right) = \frac{N}{N-1} E(\bar{\theta} - \theta_l) = 0; \quad (\text{II.55})$$

$$E(v_l \theta_l) = E_l(\theta_{r_l} \cdot \theta_l) + 0 = \frac{N}{N-1} (\bar{\theta} - \frac{1}{N} \theta_l) \theta_l; \quad (\text{II.56})$$

$$E(v_l \theta_l) = \frac{N}{N-1} \left(\bar{\theta}^2 - \frac{1}{N} \bar{\theta}^2 \right); \quad (\text{II.57})$$

$$\begin{aligned} E(v_l - \theta_l)^2 &= \bar{\theta}^2 + 2F^2 \frac{N}{N-1} (\bar{\theta}^2 - \bar{\theta}^2) - 2 \frac{N}{N-1} \left(\bar{\theta}^2 - \frac{1}{N} \bar{\theta}^2 \right) + \bar{\theta}^2 = \\ &= (1 + F^2) \frac{2N}{N-1} (\bar{\theta}^2 - \bar{\theta}^2). \end{aligned} \quad (\text{II.58})$$

²⁵Если не учитывать $r, p, q \neq l$, то $E(\theta_r) = \bar{\theta}$ и $E(\theta_r^2) = \bar{\theta}^2$, а $E(\theta_r \theta_p | r \neq p) = \frac{N\bar{\theta}^2 - \bar{\theta}^2}{N-1}$.

²⁶Если не учитывать $r, p, q \neq l$, то $E(v_l) = \bar{\theta}$ для любых стратегий.

²⁷Если не учитывать $r, p, q \neq l$, то $E(v_l^2) = \bar{\theta}^2 - \frac{2NF^2}{N-1} (\bar{\theta}^2 - \bar{\theta}^2)$.

²⁸Если не учитывать $r, p, q \neq l$, то $E(v_l - \theta_l) = 0$, $E(v_l \theta_l) = \bar{\theta}^2$, $E(v_l - \theta_l)^2 = 2 \left(1 + F^2 \frac{2N}{N-1} \right) (\bar{\theta}^2 - \bar{\theta}^2)$.

Таким образом²⁹:

$$\begin{aligned} E\left(\overline{P_v^2}\right) &= E\left[\frac{1}{N^2}\left(\sum_{k=0}^{k=N-1}\theta_k - \theta_l + v_l\right)^2\right] = \\ &= \overline{\theta^2} + \frac{2}{N}\overline{\theta}E(v_l - \theta_l) + \frac{1}{N^2}E(v_l - \theta_l)^2 = \overline{\theta^2} + \frac{2(F^2 + 1)}{N(N-1)}(\overline{\theta^2} - \overline{\theta}^2). \end{aligned} \quad (\text{II.59})$$

В результате для каждой компоненты вариации³⁰:

$$\begin{aligned} E(\text{Var}(P_v))_i &= E\left(\overline{P_v^2}\right) - E\left(\overline{P_v}\right)^2 = \\ &= \overline{\theta^2} + \frac{2F^2}{N-1}(\overline{\theta^2} - \overline{\theta}^2) - \left[\overline{\theta^2} + \frac{2(F^2 + 1)}{N(N-1)}(\overline{\theta^2} - \overline{\theta}^2)\right] = \\ &= \left(1 + \frac{2F^2}{N} - \frac{2}{N(N-1)}\right)(\overline{\theta^2} - \overline{\theta}^2), \end{aligned} \quad (\text{II.60})$$

и мы доказали утверждение 1.2.3.

(б) В популяции P_u один из элементов u_l равен

$$u_{l,j} = \begin{cases} v_{l,j} & \text{с вер. } C_r; \\ \theta_{l,j} & \text{с вер. } (1 - C_r), \end{cases} \quad (\text{II.61})$$

а все остальные элементы совпадают с элементами популяции P_θ .

Используя свойства средних для случайных величин, получаем:

$$\begin{aligned} E\left(\overline{P_u^2}\right) &= \frac{1}{N}E\left(\left(\sum_{k=0}^{k=N-1}\theta_k^2\right) - \theta_l^2 + u_l^2\right) = \\ &= \overline{\theta^2} + \frac{1}{N}E[-\theta_l^2 + C_r v_l^2 + (1 - C_r)\theta_l^2] = \overline{\theta^2} + \frac{C_r}{N}E[v_l^2 - \theta_l^2] = \\ &= \overline{\theta^2} + \frac{C_r}{N}\left[\overline{\theta^2} + 2F^2\frac{N}{N-1}(\overline{\theta^2} - \overline{\theta}^2) - \overline{\theta^2}\right] = \overline{\theta^2} + \frac{2F^2 C_r}{N-1}(\overline{\theta^2} - \overline{\theta}^2). \end{aligned} \quad (\text{II.62})$$

