



БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ
ИНСТИТУТ ПО ИНФОРМАЦИОННИ И
КОМУНИКАЦИОННИ ТЕХНОЛОГИИ

Венелин Любомиров Тодоров

**МЕТОДИ МОНТЕ КАРЛО
ЗА МНОГОМЕРНИ ИНТЕГРАЛИ И
ИНТЕГРАЛНИ УРАВНЕНИЯ
И ПРИЛОЖЕНИЯ**

АВТОРЕФЕРАТ

на дисертация за получаване на
образователна и научна степен „Доктор“
Докторска програма „Математическо моделиране и
приложение на математиката“
Професионално направление 4.5 „Математика“

Научен ръководител:

проф. д.т.н. Иван Димов

София, 2017

Дисертационният труд е обсъден и допуснат до защита на разширено заседание на секция „Паралелни алгоритми“ при ИИКТ-БАН, състояло се на 23.05.2017 г.

Дисертационният труд съдържа 184 страници, в които 26 фигури, 45 таблици и 18 страници литература, включваща 237 заглавия.

Защитата на дисертационния труд ще се състои на 2017 г. от 14:00 часа в зала 218 на ИИКТ-БАН на открито заседание.

Автор: Венелин Любомиров Тодоров

Заглавие: Методи Монте Карло за многомерни интеграли и интегрални уравнения и приложения

Увод

Методите Монте Карло (МК) са методи за приближено пресмятане на решението на задачи от изчислителната математика чрез използване на случайни процеси, като параметрите на съответния процес съвпадат с решението на задачата. Методът може да гарантира, че грешката при приближеното пресмятане на неизвестната величина е по-малка от зададена стойност с определена вероятност [7]. Методите Монте Карло се използват за оценяване на величини, които могат да бъдат представени като математическо очакване, чрез осредняване на резултатите, получени от голям брой статистически опити. Методите Монте Карло са основно средство за числено решаване на класове задачи в такива важни области като физика на елементарни частици, инженерна химия, молекулярна динамика, финансова математика. Основно научно предизвикателство в развитието на съвременните Монте Карло методи е тяхната относително бавна скорост на сходимост, която в редица случаи има асимптотика $O(N^{-1/2})$, където N е размерът на извадката. Има два подхода за подобряване на сходимостта - намаляване на дисперсията на оценяваната величина и намаляване на дискрепанса на използваните редици.

През Втората световна война, съвместната работа на големи учени като фон Нойман, Ферми, Улам и Метрополис и появата на съвременните дигитални компютри, даде силен тласък на теорията, заложена в методите Монте Карло. Статията на Metropolis и Ulam, публикувана през 1949 г. и озаглавена „The Monte Carlo Method“ [38], се възприема за първата статия, обвързваща наименованието „Монте Карло“ с използването на случайни величини [46]. Но както отбелязва Ермаков в [16] сведения за решение на реален проблем с метод Монте Карло съществува още в Стария Завет, където цар Соломон строи Божия храм. Може би най-ранното документирано използване на случайна извадка за намиране на приближена стойност на интеграл, е това на Буфон за изчисляване на числото π чрез хвърляне на игла. Терминът „Монте Карло“ е въведен от учените, работещи по разработването на ядрени оръжия в Лос Аламос през 40-те, където се е търсел отговор на въпроса дали е възможно предизвикването на ядрена реакция. Известно е, че множество неутрони, движещи се в уран, могат да предизвикат по случаен начин последваща емисия на

други неутрони, но не е можело да се предвиди теоретично дали веригите от реакции, образуващи сложна мрежа, ще предизвикат атомна експлозия. Учените използват компютъра *ENIAC*, за да моделират случайните траектории на неутроните през атомите на урановия заряд. Проектът е бил секретен с кодово наименование „Manhattan“.

Монте Карло методите предлагат простота на конструкциите и често се използват за симулация на процеси, чието поведение може да се интерпретира само в статистически смисъл. Съществуват широк клас задачи, за които методите Монте Карло са единствените възможни числени методи за решаване. При сравнението на обикновения метод Монте Карло за числено интегриране с детерминистични методи се налага изводът, че този подход е един от най-ефективните при числено пресмятане на многомерни интеграли, особено в случая на много високи размерности и без да е отчетена гладкостта на подинтегралната функция (вж. [7]). Независимо от универсалността на Монте Карло методите, техен сериозен недостатък е слабата им сходимост, основана върху грешка с порядък $O(N^{-1/2})$ при статистическа извадка с размер N , когато не се използва допълнителна информация за данните [2]. Дори скромни подобрения в тези методи имат съществен принос по отношение на ефективността и обхвата им на приложимост. Голяма част от усилията в разработването на Монте Карло методите, са насочени към конструиране на методи с намалена дисперсия. За тази цел могат да се използват различни адаптивни процедури, които използват предварителна и/или апостериорна информация, получена в процеса на пресмятанията. В общия случай, при реализациите на равномерно разпределена случайна величина може да се наблюдава струпване в определени подобласти и това да доведе до повишаване на дисперсията, тъй като реализациите не са толкова добре равномерно разпределени в областта, колкото се очаква. Затова изследванията за намаляване на дисперсията са насочени към контролирано разпределяне на реализациите в съответната област. Конструирането на метода на „слоистата“ извадка („stratified sampling“) и квази-Монте Карло методи е резултат от тези усилия. Квази-Монте Карло методите използват квазислучайни редици вместо обичайните псевдослучайни редици. Докато псевдослучайните числа са конструирани, така че да симулират поведението на истинските случайни числа, тези силно равномерно разпределени числа, на-

