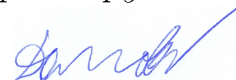


На правах рукописи



Долгов Сергей Владимирович

АЛГОРИТМЫ И ПРИМЕНЕНИЯ ТЕНЗОРНЫХ
РАЗЛОЖЕНИЙ ДЛЯ ЧИСЛЕННОГО РЕШЕНИЯ
МНОГОМЕРНЫХ НЕСТАЦИОНАРНЫХ ЗАДАЧ

01.01.07 — Вычислительная математика

АВТОРЕФЕРАТ

*диссертации на соискание ученой степени
кандидата физико-математических наук*

Москва 2014

Работа выполнена в федеральном государственном бюджетном учреждении науки Институте вычислительной математики Российской академии наук (ИВМ РАН).

Научный руководитель: член-корреспондент РАН,
доктор физико-математических наук, профессор
Тыртышников Евгений Евгеньевич.

Официальные оппоненты: Истомин Яков Николаевич,
доктор физико-математических наук, профессор.
Федеральное государственное бюджетное учреждение науки Физический институт им. П.Н. Лебедева Российской академии наук,
зам. руководителя отдела теоретической физики.

Третьяков Алексей Анатольевич,
доктор физико-математических наук.
Федеральное государственное бюджетное учреждение науки Вычислительный центр им. А.А. Дородницына Российской академии наук,
ведущий научный сотрудник.

Ведущая организация: Федеральное государственное бюджетное учреждение науки Институт прикладной математики им. М.В. Келдыша Российской академии наук.

Защита состоится ____ _____ 2014 г. в ____ час. ____ мин. на заседании диссертационного совета Д 002.045.01 при федеральном государственном бюджетном учреждении науки Институте вычислительной математики Российской академии наук (ИВМ РАН) по адресу: 119333, г. Москва, ул. Губкина, 8, ауд. 729.

С диссертацией можно ознакомиться в библиотеке и на сайте ИВМ РАН
<http://www.inm.ras.ru>.

Автореферат разослан ____ _____ 2014 г.

Ученый секретарь
диссертационного совета
д.ф.-м.н.

Бочаров Геннадий Алексеевич

Общая характеристика работы

Объект исследования и актуальность работы

Диссертация посвящена численному решению многомерных задач методами тензорных разложений. Например, мы хотим одновременно описать несколько взаимодействующих под действием случайных сил тел. Совместная функция плотности вероятности задается в d -мерном пространстве, и, как правило, требуют n^d значений после дискретизации, где d составляет количество всех координат всех частиц, а n – число степеней свободы по каждой размерности. Под “размерностью” мы и будем иметь в виду количество *конфигурационных координат* d в пространстве состояний системы. В принципе, даже случаи $d = 3$ или $d = 2$ можно рассматривать как “многомерные”, если n очень велико. Ситуация становится еще более драматичной, если физическая или математическая модель предусматривает работу с размерностями порядка десятков, сотен и более. Если исключить тривиальный случай, когда $n = 1$ для большинства координат (в этом случае нет смысла рассматривать соответствующие переменные вообще), экспоненциальный рост вычислительной сложности с размерностью делает невозможными непосредственные расчеты на любой суперкомпьютере. Например, в квантовом мире, частицы со спином $1/2$ (в определенных условиях, например в магнитно-резонансных экспериментах, рассмотрение электронов и ядер может быть ограничено спиновыми эффектами) обладают только $n = 2$ состояниями для каждой частицы, “спин вверх” и “спин вниз”. Однако, простая линейная цепочка из $d = 100$ взаимодействующих спинов (что рассматривается как модельная задача в квантовой физике) описывается уже волновой функцией с $2^{100} \sim 10^{30}$ неизвестными.

Это явление экспоненциального роста сложности в зависимости от количества конфигурационных координат называется *проклятием размерности*. Таким образом, единственным способом решения таких задач является работа только с небольшой *эффективной* частью дискретной информации, что требует гораздо меньше памяти и вычислительных операций, чем исходный массив n^d чисел.

Среди всех подходов для сжатия информации, так называемых методов с *разреженными данными*, мы можем выделить проблемно-ориентированные и общие классы. Первый класс предполагает и существенно использует специфические свойства задачи, такие как гладкость соответствующих функций, определенные правила вычисления статистических величин, и так далее. Среди наиболее известных и широко применяемых методов, например, методы Монте Карло, разреженные сетки Смоляка и т.д. Общие методы не используют в явном виде физический смысл задачи или выходных данных, полагаясь вместо этого на чисто математические инструменты для представления всего многомерного объекта с помощью правильно выбранного отображения небольшого количества данных.

Примечательно эффективным представителем класса общих методов сжатия данных является концепция разделения переменных *тензорными произведениями*, испытавшая быстрое развитие в последнее десятилетие, разрабатываемая и в настоящее время, в том числе в текущей работе. Общая идея состоит в том, чтобы представить (или приблизить) большой многомерный массив с помощью комбинации произведений меньших массивов с меньшим количеством степеней свободы. Важно отметить, что существуют определенные разложения и методы, которые требуют только исходных данных и используют только алгебраические инструмен-

ты (например, сингулярное разложение) для выделения редуцированного набора параметров.

Ключевым моментом в разделении переменных является представление многомерной функции (или ее дискретного аналога, *тензора*) в виде произведения одномерных объектов, т.е. $x(i_1, \dots, i_d) = x^{(1)}(i_1)x^{(2)}(i_2) \cdots x^{(d)}(i_d)$. Это разложение в *прямое произведение* редко выполняется точно. Обычно используют определенную комбинацию тензорных произведений, но конкретные *численные* алгоритмы нахождения такой комбинации существенно отличаются по своим свойствам.

Так, “жадные” методы тензорных аппроксимаций часто стагнируют на каком-то (неудовлетворительно большом) уровне ошибки. Методы оптимизации общего вида страдают от некорректности задачи аппроксимации тензора размерности больше двух суммой R прямых произведений. В качестве альтернативы были предложены *рекуррентные* двумерные разложения: Tensor Tree Networks, Hierarchical Tucker, Matrix Product States/Tensor Train. Последний представляет данные в виде полилинейной комбинации трехмерных массивов, и требует объема памяти $\mathcal{O}(dnr^2)$, что дает возможность избежать проклятия размерности, если ранг r не очень большой.

Любая тензорная структура требует численных методов для аппроксимации данных и других операций, таких как решение линейных систем. Одними из самых эффективных явились методы переменных направлений, а точнее семейство алгоритмов *Density Matrix Renormalization Group*, разработанных в сообществе квантовой физики для моделирования волновых функций спиновых систем в форматах тензорных произведений. Тем не менее, даже функции ошибки вида $\|x - x_\star\|^2$, по отношению к элементам тензорного разложения, могут иметь многочисленные локальные минимумы, и алгоритмы DMRG с высокой вероятностью возвращают локальный, но не глобальный минимизатор ошибки.

Другим подходом является адаптация классических итерационных алгоритмов (например, GMRES) к хранению и вычислению векторов в виде тензорных произведений. Этот путь, однако, также не является достаточно надежным, поскольку чем ближе приближенное решение к точному, тем более сложной и осциллирующей является структура ошибки и других вспомогательных (в т.ч. Крыловских) векторов, которые требуют либо грубого приближения, либо большого объема памяти для хранения их тензорных форматов.

Цели диссертационной работы

1. Разработка нового вычислительного метода для решения больших систем линейных уравнений с представлением данных в формате тензорных произведений. Теоретический анализ его свойств сходимости.
2. Применение и проверка нового метода на задачах расчета стохастической химической кинетики (основное кинетическое уравнение) и задаче моделирования Фарлей-Бунемановской неустойчивости в плазме ионосферы Земли (уравнение Власова).
3. Анализ результатов, сравнение нового алгоритма с ранее существующими техниками, сравнение модельных численных данных с имеющимися в литературе экспериментами.