С другой стороны, для $E\left(\overline{P_u^2}\right)$ имеем³¹:

$$\begin{aligned} E\left(\overline{P_u^2}\right) &= E\left[\frac{1}{N^2}\left(\sum_{k=0}^{k=N-1}\theta_k - \theta_l + u_l\right)^2\right] = \\ &= \frac{1}{N^2}E\left(C_r(N\overline{\theta} - \theta_l + v_l)^2 + (1 - C_r)(\overline{\theta} - \theta_l + \theta_l)^2\right) = \\ &= C_r\left(\overline{\theta^2} + \frac{2}{N}\overline{\theta}E(v_l - \theta_l) + \frac{1}{N^2}E(v_l - \theta_l)^2\right) + (1 - C_r)\overline{\theta^2} = \\ &= \overline{\theta^2} + \frac{2C_r(F^2 + 1)}{N(N-1)}(\overline{\theta^2} - \overline{\theta}^2). \end{aligned} \quad (\text{II.63})$$

Таким образом³²:

$$\begin{aligned} E(\text{Var}(P_u))_j &= E\left(\overline{P_u^2}\right) - E\left(\overline{P_u}\right)^2 = \\ &= \overline{\theta^2} + \frac{2F^2 C_r}{N-1}(\overline{\theta^2} - \overline{\theta}^2) - \left[\overline{\theta^2} + \frac{2C_r(F^2 + 1)}{N(N-1)}(\overline{\theta^2} - \overline{\theta}^2)\right] = \\ &= \left(1 + \frac{2F^2 C_r}{N} + \frac{2C_r}{N(N-1)}\right)\text{Var}(P_\theta)_j. \end{aligned} \quad (\text{II.64})$$

²⁹Если не учитывать $r, p, q \neq l$, то $E\left(\overline{P_v^2}\right) = \overline{\theta^2} + \frac{2}{N^2}\left(1 + \frac{NF^2}{N-1}\right)(\overline{\theta^2} - \overline{\theta}^2)$.

³⁰Если не учитывать $r, p, q \neq l$, то $E(\text{Var}(P_v))_i = \left(1 + \frac{2F^2}{N} - \frac{2}{N^2}\right)(\overline{\theta^2} - \overline{\theta}^2)$.

³¹Если не учитывать $r, p, q \neq l$, то $E\left(\overline{P_u^2}\right) = \overline{\theta^2} + \frac{2C_r}{N^2}\left(1 + F^2\frac{N}{N-1}\right)(\overline{\theta^2} - \overline{\theta}^2)$.

³²Если не учитывать $r, p, q \neq l$, то $E(\text{Var}(P_u))_j = \left(1 + \frac{2F^2 C_r}{N} + \frac{2C_r}{N^2}\right)\text{Var}(P_\theta)_j$.

и мы доказали утверждение 1.2.3. ■

Доказательство теоремы 1.2.3 (для rand/best/1/bin) (а) В популяции P_v один из элементов v_l равен

$$v_l = \theta_0 + F(\theta_p - \theta_q),$$

где l, r, p и q – различны, θ_0 – лучший элемент в текущей популяции, а все остальные элементы совпадают с соответствующими элементами популяции P_θ . В рассматриваемых вариантах алгоритма операции мутации и кроссовера действуют независимо на каждую из компонент. Поэтому в доказательстве можно ограничиться покомпонентным анализом. Для краткости далее в записи будем опускать знак вектора и индекс компоненты. Тогда покомпонентно³³:

$$E(\theta_{r=0}) = \theta_0, \quad E(\theta_p|p \neq 0) = E(\theta_q|q \neq 0) = \frac{N}{N-1}(\bar{\theta} - \frac{1}{N}\theta_0), \quad (\text{II.65})$$

$$E(\theta_{r=0}^2) = \theta_0^2, \quad E(\theta_p^2|p \neq 0) = E(\theta_q^2|q \neq 0) = \frac{N}{N-1}\left(\bar{\theta}^2 - \frac{1}{N}\theta_0^2\right), \quad (\text{II.66})$$

$$E(\theta_p\theta_q|p \neq 0, q \neq 0, q \neq p) = \frac{N(N\bar{\theta}^2 - \bar{\theta}^2) - 2(N\bar{\theta}\theta_0 - \theta_0^2)}{(N-1)(N-2)}. \quad (\text{II.67})$$

Случайная переменная $E(\tilde{F}_l)$ независима от случайных переменных θ_i для любых i и: $E(\tilde{F}) = 0$, $\text{Var}(\tilde{F}_l) = F^2$.