речени квазислучайни числа, са конструирани да бъдат толкова равномерно разпределени, колкото е математически възможно. Квазислучайните числа са конструирани да минимизират мярката за тяхното отклонение от равномерността, наречена дискрепанс. Така по-високият порядък на сходимост може да бъде получен чрез използване на детерминистични равномерно разпределени редици. Приближеното интегриране с използване на квазислучайни редици има по-бърза сходимост, с порядък $(O)(N^{-1} \log^s N)$, където s е размерността на задачата (за сравнение, обикновеният Монте Карло метод има порядък на сходимост $(O)(N^{-1/2})$). В последните десетилетия квази-Монте Карло методи известни като точкови множества от тип решетка, стават все по-популярни. Много задълбочено изследване на тези алгоритми е направено от Слоан [44]. Този метод е подходящ за гладки и периодични функции в s -мерния хиперкуб.

Съществуват два класа методи Монте Карло - преки и итерационни. Прекият метод се характеризира само с един тип грешка, наречена вероятностна. Тъй като с итерационните методи Монте Карло се приближава числено съответното итерационно приближение на точното решение, този клас методи се характеризира с два типа грешки: вероятностна, която зависи от броя на реализациите на веригата на Марков и систематична, която зависи от броя на итерациите в процеса. Задачата за балансиране на систематичната и статистическата грешка е изключително важна, когато се прилагат алгоритми от тип Монте Карло. Тази задача изисква намирането на оптимално съотношение между броя на реализациите и средната дължина на случайните траектории, което спомага за повишаване на ефективността на алгоритъма (вж. [7]).

Частните диференциални уравнения (ЧДУ) се утвърдиха като много успешно средство за математическо моделиране на различни процеси в екологията и други области. Тъй като все по-сложни процеси са обект на моделиране, това се отразява и върху съответните ЧДУ. Численият метод, построен за решаване на дадена математическа задача, която моделира реален процес от практиката, трябва да притежава редица качества. От една страна характеристики, свързани с теоретичните изисквания на числения анализ, като съвместимост, точност, устойчивост и сходимост, а от друга страна, трябва да предава адекватно свойствата на диференциалната задача - например положителност, монотонност, плътност и други. От особена важност е и реализацията на разработения

алгоритъм и неговата ефективност, която се отчита по следните показатели - точност на численото решение, компютърно време и оперативна памет, необходими за пресмятанията. Диференчните схеми са често използвани за решаването на частни диференциални уравнения, поради лесната им реализация и изчислителната им ефективност. При решаването на елиптични и параболични ЧДУ, най-често използваните диференчни схеми са от втори ред. Тези диференчни схеми се наричат стандартни и се прилагат при решаването на много задачи. При използването на стандартните диференчни схеми от втори ред, за да се постигне исканата точност, се получават големи системи, които изискват много голяма изчислителна памет и голям брой процесори върху съвременните суперкомпютри. Напоследък, голямо усилие се съсредоточава в разработването на компактни диференчни схеми с висок ред на точност, които използват само възлите на мрежата, съседни на централния възел. Идеята, която ще използваме в настоящото изследване, е да работим върху диференциалните уравнения, така че да изразим производните от по-висок ред в локалната грешка на апроксимацията. В последните години все повече нараства интересът към компактните диференчните схеми за решаване на параболични ЧДУ в екологията, поради необходимост от точно оценяване на замърсяването в големите европейски градове.

Цели и задачи на дисертационния труд

Целта на настоящата дисертация е разработването, реализирането и изследването на ефективни алгоритми Монте Карло (квази-Монте Карло) за многомерни интегрални уравнения, съответно линейни системи. Важна част е приложението на изследваните алгоритми в много области, от които основните са финанси, физика, биология, екология. Допълнителна цел е конструирането на нови числени методи с висок ред на точност на базата на метода на диференчните схеми за модели, свързани с екологичната безопасност.