Основные положения диссертации

Главным результатом данной диссертации является вычислительный метод, который сочетает в себе сильные стороны как оптимизационных тензорных алгоритмов переменных направлений, так и классических приближенных итерационных схем. В процессе DMRG итерации, мы *расширяем* тензорный формат решения с помощью тензорного формата приближенной невязки. Это обеспечивает замечательную взаимную поддержку классических методов и методов переменных направлений (DMRG). Так как решение адаптируется в процессе DMRG оптимизации по элементам формата, даже очень грубые приближения невязки (причем без других Крыловских векторов) дают высокую точность решения. С другой стороны, подключение информации о глобальной невязке в локальных шагах процесса переменных направлений обеспечивает последний правильными градиентными направлениями, помогая ему избегать локальных минимумов. Новый АМЕп (Alternating Minimal Energy) метод обладает доказанной геометрической сходимостью с точки зрения глобальных элементов тензора, аналогично методу градиентного спуска. Примечательно, что практическая скорость сходимости оказывается намного быстрее, чем теоретические оценки на основе градиентного спуска, что делает этот метод надежным даже для несимметричных линейных систем, наподобие метода полной ортогонализации.

Приложения, рассматриваемые в этой диссертации, принадлежат семейству моделей, использующих совместное распределение вероятностей. Первым примером является *уравнение Власова*, возникающее в гибридной модели *Фарлей-Бунемановской неустойчивости* в плазме ионосферы земли.

Другое приложение имеет решающую роль в высокоточном стохастическом моделировании клеточных явлений, жизненных циклов вирусов и других микробиологических процессов, где случайные флуктуации вносят существенный вклад. Это возникает в случае, когда количество реагирующих молекул очень мало по сравнению с объемом всей системы, и столкновения между молекулами происходят в случайные моменты времени. В результате кинетика подчиняется дискретному марковскому процессу, который опять же может быть описан многомерной функцией вероятности. В отличие от уравнения Власова, конфигурационные состояния здесь с самого начала дискретные, и функция вероятности подчиняется разностному *основному кинетическому уравнению*.

Обе упомянутые модели описывают эволюцию функции вероятности во времени. Таким образом, в качестве третьего приложения, мы покажем, как *схемы интегрирования по времени*, например, Эйлера или Кранка-Николсон, могут быть ускорены методами тензорных произведений. Мы рассматриваем переменную времени в качестве дополнительного измерения. Затем используем тензорный формат QTT, что дает логарифмическую вычислительную сложность в зависимости от количества временных шагов.

Представленные задачи в конечном итоге сводятся к большим линейным системам. Мы выведем несколько аналитических тензорных представлений, показывая, что матрицы и правые части в этих системах могут быть предоставлены или эффективно приближены в тензорных форматах. Остается вопрос, как теперь вычислить их решения. В численных экспериментах, мы демонстрируем, что новый итерационный алгоритм АМЕп является более быстрым и точным, чем ранее известные методы, и, наконец, может претендовать на роль многоцелевого инстру-

мента для различных задач, которые в принципе подходят для концепции разделения переменных.

Достоверность научных положений и методология

Поскольку результатом диссертации является математический инструмент (алгоритм), основными способами оценки достоверности работы и результатов являются математические методы. Так, для подтверждения сходимости алгоритма использовалось измерение текущей невязки, что известным образом дает информацию о корректности решения, а также сравнение самих решений, полученных в результате запусков алгоритмов с разными допусками точности. В тех примерах прикладных задач, для которых в литературе были доступны результаты экспериментов или моделирования другими методами, проводилось сравнение их с расчетами, проведенными автором. Проведенные исследования подтверждают корректность и применимость новых предложенных методов, при их меньших вычислительных затратах.

Научная новизна

Предложенный метод АМЕп является первым алгоритмом для решения системы уравнений в представлении тензорными произведениями, который обладает теоретически доказанной оценкой глобальной сходимости, и при этом является эффективным на практике. В основе АМЕп алгоритма лежит сочетание метода градиентного спуска и метода переменных направлений (DMRG). Однако, в отдельности метод градиентного спуска сходится весьма медленно и требует достаточно точной аппроксимации невязки (градиента), что приводит к большой вычислительной сложности. Отдельные шаги методов DMRG весьма эффективны, но итоговый результат может давать очень плохое приближение для точного решения.

С помощью нового метода было впервые получено полное решение основного кинетического уравнения для λ -фага с высокой точностью. Классический способ аппроксимации решения основного кинетического уравнения состоит в усреднении большого набора случайных реализаций. Он хорошо применим в случаях, когда требуется невысокая точность. С ростом числа реализаций M метод сходится с достаточно медленной скоростью $\mathcal{O}(M^{-1/2})$, что отрицательно сказывается на вычислительной сложности в высокоточных расчетах. В качестве альтернативы, были предложены и методы тензорных представлений. Однако ранее существовавшие алгоритмы позволяли моделировать только небольшие промежутки времени, что непригодно для реальных задач в системной биологии.

Новый метод также принес существенный вклад в ускорение расчета Фарлей-Бунемановской неустойчивости, позволив провести моделирование четырехмерного уравнения Власова на персональном компьютере, тогда как затраты памяти на хранение решения в виде обычного четырехмерного массива неизбежно требовали применения распределенных высокопроизводительных систем.

Научная и практическая значимость полученных результатов

Предложенный подход может применяться в различных задачах моделирования стохастических и квантовых систем как метод достаточной степени общности. Так,

для того, чтобы проверить некоторый новый случай системы уравнений, нужно уметь представлять матрицу и правую часть в тензорных произведениях, но сам алгоритм ее решения не требует изменений. Поэтому, даже при наличии специализированных эффективных методов для конкретных задач, методы тензорных представлений могут использоваться для проверочных целей.

Другой особенностью малоранговых аппроксимаций являются весьма умеренные требования к памяти, хотя число операций, требуемых для решения задачи, может быть большим. Примером является моделирование Фарлей-Бунемановской неустойчивости в данной диссертации, где вычислительные времена с использованием стандартной схемы расщепления и разделения переменных являются качественно сравнимыми, но затраты памяти отличаются на полтора порядка. Важное применение предложенных методов ожидается в задачах магнитного резонанса и квантовом управлении. Предварительные результаты моделирования ядерного магнитного резонанса белков (задача сводится к расчету квантовых конфигураций спиновых систем) показывают высокую эффективность представлений тензорными произведениями и АМЕп алгоритма для решения уравнений с большим количеством переменных.

Апробация результатов диссертации и публикации

Основные результаты работы апробированы в следующих докладах на конференциях и семинарах.

- [1] Solution of the chemical master equation by the separation of variables and alternating optimization methods. *European Conference on Mathematical and Theoretical Biology*, Gothenburg, June 16, 2014.
- [2] A new family of the alternating linear optimization schemes in tensor product representations. *Seminar of the group Computational Methods in Systems and Control Theory*, Max Planck Institute Magdeburg, March 04, 2014.
- [3] Alternating minimal energy methods for linear systems in higher dimensions. Part II: Faster algorithm and application to nonsymmetric systems. *ENUMATH Conference*, EPFL Lausanne, August 26-30, 2013.
- [4] Alternating minimal energy methods for linear systems in higher dimensions. Part II: Faster algorithm and application to nonsymmetric systems. *MAFELAP*, Brunel University, London, June 11, 2013.
- [5] Fast adaptive alternating linear schemes in higher dimensions. Part 2: eigenproblems. *NASCA Conference*, University of Calais, June 24, 2013.
- [6] Fast adaptive tensor product approach to eigenvalue problems in higher dimensions. *Seminar of the Department of Chemistry*, University of Southampton, June 27, 2013.
- [7] Alternating minimal energy methods for linear systems in higher dimensions. Part II: Faster algorithm and application to nonsymmetric systems. *Swiss numerics colloquium*, EPFL, Lausanne, April 05, 2013.