Для стратегии "rand/best/1/bin" $E(v_l) = E_l(\theta_{r_l} = \theta_0) + 0 = \theta_0$. Рассчитаем³⁴:

$$\begin{aligned} E(v_l^2|p \neq 0, q \neq 0, q \neq p) &\equiv E(v_l^2) = \\ &= \theta_0^2 + 2F^2 \left[\frac{N}{N-2}(\bar{\theta}^2 - \bar{\theta}^2) - \frac{N(\theta_0 - \bar{\theta})^2}{(N-1)(N-2)} \right]. \end{aligned} \quad (\text{II.68})$$

Теперь рассчитаем $E(\overline{P_v^2})$. Для стратегии *rand/best/1/bin*

$$\begin{aligned} E(\overline{P_v^2}) &= \frac{1}{N} \left[\sum_{k=0}^{k=N-1} \theta_k^2 - E_l(\theta_l^2) + E_l(v_l^2) \right] = \\ &= \bar{\theta}^2 + \frac{1}{N} \left\{ \theta_0^2 - \bar{\theta}^2 + 2F^2 \left[\frac{N}{N-2}(\bar{\theta}^2 - \bar{\theta}^2) - \frac{N(\theta_0 - \bar{\theta})^2}{(N-1)(N-2)} \right] \right\}. \end{aligned} \quad (\text{II.69})$$

Следующее, что нужно рассчитать, это $E(\overline{P_v^2})$. Для расчета нам понадобятся соотношения ($l = \text{rand}()$)³⁵:

$$E\left(\sum_{k \neq l} \theta_k\right) = (N-1)\bar{\theta}, \quad (\text{II.70})$$

$$E\left(\theta_l \sum_{k \neq l} \theta_k\right) = N\bar{\theta}^2 - \bar{\theta}^2, \quad (\text{II.71})$$

$$E\left(v_l \sum_{k \neq l} \theta_k\right) = (N-1)\theta_0\bar{\theta}, \quad (\text{II.72})$$

$$E\left(\left(\sum_{k \neq l} \theta_k\right)^2\right) = (N^2 - 2N)\bar{\theta}^2 + \bar{\theta}^2. \quad (\text{II.73})$$

³³В данном доказательстве не учтено $l \neq 0 = r$, и $p, q \neq l$. Если же пренебречь тем, что $p, q \neq 0 = r$, то $E(\theta_r) = \bar{\theta}$ и $E(\theta_r^2) = \bar{\theta}^2$, а $E(\theta_p\theta_q|q \neq p) = \frac{N\bar{\theta}^2 - \bar{\theta}^2}{N-1}$.

³⁴Если не учитывать $p, q \neq 0$, то $E(v_l^2) = \bar{\theta}^2 - \frac{2NF^2}{N-1}(\bar{\theta}^2 - \bar{\theta}^2)$.

³⁵Если не учитывать $r, p, q \neq l$, то $E(v_l - \theta_l) = 0$, $E(v_l\theta_l) = \bar{\theta}^2$, $E(v_l - \theta_l)^2 = 2\left(1 + F^2\frac{2N}{N-1}\right)(\bar{\theta}^2 - \bar{\theta}^2)$.

Таким образом³⁶:

$$\begin{aligned}
E(\overline{P_v^2}) &= E\left[\frac{1}{N^2}\left(\sum_{k=0, k \neq l}^{k=N-1} \theta_k + v_l\right)^2\right] = \\
&= \frac{1}{N^2}\left[E\left(\left(\sum_{k=0, k \neq l}^{k=N-1} \theta_k\right)^2\right) + 2E\left(v_l \sum_{k=0, k \neq l}^{k=N-1} \theta_k\right) + E(v_l^2)\right] = \\
&= \frac{1}{N^2}\left\{(N^2 - 2N)\overline{\theta^2} + \overline{\theta^2} + 2(N-1)\theta_0\overline{\theta} + \theta_0^2 + 2F^2\left[\frac{N(\overline{\theta^2} - \overline{\theta}^2)}{N-2} - \frac{N(\theta_0 - \overline{\theta})^2}{(N-1)(N-2)}\right]\right\}
\end{aligned} \tag{II.74}$$

В результате, для каждой компоненты вариации³⁷:

$$\begin{aligned}
E(\text{Var}(P_v))_i &= E(\overline{P_v^2}) - E(\overline{P_v})^2 = \\
&= \left(1 + \frac{2F^2(N-1)}{N(N-2)} - \frac{N+1}{N^2}\right)(\overline{\theta^2} - \overline{\theta}^2) + \left(\frac{N-1}{N^2} - \frac{2F^2}{N(N-2)}\right)(\theta_0 - \overline{\theta})^2, \tag{II.75}
\end{aligned}$$

и мы доказали утверждение 1.2.3.