- [8] Fast adaptive alternating linear solvers. Implementation hints and application to Fokker-Planck and master equations. *Workshop on algorithms for high-dimensional problems in quantum chemistry*, University Southampton, February 26, 2013.
- [9] Alternating minimal residual methods for linear systems in higher dimensions. Part II: heuristics and experiments. *29 GAMM Seminar on Uncertainty Quantification*, MPI MIS, Leipzig, January 22, 2013.
- [10] Advanced tensor representation and solution techniques with application to Fokker-Planck and master equations. *CMAM-5*, Humbolt University, Berlin, August 16–18, 2012.
- [11] Advantages and difficulties of use of tensor methods in solution to the Fokker-Planck equation. *28 GAMM Seminar on Analysis and Numerical Methods in Higher Dimensions*, MPI MIS, Leipzig, January 16–18, 2012.
- [12] A gray-box DMRG algorithm for tensor structured solution to linear systems. *17th Conference of the International Linear Algebra Society*, University Braunschweig, August 26, 2011.
- [13] On a solution to a parabolic equation in the QTT format using the DMRG approach. *4th Workshop on High-Dimensional Approximation*, University Bonn, June 27, 2011.
- [14] Use of the DMRG scheme for structured linear system solution. *3rd International Conference on Matrix Methods in Mathematics and Applications*, ИВМ РАН, Москва, 24 июня, 2011.
- [15] Linear solvers in TT formats. *27th GAMM-Seminar on Approximation of Multiparametric functions*, MPI MIS, Leipzig, January 26–28, 2011.
- [16] TT-GMRES: о решении систем линейных уравнений в тензорных форматах. *53 научная конференция МФТИ*, Москва, Ноябрь 26–28, 2010.

По результатам работы опубликовано 6 статей в международных рецензируемых журналах, и 1 статья в материалах конференций [3].

Структура диссертации

Диссертация состоит из введения, пяти глав, заключения, списка обозначений и списка литературы. Текст изложен на 161 странице, содержит 25 рисунков. Список литературы включает 253 наименования.

Благодарности

Эта диссертация содержит основные результаты моей научной работы с 2011 по 2014 год. Прежде всего, я хотел бы выразить благодарность моему научному руководителю, профессору Евгению Евгеньевичу Тыртышникову за внимательное руководство и неизменную поддержку в институте вычислительной математики.

Я глубоко признателен моим коллегам Ивану Валерьевичу Оселедцу и Дмитрию Валерьевичу Савостьянову за очень плодотворное сотрудничество и обмен

идеями. Их высокий профессионализм и поддержка были неоценимы в процессе исследования многомерных областей математики.

Я особо благодарен гостеприимству математического института им. Макса Планка в Лейпциге, и лично профессору Борису Николаевичу Хоромскому за вдохновляющую атмосферу и уникальные условия, великолепно способствующие научной деятельности.

Я высоко ценю замечательное и дружелюбное сотрудничество, установившееся с нашими коллегами в различных группах, занимающихся как прикладными задачами, так и разработками вычислительных методов, в особенности профессору Илье Купрову, University of Southampton, профессору Александру Павловичу Смирнову, факультет вычислительной математики и кибернетики МГУ, профессору Геннадию Алексеевичу Бочарову, институт вычислительной математики РАН, профессорам Томасу Шульте-Хербрюггену и Томасу Хукле, Technical University Munich, профессорам Удо Райхлю и Питеру Беннеру, институт сложных динамических систем им. Макса Планка, Магдебург. Объективная оценка текущей работы и дальнейших направлений исследований были бы невозможны без мотивирующих дискуссий о последних математических инструментах и интригующих практических приложениях.

Содержание работы

Во **введении** излагается актуальность темы работы, история методов тензорных аппроксимаций, цели и значимость данной диссертации.

В **первой главе** мы описываем три основных приложения: уравнение Власова для Фарлей-Бунемановской неустойчивости, основное кинетическое уравнение для стохастической химической кинетики, и схема одновременной пространственно-временной дискретизации динамических уравнений.

Фарлей-Бунемановская (ФБ) неустойчивость возникает в слабо ионизированной плазме E-области ионосферы Земли. Эта неустойчивость порождается в плазме с замагниченными электронами и незамагниченными ионами в электрическом поле, направленном перпендикулярно геомагнитному полю. В E-области электроны подвержены заметному влиянию геомагнитного поля, в отличие от ионов, которые не замагничиваются из-за частых столкновений с нейтральными частицами газа. В результате, посредством скорости дрейфа, обусловленной электрическим полем, распределение электронов по скоростям сдвигается относительно ионного распределения. Соответствующие условия для возникновения неустойчивости возникают в экваториальных и полярных зонах E-области ионосферы Земли на высотах порядка 90–130 км, где нестабильность проявляется как низкочастотные колебания плазмы с длинами волн порядка метра.

Плазма состоит из электронов и положительно заряженных ионов. Условия в E-слое ионосферы (величины фонового электрического поля и магнитного поля земли) позволяют использовать так называемое жидкостное уравнение на плотность концентрации электронов, представляющее собой обычное двух- или трехмерное уравнение диффузии-адвекции. Мы рассматриваем конкретные параметры, используемые в предыдущих работах, которые позволяют ограничиться двумя пространственными координатами (в плоскости перпендикулярной магнитному

полю). Но поведение ионов требует кинетического описания, которое дается четырехмерным уравнением Власова (две пространственные, две скоростные координаты):

$$\frac{\partial \psi}{\partial t} + v \frac{\partial \psi}{\partial x} + w \frac{\partial \psi}{\partial y} - \frac{\partial \phi}{\partial x} \frac{\partial \psi}{\partial v} + \left(\frac{e E_0}{m_i v_{T_i} \nu_{in}} - \frac{\partial \phi}{\partial y} \right) \frac{\partial \psi}{\partial w} = \psi_0 - \psi, \quad (1)$$

где $\psi = \psi(x, y, v, w, t)$ является функцией распределения ионов по скоростям v, w в каждой пространственной точке x, y во время t , ϕ обозначает электростатический потенциал, возникающий из-за неравновесности концентраций ионов и электронов, $\Delta \phi = \frac{e^2}{\varepsilon_0 m_i \nu_{in}^2} (n_e - n_i)$, E_0 это внешнее электрическое поле, e – заряд электрона, ε_0 – диэлектрическая постоянная, m_i – масса иона, ν_{in} – частота столкновений ионов с нейтральными частицами, $v_{T_i} = \sqrt{T_i/m_i}$ – характерная тепловая скорость ионов с температурой T_i . В правой части уравнения (1) стоит упрощенный оператор столкновений Bhatnagar-Gross-Krook, где

$$\psi_0 = \frac{n_i}{2\pi} \exp\left(-\frac{v^2 + w^2}{2}\right), \quad n_i = \int_{\mathbb{R}^2} \psi(x, y, v, w) dv dw \quad (2)$$

определяет правило вычисления нейтрального распределения: оно соответствует распределению плотности ионов в пространстве, но предполагается максвелловским по скоростям.

Решение уравнения (1) является четырехмерной функцией, и требует очень больших объемов памяти при непосредственном хранении. Однако, динамика распределения не сильно далеко уходит от максвелловского. Это дает возможность разделения переменных с помощью ТТ формата

$$\psi_{i,j,k,m} = \psi^{(1)}(i, j) \psi^{(3)}(k) \psi^{(4)}(m) = \sum_{\alpha_1, \alpha_3} \psi_{\alpha_1}^{(1)}(i, j) \psi_{\alpha_1, \alpha_3}^{(3)}(k) \psi_{\alpha_3}^{(4)}(m) \quad (3)$$

с разумными ТТ рангами, и сокращения затрат памяти в десятки раз.

Для дискретизации по пространственным переменным x, y , а также по скоростям v, w в уравнении (1) мы используем разностные схемы, принимая во внимание периодические граничные условия. Хотя областью определения для скоростей является вся плоскость, на практике функция распределения ψ в (1), будучи возмущенным распределением Максвелла, быстро убывает с ростом v и w , и мы можем ограничить область до квадрата $(v, w) \in [-v_{max}, v_{max}]^2$, накладывая периодические граничные условия.

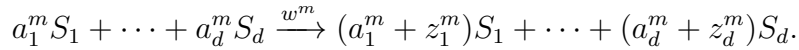
Для решения нелинейных уравнений модели, мы используем расщепления по физическим процессам и координатам:

1. диффузия электронной плотности;
2. конвекция электронной плотности;
3. Уравнение Пуассона для электростатического потенциала.
4. конвекция ионного распределения;
5. релаксация ионного распределения за счет столкновений.

Для получения второго порядка аппроксимации по времени, мы используем двуциклическую схему Марчука-Стрэнга для *квазилинейных* задач. Конвекция ионного распределения выполняется с помощью явной схемы по времени, построенной на основе приближенного метода характеристик, а оператор релаксации за счет вида (2) приближенно является ортогональным проектором, что позволяет написать упрощенное выражение для матричной экспоненты. Таким образом, оба шага интегрирования ионного уравнения требуют приближенного умножения многоуровневой матрицы на вектор в тензорном формате, для чего в главе 4 предложен специальный быстрый алгоритм.

Основное кинетическое уравнение является достаточно точным стохастическим описанием системы химических или биологических реакций, в которой количество молекул веществ находится в пределах сотен, и стохастический шум вносит уже значительный вклад, который не может быть должным образом учтен в детерминистских ОДУ. При таких малых концентрациях, стохастические колебания числа молекул могут достигать уровня $\mathcal{O}(10^{-1})$, так как столкновения между частицами и соответствующие реакции возникают принципиально случайным образом. Более того, сам по себе такой биологический шум играет важную роль в межклеточных и внутриклеточных функциях. Такие системы могут демонстрировать неожиданные поведения, например, наличие нескольких метастабильных состояний.

Динамическую эволюцию стохастической системы реакций можно рассматривать как дискретный марковский процесс. Предположим, что d различных активных химических соединений S_1, \dots, S_d в хорошо перемешиваемой среде могут реагировать посредством M реакционных каналов. Каждый канал задается *стехиометрическим вектором* $\mathbf{z}^m \in \mathbb{Z}^d$, и функцией скорости $w^m(\mathbf{i}) : \mathbb{R}_+^d \rightarrow \mathbb{R}_+$, $m = 1, \dots, M$, $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$, так что m -я реакция может быть записана в классическом виде



Чтобы ввести стохастическое описание, обозначим состояния мультииндексом $\mathbf{i} = (i_1, \dots, i_d)$, и всегда будем иметь в виду числа молекул, так что i_k является неотрицательным целым числом, $i_k \in (\{0\} \cup \mathbb{N})$. Вероятностная роль функции скорости следующая: для бесконечно малого интервала времени dt , $W^m(\mathbf{i}, t, dt) = w^m(\mathbf{i}) dt$ является вероятностью того, что при числе молекул \mathbf{i} в момент времени t , в последующем интервале $[t, t + dt)$ в системе произойдет одна реакция по каналу m . Количественной характеристикой состояния \mathbf{i} является вероятность того, что количества молекул веществ S_1, \dots, S_d в момент времени t принимают значения i_1, \dots, i_d , $\Psi(\mathbf{i}, t) : (\{0\} \cup \mathbb{N})^d \cup [0, T] \rightarrow \mathbb{R}_+$. Теперь, принимая dt достаточно малым, так что вероятность того, что больше, чем одна реакция будет происходить в интервале $[t, t + dt)$ пренебрежимо мала, можно написать распределение в конце интервала, $\Psi(\mathbf{i}, t + dt)$, используя законы сложения и умножения вероятностей для независимых и взаимоисключающих событий. Переходя к пределу $dt \rightarrow 0$, получаем основное кинетическое уравнение:

$$\frac{d\Psi(\mathbf{i}, t)}{dt} = \sum_{m=1}^M w^m(\mathbf{i} - \mathbf{z}^m) \Psi(\mathbf{i} - \mathbf{z}^m, t) - w^m(\mathbf{i}) \Psi(\mathbf{i}, t). \quad (4)$$

Потенциально возможны любые числа молекул, т.е. уравнение (4) является бесконечномерным ОДУ. Разумеется, для проведения численного моделирования нам нужно ограничить его до конечной задачи. Так называемый алгоритм *проекции в конечное пространство* (Finite State Projection, FSP) использует тот факт, что очень большие количества молекул имеют очень малую вероятность возникнуть за конечное время: $\Psi(\mathbf{i}, t) \rightarrow 0$, $|\mathbf{i}| \rightarrow \infty$. Таким образом, мы полагаем что каждый i_k лежит в конечном диапазоне, $i_k = 0, \dots, n_k - 1$, и выбираем n_k достаточно большими, так что, например, при $i_k > n_k$ вероятность $\Psi(\mathbf{i}, t)$ меньше машинной точности, и можно пренебречь ошибкой, внесенной при ограничении пространства состояний. Однако, даже если каждый $n_k = \mathcal{O}(n)$ не превышает десятков, общее число степеней свободы растет как n^d , и сжатое хранение и обработка для Ψ принципиально необходимы.

Схемы интегрирования по времени, например классические методы Кранка-Николсон и Эйлера для линейных ОДУ, могут быть переписаны в виде блочных систем линейных уравнений. Пусть в качестве базисной используется схема Кранка-Николсон,

$$\left(I + \frac{\delta t}{2}A\right)x_{p+1} = \left(I - \frac{\delta t}{2}A\right)x_p + \frac{\delta t}{2}(f_p + f_{p+1}), \quad p = 0, \dots, N_t - 1, \quad (5)$$

где $x_p \approx x(t_p)$, $f_p = f(t_p)$, $t_p = p\delta t$ задает сетку по времени, тогда мы можем собрать все временные шаги в одну глобальную линейную систему:

$$\begin{bmatrix} I + \frac{\delta t}{2}A & & & & \\ -I + \frac{\delta t}{2}A & I + \frac{\delta t}{2}A & & & \\ & \ddots & \ddots & & \\ & & & -I + \frac{\delta t}{2}A & I + \frac{\delta t}{2}A \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N_t} \end{bmatrix} = \begin{bmatrix} x_0 - \frac{\delta t}{2}Ax_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \frac{\delta t}{2}g, \quad (6)$$

где $g = [f_0 + f_1 \quad f_1 + f_2 \quad \dots \quad f_{N_t-1} + f_{N_t}]^\top$. Конечно, эта система не несет смысла, если мы будем держать ее в стандартном виде – пошаговое решение (5) является оптимальным методом для двухдиагональной матрицы. Тем не менее, предположим что можно вычислить ее решение более эффективно в *структурированном* представлении. Тогда становится заманчиво получить сразу всю историю динамики, и использовать очень мелкие шаги по времени, чтобы гарантировать достаточную точность для всех спектральных компонент матрицы A . С использованием приближений тензорными произведениями, мы можем ввести дополнительные виртуальные переменные, и в конечном итоге аппроксимировать и пространственную, и временную части системы (6) с логарифмической редукцией вычислительной сложности, $\mathcal{O}(\log(N) \log(N_t))$, которая действительно подтверждается в численных экспериментах.

Глава 2 посвящена описанию различных методов разделения переменных. Дается обзор существующих тензорных разложений: канонического (CP), формата Таккера, и рекуррентных представлений иерархический Таккер и MPS/TT, а также их свойств и связанных с ними алгоритмов. *Рекуррентные* тензорные представления являются эффективным инструментом сжатия данных. На первом шаге построения такого формата мы используем *группировку индексов*, чтобы уменьшить число переменных формально (но получить большее число степеней свободы

в новой, сдвоенной переменной), а затем находим для них более узкие подпространства с помощью двумерных матричных малоранговых приближений (например, сингулярного разложения), что и дает фактическое сокращение числа неизвестных.

В особенности, мы сфокусируемся на MPS/ТТ представлении, в котором размерности группируются в соответствии с *линейным* деревом. Практический опыт позволяет полагать, что наиболее важным является *порядок* индексов (а он может быть определен посредством физических соображений, или с использованием адаптивных алгоритмов), тогда как для группировки размерностей в большинстве случаев достаточно линейного дерева. Его простота обеспечивает определенную элегантность обозначений, что особенно важно в многомерных задачах из-за обилия индексов.