(б) В популяции P_u один из элементов u_l равен

$$u_{l,j} = \begin{cases} v_{l,j} & \text{с вер. } C_r; \\ \theta_{l,j} & \text{с вер. } (1 - C_r), \end{cases} \tag{II.76}$$

а все остальные элементы совпадают с элементами популяции P_θ . Используя свойства средних для случайных величин, получаем:

$$\begin{aligned}
E(\overline{P_u^2}) &= \frac{1}{N}E\left(\left(\sum_{k=0, k \neq l}^{k=N-1} \theta_k^2\right) + u_l^2\right) = \\
&= \frac{1}{N}E\left(\left(\sum_{k=0, k \neq l}^{k=N-1} \theta_k^2\right) + C_r v_l^2 + (1 - C_r)\theta_l^2\right) = E\left(\frac{1}{N}\left(\sum_{k=0}^{k=N-1} \theta_k^2\right)\right) + \frac{C_r}{N}E(v_l^2 - \theta_l^2) = \\
&= \overline{\theta^2} + \frac{C_r}{N}(\theta_0^2 - \overline{\theta^2}) + \frac{2F^2 C_r}{N}\left[\frac{N}{N-2}(\overline{\theta^2} - \overline{\theta}^2) - \frac{N(\theta_0 - \overline{\theta})^2}{(N-1)(N-2)}\right]. \tag{II.77}
\end{aligned}$$

³⁶Если не учитывать $r, p, q \neq l$, то $E(\overline{P_v^2}) = \overline{\theta^2} + \frac{2}{N^2}\left(1 + \frac{NF^2}{N-1}\right)(\overline{\theta^2} - \overline{\theta}^2)$.

³⁷Если не учитывать $r, p, q \neq l$, то $E(\text{Var}(P_v))_i = \left(1 + \frac{2F^2}{N} - \frac{2}{N^2}\right)(\overline{\theta^2} - \overline{\theta}^2)$.

С другой стороны, для $E(\overline{P_u^2})$ имеем:³⁸:

$$\begin{aligned}
E(\overline{P_u^2}) &= E\left[\frac{1}{N^2}\left(\sum_{k=0, k \neq l}^{k=N-1} \theta_k + u_l\right)^2\right] = \\
&= \frac{1}{N^2}\left(E\left(\left(\sum_{k=0, k \neq l}^{k=N-1} \theta_k\right)^2\right) + 2E\left(u_l \sum_{k=0, k \neq l}^{k=N-1} \theta_k\right) + E(u_l^2)\right) = \\
&= E\left(\frac{1}{N^2}\left(\sum_{k=0, k \neq l}^{k=N-1} \theta_k\right)^2\right) + 2C_r E\left(\frac{1}{N^2} v_l \sum_{k=0, k \neq l}^{k=N-1} \theta_k\right) + \\
&+ 2(1 - C_r)E\left(\frac{1}{N^2} \theta_l \sum_{k=0, k \neq l}^{k=N-1} \theta_k\right) + C_r E\left(\frac{1}{N^2} v_l^2\right) + (1 - C_r)E\left(\frac{1}{N^2} \theta_l^2\right) = \\
&= \dots = \frac{N - 2C_r \bar{\theta}^2}{N} + \frac{C_r \bar{\theta}^2}{N^2} + \frac{2C_r(N-1)}{N^2} \theta_0 \bar{\theta} + \frac{C_r}{N^2} E(v_l^2). \quad (\text{II.78})
\end{aligned}$$

Таким образом³⁹:

$$\begin{aligned}
E(\text{Var}(P_u))_j &= E(\overline{P_u^2}) - E(\overline{P_u})^2 = \\
&= \bar{\theta}^2 + \frac{C_r}{N} (\theta_0^2 - \bar{\theta}^2) + \frac{2F^2 C_r}{N} \left[\frac{N}{N-2} (\bar{\theta}^2 - \bar{\theta}^2) - \frac{N(\theta_0 - \bar{\theta})^2}{(N-1)(N-2)} \right] - \\
&\quad - \frac{N - 2C_r \bar{\theta}^2}{N} - \frac{C_r \bar{\theta}^2}{N^2} - \frac{2C_r(N-1)}{N^2} \theta_0 \bar{\theta} - \frac{C_r}{N^2} E(v_l^2) = \dots = \\
&= \left(1 + \frac{2F^2 C_r(N-1)}{N(N-2)} - \frac{C_r}{N^2}\right) (\bar{\theta}^2 - \bar{\theta}^2) + \left(\frac{(N-1)C_r}{N^2} - \frac{2F^2 C_r}{N(N-2)}\right) (\theta_0 - \bar{\theta})^2 = \\
&= \left(1 + \frac{2F^2 C_r(N-1)}{N(N-2)} - \frac{C_r}{N^2}\right) (\bar{\theta}^2 - \bar{\theta}^2) + \left(\frac{(N-1)C_r}{N^2} - \frac{2F^2 C_r}{N(N-2)}\right) (\theta_0 - \bar{\theta})^2, \quad (\text{II.79})
\end{aligned}$$