Говорят, что тензор x представлен (или аппроксимирован) в формате *Matrix Product States* (MPS), или *Tensor Train* (ТТ), если выполняется

$$x(i_1, \dots, i_d) \approx \sum_{\alpha_1, \dots, \alpha_{d-1}} x_{\alpha_1}^{(1)}(i_1) x_{\alpha_1, \alpha_2}^{(2)}(i_2) \cdots x_{\alpha_{d-2}, \alpha_{d-1}}^{(d-1)}(i_{d-1}) x_{\alpha_{d-1}}^{(d)}(i_d), \quad (7)$$

где $x^{(k)} \in \mathbb{C}^{r_{k-1} \times n_k \times r_k}$ называются ТТ *ядрами* (*блоками*, или *факторами*), диапазоны $r_k = r_k(x) \leq r(x)$ *ранговых индексов* $\alpha_k = 1, \dots, r_k$ называются *ТТ рангами*. Для однообразия записи можно полагать $r_0 = r_d = 1$. Любопытно, что это представление было независимо открыто несколько раз: как MPS в квантовой физике с 1980'х, и как ТТ в 2009 в вычислительной линейной алгебре. Асимптотические затраты памяти для ТТ формата составляют $\mathcal{O}(dnr^2)$. Для небольших диапазонов исходных индексов i_k (т.н. *модовых размеров* n) и более высоких рангов, ТТ формат содержит меньше неизвестных, чем более сложные древовидные разложения (последние имеют асимптотическую сложность не ниже $\mathcal{O}(r^3)$), что делает его особенно эффективным для описания, например, спиновых систем.

ТТ формат для матриц, также известный как *Matrix Product Operator*, вводится с использованием перестановки индексов,

$$A(i_1, \dots, i_d, j_1, \dots, j_d) = \sum_{\gamma_1, \dots, \gamma_{d-1}} A_{\gamma_1}^{(1)}(i_1, j_1) A_{\gamma_1, \gamma_2}^{(2)}(i_2, j_2) \cdots A_{\gamma_{d-1}}^{(d)}(i_d, j_d). \quad (8)$$

С использованием дистрибутивности, произведение матрицы на вектор пишется независимо для каждого ТТ блока: результат получается в аналогичном ТТ формате $y(\mathbf{i}) = y^{(1)}(i_1) \cdots y^{(d)}(i_d)$, где

$$y_{\beta_{k-1}, \beta_k}^{(k)} = A_{\gamma_{k-1}, \gamma_k}^{(k)} x_{\alpha_{k-1}, \alpha_k}^{(k)}, \quad \text{т.е.} \quad y_{\beta_{k-1}, \beta_k}^{(k)}(i_k) = \sum_{j_k} A_{\gamma_{k-1}, \gamma_k}^{(k)}(i_k, j_k) x_{\alpha_{k-1}, \alpha_k}^{(k)}(j_k), \quad (9)$$

причем $\beta_k = \overline{\alpha_k}, \gamma_k = 1, \dots, r_k(A)r_k(x)$, $k = 0, \dots, d$.

Для единого описания форматов (7) и (8), удобно использовать следующее **Определение**. Пусть даны ТТ форматы $\{x^{(k)}(i_k)\}$ или $\{A^{(k)}(i_k, j_k)\}$. Векторное *ТТ отображение* раскрывает ТТ представление в полный тензор следующим образом:

$$\tau(x^{(p)}, \dots, x^{(q)}) : \{x^{(k)}\}_{k=p}^q \rightarrow x^{(p, \dots, q)} \in \mathbb{C}^{r_{p-1} \times n_p \cdots n_q \times r_q}, \quad \text{где}$$

$$x_{\alpha_{p-1}, \alpha_q}^{(p, \dots, q)}(\overline{i_p, \dots, i_q}) = \sum_{\alpha_p, \dots, \alpha_{q-1}} x_{\alpha_{p-1}, \alpha_p}^{(p)}(i_p) \cdots x_{\alpha_{q-1}, \alpha_q}^{(q)}(i_q), \quad (10)$$

при $1 \leq p \leq q \leq d$. Для граничных случаев мы можем дополнительно обозначать

$$x^{(1, \dots, q)} = x^{(\leq q)} = x^{(< q+1)} \in \mathbb{C}^{n_1 \dots n_q \times r_q}, \quad x^{(p, \dots, d)} = x^{(\geq p)} = x^{(> p-1)} \in \mathbb{C}^{r_{p-1} \times n_p \dots n_d}.$$

Матричное ТТ отображение $\tau(A^{(p)}, \dots, A^{(q)})$ аналогично соответствует (8).

Другие полилинейные алгебраические операции, например, $z = x + y$, могут быть также перенесены с элементов исходных тензоров на элементы их тензорных форматов,

$$z^{(1)}(i_1) = [x^{(1)}(i_1) \quad y^{(1)}(i_1)], \quad z^{(d)}(i_d) = \begin{bmatrix} x^{(d)}(i_d) \\ y^{(d)}(i_d) \end{bmatrix}, \quad z^{(k)}(i_k) = \begin{bmatrix} x^{(k)}(i_k) & \\ & y^{(k)}(i_k) \end{bmatrix}$$

для $k = 2, \dots, d-1$. Все операции обладают линейной сложностью по размерности d , и полиномиальными сложностями по n и r .

Главным преимуществом тензорных деревьев по сравнению с простой суммой R прямых произведений является процедура *рекомпрессии* или *округления*. Алгебраические операции могут увеличивать ТТ ранги, при этом они могут получаться завышенными для данных тензоров и точностей. ТТ округление позволяет понижать ранги до квазиоптимальных значений при заданной точности. Эта процедура использует известные QR и сингулярное разложения матриц, и может быть выполнена за $\mathcal{O}(dnr^3)$ операций.

Отдельно стоит отметить понятие *тензоризации*, или *квантизации*, предложенной Оселедцем и Хоромским. Рассмотрим для краткости одномерный вектор, $x = [x(i)]$. Пусть число допустимых значений для i составляет степень 2, т.е. $n = 2^L$. Рассмотрим позиционную запись i в двоичной системе исчисления, т.е.

$$i = \sum_{l=1}^L i_l \cdot 2^{l-1}, \quad i_l \in \{0, 1\}. \quad (11)$$

Это соответствует перегруппировке вектора в тензор с L виртуальными (*квантизованными*) размерностями. Теперь можно применить ТТ разложение:

$$x(i) = x(i_1, \dots, i_L) \approx x^{(1)}(i_1) \cdots x^{(L)}(i_L).$$

Если ТТ ранги этого тензора малы, требуемый объем памяти составляет примерно *логарифм* от исходного, $\mathcal{O}(L \cdot 2 \cdot r^2) = \mathcal{O}(\log n)$. Это позволяет добиться большей эффективности сжатия данных, в частности, в “маломерных”, двух- и трехмерных задачах. Такая разновидность ТТ формата была названа *Quantized Tensor Train* (QTT) формат. Название Quantized связано с термином “квант”, минимально возможный долей информации, получаемой при задании каждой цифры i_l .

При работе с малопараметрическими представлениями данных, неизбежно задается вопрос о конкретном значении ТТ (или QTT) рангов, r . В конце главы предлагается новое комбинированное тензорное представление (QTT-Tucker), объединяющее аналитические и практические преимущества Таккеровского, ТТ и QTT форматов. В основе нового формата лежит разложение Таккера:

$$x(i_1, \dots, i_d) = \sum_{\gamma_1, \dots, \gamma_d} x^{(c)}(\gamma_1, \dots, \gamma_d) x_{\gamma_1}^{(1)}(i_1) \cdots x_{\gamma_d}^{(d)}(i_d). \quad (12)$$

Проклятие размерности снимается за счет хранения Таккеровского ядра в ТТ формате:

$$x^{(c)}(\gamma_1, \dots, \gamma_d) = \sum_{\alpha_1, \dots, \alpha_{d-1}} x_{\alpha_1}^{c(1)}(\gamma_1) x_{\alpha_1, \alpha_2}^{c(2)}(\gamma_2) \cdots x_{\alpha_{d-1}}^{c(d)}(\gamma_d). \quad (13)$$

Наконец, если Таккеровские факторы $x_{\gamma_k}^{(k)}(i_k)$ из-за больших размеров n_k занимают слишком много памяти, для индекса i_k можно ввести тензоризацию:

$$x_{\gamma_k}^{(k)}(i_k) = \sum_{\gamma_{k,L}, \dots, \gamma_{k,1}} x_{\gamma_{k,L-1}}^{f(k,L)}(i_{k,L}) x_{\gamma_{k,L-1}, \gamma_{k,L-2}}^{f(k,L-1)}(i_{k,L-1}) \cdots x_{\gamma_{k,1}, \gamma_k}^{f(k,1)}(i_{k,1}). \quad (14)$$

На практике, этот формат может приносить дополнительное сокращение сложности, поскольку QTT ранги в Таккеровских факторах требуют меньших значений для поддержания той же точности аппроксимации, чем соответствующие QTT ранги в глобальном линейном формате. Еще одной интересной особенностью нового формата является то, что он может наследовать оценки рангов из формата Таккера, которые могут быть гораздо легче доказуемы, чем оценки в ТТ формате, за счет использования полиномиальной интерполяции. Вычислительные алгоритмы для формата QTT-Tucker удобно описывать с точки зрения двухуровневого ТТ представления. Это позволяет использовать ТТ версии многих методов (например, переменных направлений).

В главе 3 мы объединяем прикладные задачи, изложенные в первой главе, и методы тензорных представлений из второй главы. Мы выведем несколько явных разложений низкого ТТ/QTT ранга для матриц и векторов (тензоров), имеющих отношение к дифференциальным и разностным операторам и некоторым типичным функциям в рассматриваемых приложениях, поскольку в процессе эксплуатации различных тензорных форматов полезно уметь строить простые и широко используемые тензоры явно, задавая непосредственно элементы формата.

В качестве наиболее примечательных примеров, следует отметить следующие.

Матрицы конечных разностей в QTT формате: пусть J_n есть Жорданов блок размером $n \times n$, $G_n = I_n - J_n^\top$ есть матрица жесткости в схеме Кранка-Николсон (6), и $M_n = I_n + J_n^\top$ есть матрица масс. Если $n = 2^L$, то можно записать следующее QTT представление с рангами 2:

$$\begin{bmatrix} G_n(\mathbf{i}, \mathbf{j}) \\ M_n(\mathbf{i}, \mathbf{j}) \end{bmatrix} = \begin{bmatrix} G_{i_1, j_1} & J_{i_1, j_1} \end{bmatrix} \cdots \begin{bmatrix} I_{i_k, j_k} & J_{i_k, j_k} \\ J_{i_k, j_k}^\top & J_{i_k, j_k} \end{bmatrix} \cdots \begin{bmatrix} I_{i_L, j_L} \\ J_{i_L, j_L}^\top \end{bmatrix},$$

где $G = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ и $M = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ суть элементарные (2×2) матрицы конечных разностей и масс.

Матрицы перехода для ионного уравнения Фарлей-Бунемановской неустойчивости. Например, одномерная матрица интегрирования на один шаг по времени конвекционной части уравнения Власова (1) пишется в виде

$$M_c = \alpha_{-2} S_{-2} + \alpha_{-1} S_{-1} + \alpha_0 S_0 + \alpha_1 S_1 + \alpha_2 S_2,$$

где S_{-2}, \dots, S_2 суть матрицы периодического сдвига, а коэффициенты $\alpha_{-2}, \dots, \alpha_2$ определяют веса интерполяции с характеристики. Однако, в нашем случае каждый вес α_i параметризован значением скорости v . То есть, действие α_i задается

диагональной матрицей

$$\Lambda_i = [I \otimes I] \otimes \text{diag}(\alpha_i(v)) \otimes I,$$

и в итоге для матрицы перехода в исходном ионном уравнение получаем ТТ представление ранга 5:

$$\begin{aligned} M_x(\delta t) &= [S_{-2} \otimes I] \otimes \text{diag}(\alpha_{-2}(v)) \otimes I + [S_{-1} \otimes I] \otimes \text{diag}(\alpha_{-1}(v)) \otimes I \\ &+ [S_0 \otimes I] \otimes \text{diag}(\alpha_0(v)) \otimes I \\ &+ [S_1 \otimes I] \otimes \text{diag}(\alpha_1(v)) \otimes I + [S_2 \otimes I] \otimes \text{diag}(\alpha_2(v)) \otimes I, \end{aligned}$$

где Кронекеровские произведения по x и y (в квадратных скобках) вычисляются явно (мы не разделяем пространственные переменные), тогда как остальные подразумеваются неявно, и сомножители хранятся отдельно, как ТТ блоки.

Третьим интересным примером аналитических ТТ построений является матрица основного кинетического уравнения для *каскадной цепи реакций*. Она возникает при моделировании систем, в которых каждый из генов воздействует на экспрессию следующего гена, что выражается в определенном виде скоростей реакций. При этом, матрица основного кинетического уравнения является суммой попарных произведений элементарных матриц, и имеет эффективное ТТ представление, как показано в следующей лемме.

Лемма. Пусть даны элементарные векторы или матрицы $E_k(i_k)$, $F_k^k(i_k)$, $F_k^{k+1}(i_k)$ для $k = 1, \dots, d$. Сумма попарных произведений

$$H = F_1^1 \otimes \left(\bigotimes_{j=2}^d E_j \right) + \sum_{k=2}^d \left(\bigotimes_{j=1}^{k-2} E_j \right) \otimes F_{k-1}^k \otimes F_k^k \otimes \left(\bigotimes_{j=k+1}^d E_j \right) \quad (15)$$

обладает точным ТТ разложением $H(i) = H^{(1)}(i_1) \dots H^{(d)}(i_d)$ ранга 3, вида

$$\begin{aligned} H^{(1)}(i_1) &= [E_1(i_1) \quad F_1^2(i_1) \quad F_1^1(i_1)], \quad H^{(d)}(i_d) = \begin{bmatrix} 0 \\ F_d^d(i_d) \\ E_d(i_d) \end{bmatrix}, \\ H^{(k)}(i_k) &= \begin{bmatrix} E_k(i_k) & F_k^{k+1}(i_k) & 0 \\ 0 & 0 & F_k^k(i_k) \\ 0 & 0 & E_k(i_k) \end{bmatrix}, \quad \text{если } k = 2, \dots, d-1. \end{aligned} \quad (16)$$

Разложение Таккера (12) также выполняется с рангами не более 3.

В **главе 4** изложены методы решения линейных уравнений в тензорных форматах: классические итерационные методы с тензорными округлениями, итерационные методы оптимизации переменных направлений, разработанные в квантовой физике (DMRG) и вычислительной математике (метод наименьших квадратов в переменных направлениях), в также новый улучшенный алгоритм (AMEn), которая дополняет схему переменных направлений классическим градиентным шагом с использованием глобальной невязки системы. Мы анализируем его сходимость, доказываем геометрическую скорость сходимости, и обсуждаем некоторые практические аспекты.

Исходная алгебраическая задача, такая как аппроксимация или решение линейной системы, может часто рассматриваться как оптимизация некоторой квадратичной целевой функции. Наиболее важными являются:

1. ошибка $E_{x_*}(x) = \|x_* - x\|^2 = J_{I, x_*}$,
2. энергия $J_{A,b}(x) = \|x_* - x\|_A^2 = (x, Ax) - 2\text{Re}(x, b) + \text{const}$, где $b = Ax_*$, и
3. невязка $R_{A,b}(x) = \|Ax - b\|^2 = J_{A^*A, A^*b}(x) + \text{const}$.

Эти три функции связаны с решением некоторой (может быть, с единичной матрицей) линейной системы. Функция энергии определяется для симметричной положительно определенной (SPD) матрицы, $A = A^* > 0$, с A -скалярным произведением, и A -нормой, введенными обычным образом: $(x, y)_A = (x, Ay)$, $\|x\|_A = \sqrt{(x, x)_A}$.

Одним из наиболее простых тензорных оптимизационных методов является метод наименьших квадратов в переменных направлениях (ALS), так же названный *линейная схема* в переменных направлениях, и совпадающий с (одноблочным) *Density Matrix Renormalization Group* (DMRG) подходом из квантовой физики. Оригинальная DMRG схема была предложена для вычисления основного состояния системы путем минимизации отношения Рэля. Позднее она была применена для решения SPD линейной системы $Ax = b$, где A и b заданы в ТТ формате, путем минимизации функции энергии.