и мы доказали утверждение 1.2.3. \blacksquare

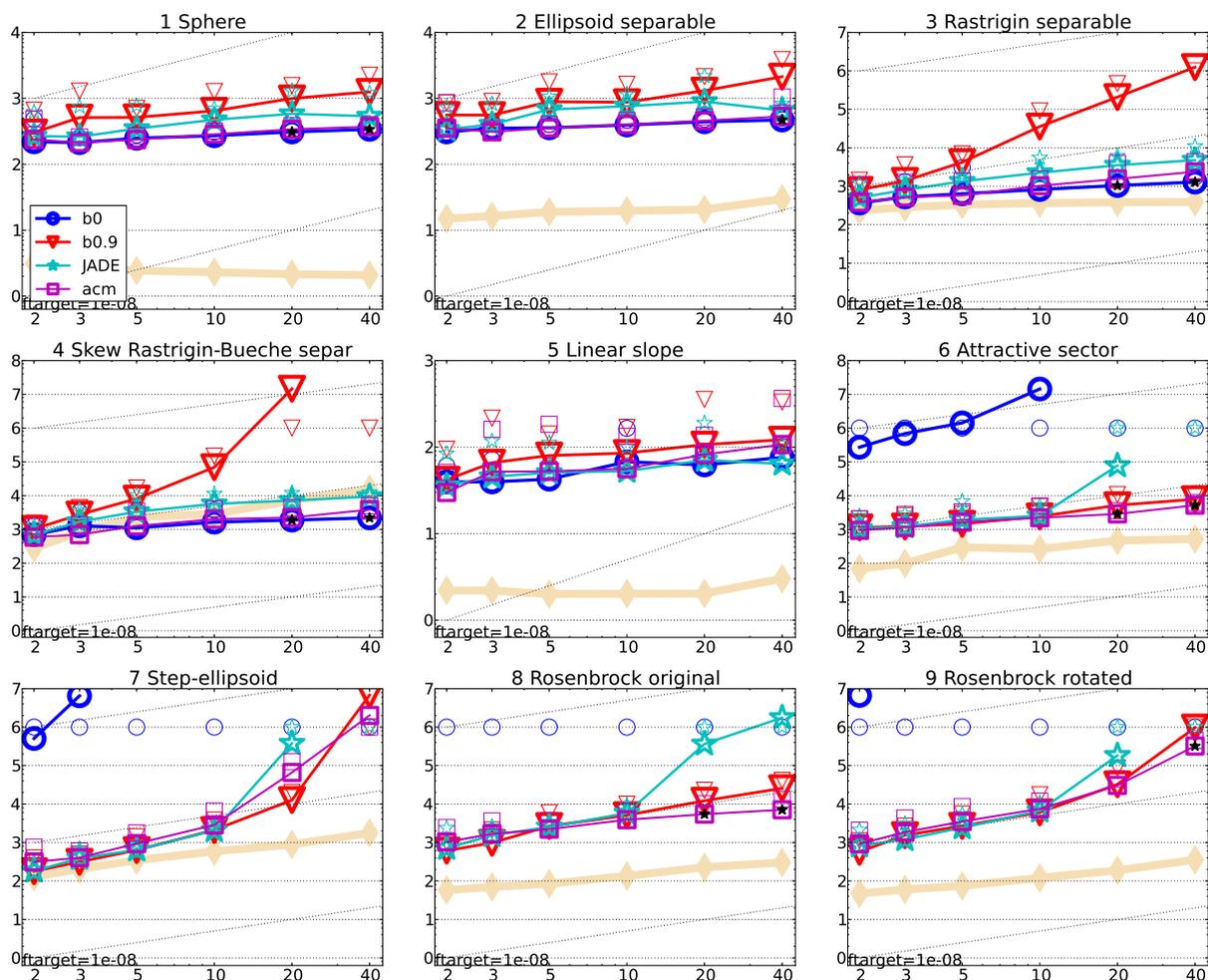
³⁸Если не учитывать $r, p, q \neq l$, то $E(\overline{P_u^2}) = \bar{\theta}^2 + \frac{2C_r}{N^2} \left(1 + F^2 \frac{N}{N-1}\right) (\bar{\theta}^2 - \bar{\theta}^2)$.

³⁹Если не учитывать $r, p, q \neq l$, то $E(\text{Var}(P_u))_j = \left(1 + \frac{2F^2 C_r}{N} + \frac{2C_r}{N^2}\right) \text{Var}(P_\theta)_j$.