Итак, мы ограничиваем оптимизацию $J_{A,b}(x)$ на векторы $x = \tau(x^{(1)}, \dots, x^{(d)})$, представленные в ТТ формате с *фиксированными* ТТ рангами $\mathbf{r} = (r_1, \dots, r_{d-1})$, и выполняем фактически вычисления посредством последовательности *микрошагов*, т.е. *последовательных* оптимизаций по ТТ блокам $x^{(k)}$. Каждая такая *локальная задача* ставится следующим образом:

$$u^{(k)} = \arg \min_{x^{(k)}} J_{A,b}(\tau(x^{(1)}, \dots, x^{(d)})) \quad \text{по } x^{(k)} \in \mathbb{C}^{r_{k-1} \times n_k \times r_k}. \quad (17)$$

ТТ блок $x^{(k)}$ затем заменяется на $u^{(k)}$, и процесс переходит к следующему ядру; обычно блоки пересчитываются в последовательном проходе по размерностям, например $k = 1, \dots, d$ (*прямой полу-проход*), или $k = d, \dots, 1$ (*обратный полу-проход*), и так далее вплоть до сходимости x .

Используя определение ТТ отображения (10) для частичного ТТ представления, мы можем определить *фрейм-матрицу*:

$$X_{\neq k} = x^{(<k)} \otimes I_{n_k} \otimes (x^{(>k)})^\top \in \mathbb{C}^{n_1 \dots n_d \times r_{k-1} n_k r_k}. \quad (18)$$

Она задает соотношение *линейности* ТТ формата, $x = X_{\neq k} x^{(k)}$. Это позволяет переписать (17) в виде решения линейной системы

$$\begin{aligned} u^{(k)} &= A_k^{-1} b_k, \\ A_k &= X_{\neq k}^* A X_{\neq k} \in \mathbb{C}^{(r_{k-1} n_k r_k) \times (r_{k-1} n_k r_k)}, \quad b_k = X_{\neq k}^* b \in \mathbb{C}^{r_{k-1} n_k r_k}. \end{aligned} \quad (19)$$

Недостатком одноблочного алгоритма DMRG является то, что ТТ ранги остаются постоянными в процессе расчета. Мы должны угадывать правильные значения рангов решения *априори*, что может быть трудно; Кроме того, сходимость даже к квази-оптимальному решению при данных ТТ рангах может быть очень медленной. Для адаптации ТТ рангов в процессе расчета был разработан *двухблочный* DMRG. Данный метод работает с вектором в следующем виде:

$$x = \tau(x^{(1)}, \dots, x^{(k-1)}, x^{(k,k+1)}, x^{(k+2)}, \dots, x^{(d)}), \quad (20)$$

где $x^{(k,k+1)} = x^{(k,\dots,k+1)} = \tau(x^{(k)}, x^{(k+1)})$ пишется в соответствии с определением (10). Шаг локальной оптимизации проводится аналогично (17), но по элементам *суперядра* $x^{(k,k+1)}$ следующим образом:

$$u^{(k,k+1)} = \arg \min_{x^{(k,k+1)}} J_{A_{k,k+1}, b_{k,k+1}}(x^{(k,k+1)}) \quad \text{по } x^{(k,k+1)} \in \mathbb{C}^{r_{k-1}n_k n_{k+1} r_{k+1}}. \quad (21)$$

Решение системы (21) $u^{(k,k+1)}$ затем разделяется обратно на $u^{(k)}$ и $u^{(k+1)}$ для восстановления первоначальной ТТ структуры с помощью матричного малорангового разложения. Так как элементы $u^{(k,k+1)}$ отличаются от элементов предыдущего приближения, ранг r_k может быть изменен. Тем не менее, из-за *локального* характера оптимизации, методы DMRG (даже двухблочные) могут терять важную часть информации о направлении к точному решению, и стагнировать в каком-то локальном минимуме, далеко от желаемого порога.

Одноблочный DMRG, аналогично методу *градиентного спуска*, является *вариационным* для функции энергии: на каждом шаге, эти методы ищут минимизатор функции энергии с ограничениями. В этом смысле DMRG метод аналогичен классическим проекционным методам, тому же алгоритму градиентного спуска или GMRES. Разница состоит в том, что фрейм-матрицы текущего решения не приближают Крыловские базисы, и, следовательно, не гарантируют сходимости метода. В данной диссертации, мы предлагаем новый метод (*AMEn*), который сочетает классические алгоритмы и схемы переменных направлений, путем обогащения фрейм-матриц информацией Крыловского типа, а точнее, фрейм-матрицей *невязки*. По аналогии с алгоритмом *минимальных невязок* (Minimal Residual, MR), мы можем назвать метод градиентного спуска также методом *минимальных энергий* (Minimal Energy, MEn). Новый метод решает линейную систему в ТТ формате методом DMRG, дополненным шагом вида градиентного спуска. Это и послужило причиной названия AMEn: *Alternating Minimal Energy*.

Для анализа алгоритмов переменных направлений (а именно, мы рассматриваем AMEn) удобно переписать их в следующем рекуррентном виде.

1. Построить $A_1 = X_{\neq 1}^* A X_{\neq 1}$, $b_1 = X_{\neq 1}^* b$.
2. Решить $A_1 u^{(1)} = b_1$, принять $u = \tau(u^{(1)}, x^{(2)}, \dots, x^{(d)})$.
3. **Объединить** $x^{(1)} := [u^{(1)} \quad z^{(1)}]$, $x^{(2)} := \begin{bmatrix} x^{(2)} \\ 0 \end{bmatrix}$, где $z^{(1)}$ является ТТ блоком *приближенной невязки*, $\tilde{z} = \tau(z^{(1)}, \dots, z^{(d)}) \approx b - Au$.
4. Перезапустить алгоритм для приближенного решения $A_{\geq 2} x^{(\geq 2)} = b_{\geq 2}$ (рекурсия). *Редуцированная* матрица $A_{\geq 2}$ и правая часть $b_{\geq 2}$ собираются так:

$$\begin{aligned} A_{\geq 2} &= X_{\neq \{2, \dots, d\}}^* A X_{\neq \{2, \dots, d\}}, \quad b_{\geq 2} = X_{\neq \{2, \dots, d\}}^* b, \quad \text{где} \\ X_{\neq \{2, \dots, d\}} &\equiv X_1 = x^{(1)} \otimes I_{n_2 \dots n_d} \in \mathbb{C}^{(n_1 \dots n_d) \times (r_1 n_2 \dots n_d)}. \end{aligned} \quad (22)$$

Третий шаг принципиально отличает алгоритм AMEn и алгоритмы DMRG. Действительно, редуцированная система на шаге 4 является системой Галеркина в базисе столбцов матрицы $X_{\neq \{2, \dots, d\}}$: полное решение удовлетворяет

$$x = X_1 x^{(\geq 2)}, \quad \text{где } x^{(\geq 2)} \approx (X_1^* A X_1)^{-1} (X_1^* b),$$

т.е. приближенно выполняет условия Галеркина, $X_1^*(b - Ax) \approx 0$ (приближенность возникает из-за того, что система $A_{\geq 2}x^{(\geq 2)} = b_{\geq 2}$ в свою очередь решается АМЕп методом, т.е. неточно).

Однако, заметим, что $X_1 = [u^{(1)} \quad z^{(1)}] \otimes I_{n_2 \dots n_d} = [U_1 \quad Z_1]$, где $Z_1 = z^{(1)} \otimes I_{n_2 \dots n_d}$ является фрейм-матрицей приближенной невязки, $\tilde{z} = Z_1 z^{(1)}$. Следовательно, Галеркинский базис содержит вектор неточной невязки, $\tilde{z} \in \text{span}(X_1)$. Это дает возможность доказать *глобальную сходимость* метода АМЕп на основе теории сходимости неточного градиентного спуска.