П. III. Результативность АДЭ-АСМ метода

На рисунках ниже представлены численные результаты тестирования стратегий АДЭ, АДЭ-АКМ и алгоритма JADE, который использует адаптивную схему для кроссовера и отнсительно которого было показано, что он имеет более высокую производительность по сравнению с другими адаптивными DE-алгоритмами [69, 80] в соответствии с ВВОВ-2012 [41]. Попытка определялась как успешная и вычисления заканчивались, если алгоритм находил значение целевой функции ниже $f(\vec{\theta}^*) + 10^{-8}$. Ожидаемое время работы алгоритма (ERT) при ненулевой вероятности успеха рассчитывается, как отношение числа вызовов целевой функции во всех испытаниях к числу успешных испытаний.

По горизонтали: размерность задачи. По вертикали: \log_{10} от ожидаемого времени работы (Estimated running time, ERT), измеренного в количестве вычислений функции f , деленном на размерность целевой функции и умноженном на 10^{-8} . Различные символы соответствуют различным алгоритмам, приведенным в легенде для функций f_1 и f_{24} . Символы, изображенные тонкой линией показывают максимальное количество оценок целевой функции из самого длинного испытания деленное на размерность. Горизонтальные линии означают линейное масштабирование, наклонные пунктирные линии показывают квадратичную зависимость. Черные звезды указывают статистически лучший результат по сравнению со всеми другими алгоритмами.



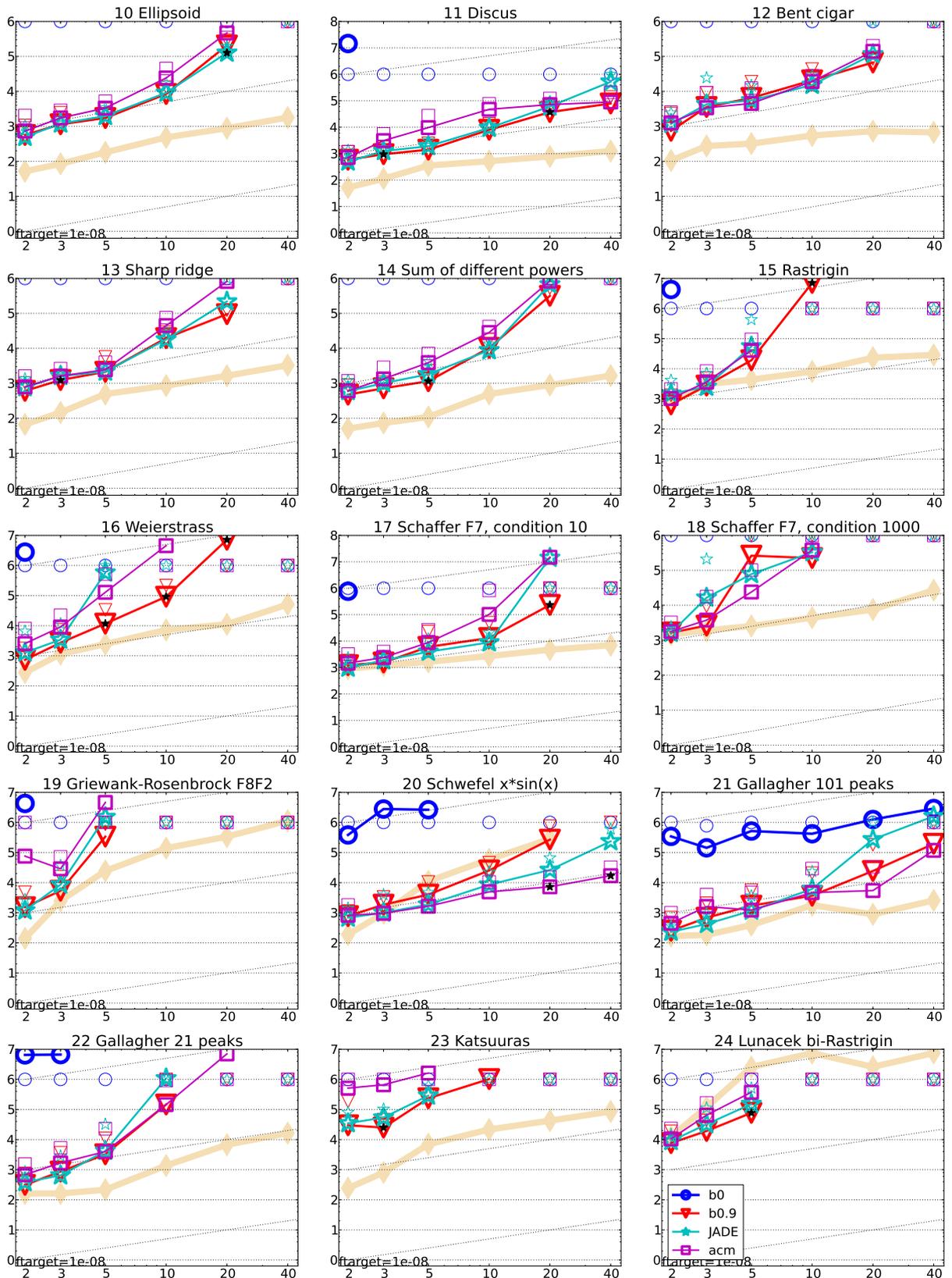
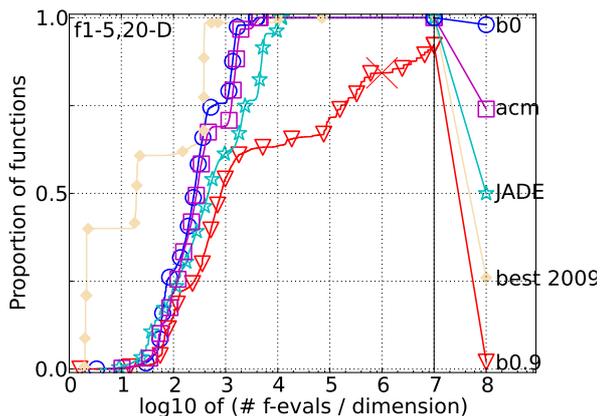
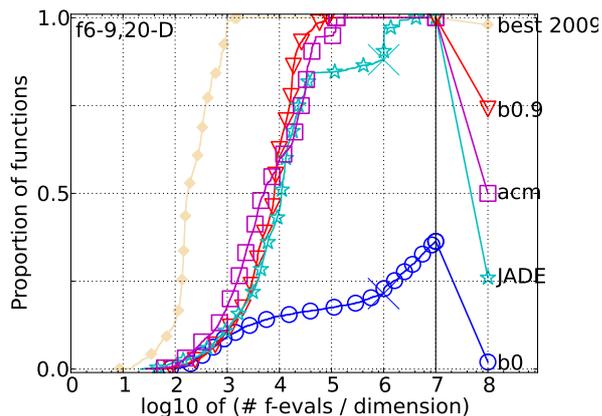


Рис. III.1. По горизонтали: размерность задачи D . По вертикали: $\log_{10}(ERT/D \cdot 10^{-8})$.
 Легенда: \circ :bin0 — АДЭ-Р с кроссовером $C_r = 0,1$; ∇ :bin0.9 — АДЭ-Р с кроссовером $C_r = 0,9$; $*$:JADE — адаптивный вариант ДЭ — JADE; \square :acm — АДЭ с кроссовером, задаваемым адаптивной корреляционной матрицей АДЭ-АКМ [10].

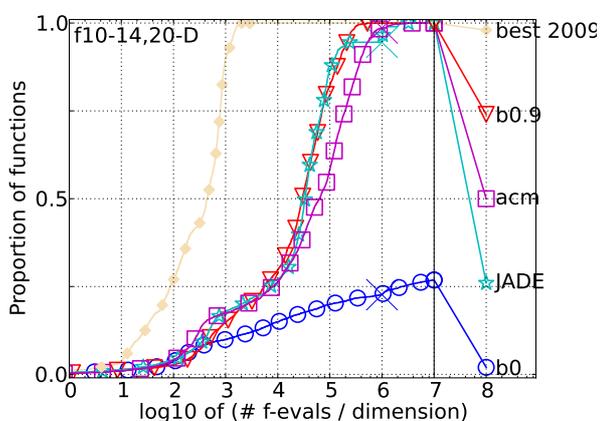
сепарабельные функции



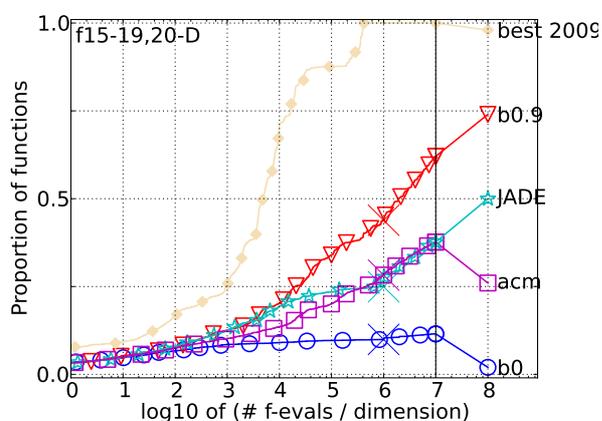
умеренные функции



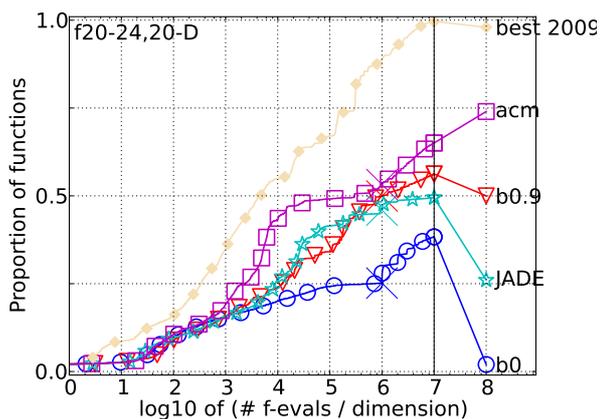
плохообусловленные функции



мультимодальные функции



слабо структурированные мультимодальные функции



все функции

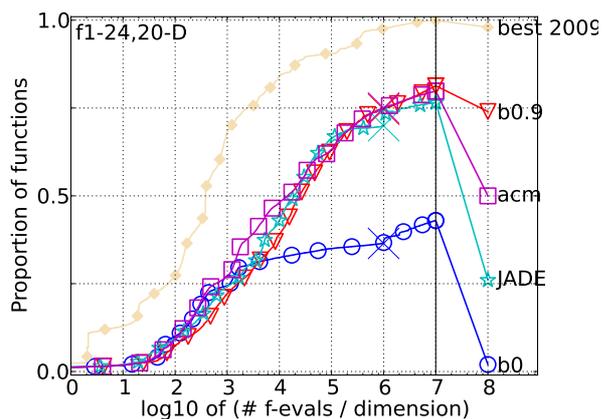


Рис. III.2. Горизонтальная ось: $\log_{10}(N_{FE}/D)$ (допустимое количество оценок целевой функции деленное на размерность задачи). Вертикальная ось: доля решенных функций из набора (15 попыток для каждой функции с попаданием в окрестность минимума с заданной точностью $10^{[-8..2]}$) для всех функций и подгрупп в классе задач с размерностью $D = 20$. Линия «best-2009» соответствует наилучшему ERT среди тестируемых в ходе WBOB-2009 алгоритмов для каждой задачи.

Легенда: \circ :bin0 — АДЭ-Р с кроссовером $C_r = 0,1$; ∇ :bin0.9 — АДЭ-Р с кроссовером $C_r = 0,9$; $*$:JADE — адаптивный вариант ДЭ — JADE; \square :acm — адаптивный АДЭ с кроссовером, задаваемым адаптивной корреляционной матрицей АДЭ-АКМ [10]

Благодарности

Работа над первой частью диссертации частично поддержана Программой сотрудничества между ОИЯИ и Болгарским научным центром. Исследования, представленные в главе 2 выполнялись при частичной финансовой поддержке грантов РФФИ 12-01-00396 и 13-01-00060, а результаты, полученные в главе 3 — при частичной финансовой поддержке Российского научного фонда (грант № 14 12 00516).