Утверждение. Пусть дана SPD линейная система $Ax = b$ и начальное приближение t , рассмотрим невязку $z = b - At$ и вектор \tilde{z} , такой, что $\angle(z; \tilde{z}) \leq \theta < \pi/2$. Тогда неточный метод градиентного спуска

$$x = t + \tilde{z}h, \quad h = \arg \min_{h'} J_{A,b}(t + \tilde{z}h') = \frac{(\tilde{z}, z)}{(\tilde{z}, A\tilde{z})} \quad (23)$$

сходится, и скорость сходимости оценивается следующим образом ($x_* = A^{-1}b$):

$$\omega_{z, \tilde{z}} = \frac{\|x_* - x\|_A}{\|x_* - t\|_A} \leq \frac{\tilde{\kappa} - 1}{\tilde{\kappa} + 1} = \tilde{\Omega} < 1, \quad \tilde{\kappa} = \text{cond}(A) \frac{1 + \sin \theta}{1 - \sin \theta}. \quad (24)$$

Рекуррентный вид метода АМЕп позволяет собрать оценки сходимости, гарантируемые неточным градиентным спуском на каждом уровне, в следующий результат.

Теорема. АМЕп алгоритм является сходящимся, если ошибки аппроксимации невязок $\tilde{z}_k \approx z_k = b_{\geq k} - A_{\geq k}u^{(\geq k)}$, вносимые на всех шагах рекурсии, удовлетворяют условию $\theta = \max_{k=1, \dots, d-1} \angle(z_k; \tilde{z}_k) < \pi/2$. Скорость сходимости одной АМЕп итерации ограничена сверху так, что имеет место неравенство

$$\frac{\|x_* - x\|_A^2}{\|x_* - t\|_A^2} \leq 1 - (1 - \tilde{\Omega}^2)^{d-1}, \quad \tilde{\Omega} = \frac{\tilde{\kappa} - 1}{\tilde{\kappa} + 1}, \quad \tilde{\kappa} = \text{cond}(A) \frac{1 + \sin \theta}{1 - \sin \theta}, \quad (25)$$

где t является начальным приближением, x – приближением после АМЕп итерации, а $x_* = A^{-1}b$ – точное решение.

В конце четвертой главы разбираются практические аспекты эффективной реализации алгоритма АМЕп, и приводится сравнение его с другим типом адаптивной одноблочной схемы DMRG, разработанным в сообществе квантовой физики. Показаны основные преимущества АМЕп: отсутствие возмущения решения в процессе внесения блока невязки, как следствие отсутствие эвристических параметров для масштабирования поправочных данных, и в итоге возможность доказательства глобальной сходимости. Приведена ссылка на публикацию автора [3], в которой также подтверждается и более высокая эффективность АМЕп в практических расчетах основных состояний спиновых цепей.

В **главе 5** излагаются численные эксперименты. Рассмотрены несколько важных биологических систем, описываемых с помощью основного кинетического уравнения, а также гибридная многомерная модель плазменной неустойчивости, исследованы их особенности, связанные с концепцией тензорных произведений, проведено сравнение вычислительных алгоритмов. Данными численными экспериментами подтверждено, что предлагаемые тензорные форматы и методы применимы в качестве эффективных и высокоточных инструментов для обсуждаемых приложений.

Заключение: основные результаты работы

В данной диссертации, мы предложили и исследовали представления и алгоритмы для аппроксимаций функций и операторов, и решения многомерных линейных систем в форматах тензорных произведений. Особенно интересные многомерные модели возникают в виде нестационарных дифференциальных уравнений для описания стохастических динамических систем и химической кинетики. В течение долгого времени, прямой расчет функции плотности вероятности уравнениями Власова, Фоккера-Планка или основным кинетическим (управляющим) избегался в связи с проклятия размерности, за исключением случаев малого числа переменных. Форматы разреженного представления данных общего вида, косвенно хранящие все n^d элементов, открывают путь к быстрым и точным методам стохастического моделирования в машиностроении, биологии, химии и физики.

- Центральным результатом диссертации является новый вычислительный алгоритм (AMEn) для решения больших систем линейных уравнений и аппроксимации данных в форматах тензорных произведений. Этот метод был получен путем объединения DMRG схемы оптимизации в переменных направлениях, и классического метода градиентного спуска, и позволяет вычислять приближенные решения систем уравнений значительно быстрее и точнее, чем каждая из используемых схем в отдельности.

В частности, быстрая сходимость метода AMEn позволяет решать и системы с сильно несимметричными матрицами без изменения алгоритма, хотя его формулировка и анализ приводятся для симметричных положительно определенных матриц. Это находит особенно важное применение в решение нестационарных уравнений, где неявные схемы дискретизации по времени приводят к несимметричным задачам.

- Отдельным результатом диссертации является блочная версия стандартных схем Кранка-Николсон и Эйлера, которая, с использованием аппроксимаций тензорными произведениями, позволяет достичь логарифмической сложности в зависимости от числа временных шагов. Мы показали, что этот подход действительно эффективен для практических применений, и позволяет быстро вычислять эволюцию сложных систем с высокой точностью.

Мы убедились, что достижение прогресса в ускорении расчетов или повышении точности стало возможным только после усовершенствования вычислительных методов: ни DMRG алгоритмы переменных направлений, ни классические итерации сами по себе не способны справиться с предложенными задачами. Но в новых комбинированных схемах они демонстрируют замечательную кооперацию, позволяя вычислять решения эффективно и с высокой точностью даже в существенно многомерных примерах. Это достаточно редкий случай, когда инструмент для многомерных задач одновременно является эффективным на практике, и обладает теоретическим анализом глобальной сходимости.

- Численные эксперименты, представленные в диссертации, включают в себя примеры стохастических моделей в биологии и в физике плазмы, и показывают возможность существенного сокращения вычислительных затрат. В частности, применение предложенных методик к основному кинетическому уравнению дает

принципиально новую возможность решать *стохастическое* уравнение *детерминистским* алгоритмом.

Методика, предложенная автором, успешно применяется и в других задачах, не вошедших в рамки данной работы. Совсем недавно было успешно проведено квантовое моделирование спиновой динамики для ядерного магнитного резонанса – использование АМЕп метода, предложенного в данной работе, позволяет рассчитывать структуры сложных белков на одном процессоре рабочей станции при весьма общих начальных знаниях о системе. До сих пор, единственной применимой альтернативой был метод SSR, существенно использующий интуитивные химические соображения для построения редуцированной модели, и тем не менее требующий до терабайта общей памяти.

Публикации автора по теме диссертации

- [1] Computation of extreme eigenvalues in higher dimensions using block tensor train format / S. V. Dolgov, B. N. Khoromskij, I. V. Oseledets, D. V. Savostyanov // Computer Phys. Comm. — 2014. — Vol. 185, no. 4. — P. 1207–1216.
- [2] Dolgov S. V., Smirnov A. P., Tyrtysnikov E. E. Low-rank approximation in the numerical modeling of the Farley-Buneman instability in ionospheric plasma // J. Comp. Phys. — 2014. — Vol. 263. — P. 268–282.
- [3] Dolgov S. V., Savostyanov D. V. Corrected one-site density matrix renormalization group and alternating minimal energy algorithm // Proc. ENUMATH 2013, accepted. — 2014. — URL: <http://arxiv.org/abs/1312.6542>.
- [4] Dolgov S., Khoromskij B. Two-Level QTT-Tucker Format for Optimized Tensor Calculus // SIAM J. on Matrix An. Appl. — 2013. — Vol. 34, no. 2. — P. 593–623.
- [5] Dolgov S. V. TT-GMRES: solution to a linear system in the structured tensor format // Russ. J. Numer. Anal. Math. Modelling. — 2013. — Vol. 28, no. 2. — P. 149–172.
- [6] Dolgov S. V., Khoromskij Boris N., Oseledets Ivan V. Fast solution of multi-dimensional parabolic problems in the tensor train/quantized tensor train-format with initial application to the Fokker-Planck equation // SIAM J. Sci. Comput. — 2012. — Vol. 34, no. 6. — P. A3016–A3038.
- [7] Dolgov S. V., Oseledets I. V. Solution of linear systems and matrix inversion in the TT-format // SIAM J. Sci. Comput. — 2012. — Vol. 34, no. 5. — P. A2718–A2739.