Я искренне благодарна моим научным руководителям: Елене Валерьевне Земляной, за постоянное внимание к выполняемым исследованиям, за поддержку и помощь во время написания диссертации; Киселеву Михаилу Алексеевичу за постановку задач, решение которых продемонстрировало возможности нового алгоритма, за полезные обсуждения и поддержку.

Особую признательность выражаю соавторам работ, лежащих в основе диссертации, Елене Земляной, Михаилу Жабицкому, Валерию и Константину Лукьяновым, Михаилу Киселеву, Виктору Аксенову, за плодотворное сотрудничество, за совместную работу и полезные обсуждения.

Считаю своим приятным долгом поблагодарить Пузынина И. В., Иванова В. В., Севастьянова Л. А., Гердта В. П. и Чулуунбаатар О. за внимание к работе, полезные советы и конструктивные критические замечания.

Благодарю весь коллектив сотрудников Отдела вычислительной физики ЛИТ за готовность помочь, обстановку взаимопонимания и творчества, за идеи и ценные советы, которыми они щедро делились. Благодаря им я узнала много нового о различных вычислительных методах.

Я признательна руководству ЛИТ ОИЯИ за создание благоприятной и творческой обстановки для выполнения научной работы в Лаборатории.

Мне дорог интерес к моей работе и поддержка со стороны моих родных. Благодаря Татьяне Яковлевне Жабицкой, многие места в диссертации приобрели ясность изложения и четкость формулировок, а Вячеслав Михайлович помог мне лучше понять и осмыслить полученные результаты. Я благодарна моим родителям, Светлане и Игорю Вишневым, за их веру в меня. Моего мужа Михаила благодарю за плодотворную совместную работу, помощь в постановке задач и продуктивные обсуждения, а также за любовь, поддержку и почти безграничное терпение. В заключение хочу сказать спасибо Никите, Славе и Сереже за их терпение и понимание.