Конкретните задачи за постигането на тази цел са:

1. Да се разработи и изследва нов почти оптимален алгоритъм Монте Карло за интегрални уравнения, базиран на балансиране на систематичната и стохастичната грешка. Да се получат оценки за броя на реализациите и

броя на итерациите и да се изследва ефективността на алгоритъма върху интегрални уравнения с приложен характер.

2. Да се реализира квази-МК метод от тип решетки с генериращ вектор обобщената редицата на Фибоначи, МК метод базиран на извадката латински хиперкуб и адаптивен алгоритъм МК. Да се изследва кой от алгоритмите е най-ефективен в зависимост от размерността на интеграла и гладкостта на подинтегралната функция.
3. Да се приложат разгледаните методи върху многомерни интеграли с приложен характер във финансите и в статистиката, както и да се получат оценки за ядрото на Вигнер в квантовата механика и да се изследва кой от методите е най-ефективен за решаване на задачата на Ричард Файнман.
4. Да се конструира нов метод МК за линейни системи на базата на метода МК за линейни системи „случайно блуждаене по уравненията“ и да се направи сравнение между двата алгоритъма и рафинирания алгоритъм МК.
5. Да се разработи нова компактна схема с четвърти ред на точност за системи от параболични ЧДУ. Да се повиши точността с помощта на екстраполация по Ричардсон. Да се приложи компактната схема върху модел на далечен пренос на замърсители във въздуха, както и за други примери.

Методология на изследването

Методологията на настоящите изследвания се основава на фундаментални резултати от следните области:

- функционален анализ [27] - функционални редове (сходимост, ред на Neumann, грешка от „отрязването“ на ред („truncation error“), ред на Тейлър);
- теория на вероятностите [4] - случайни величини, гранични теореми, изместени и неизместени оценки, вериги на Марков;
- числен анализ [1] - апроксимация, квадратурни формули;

- числени методи за диференциални уравнения и теория на диференчните схеми [3].

Програмните кодове на конструираните алгоритми са написани на програмната среда за пресмятания MATLAB [56].

Апробация на резултатите

Резултати, включени в дисертацията, са докладвани на: съвместен семинар на секции „Паралелни алгоритми“ и „Научни пресмятания“ на ИИКТ-БАН;

Част от резултатите са представени на следните конференции:

1. 9th International Conference on „Large-Scale Scientific Computations“ (LSSC’13), Созопол, България, 2013;
2. 8th International Conference on Numerical Methods and Applications: NMA’14, Боровец, 2014;
3. 10th IMACS Seminar on Monte Carlo Methods, (MCM 2015), Linz, Austria, July 6-10, 2015;
4. Advanced Computing for Innovation, AComIn 2015, 10-11 ноември, 2015, София;
5. Международната конференцията Numerical Methods for Scientific Computations and Advanced Applications, NMSCAA’16, 29 май - 2 юни, Хисаря, 2016;
6. APPLICATION OF MATHEMATICS IN TECHNICAL AND NATURAL SCIENCES: 8th International Conference for Promoting the Application of Mathematics in Technical and Natural Sciences - AMiTaNS’16, 22–27 юни 2016, Албена.

Списък на публикациите по дисертацията

Основните резултати по дисертацията са публикувани в:

1. I.T. Dimov, R. Georgieva, V. Todorov, Balancing of Systematic and Stochastic Errors in Monte Carlo Algorithms for Integral Equations, 8th International

Conference Numerical Methods and Applications, NMA 2014, Borovets, Bulgaria, August 20-24, 2014, Numerical Methods and Applications (I. Dimov, S. Fidanova, and I. Lirkov - Eds.), LNCS 8962, Springer, 2015, 44–51. *DOI* : 10.1007/978 – 3 – 319 – 15585 – 2_5. ISSN: 0302-9743. **SJR (2015): 0.252.**

2. Dimov, I.T., Todorov, V., Error Analysis of Biased Stochastic Algorithms for the Second Kind Fredholm Integral Equation, International Conference on Advanced Computing for Innovation, AComIn 2015, Sofia; Bulgaria, 10-11 November 2015, 2015 (Margenov S., Angelova G., Agre G. eds.), Studies in Computational Intelligence 648, 2016, 3-16. Springer Verlag. *DOI* : 10.1007/978 – 3 – 319 – 32207 – 0_1. ISSN: 1860-949X. **SJR(2015): 0.187.**
3. V. Todorov, I. Dimov. Monte Carlo methods for multidimensional integration for European option pricing, DOI:10.1063/1.4965003, AIP Conf. Proc. 1773, 100009, ISSN 0094243X, (2016), **SJR (2016): 0.198.**
4. Dimov I., J. Kandilarov, V. Todorov, L. Vulkov. High-Order Compact Difference Schemes with Richardson Extrapolation for Semilinear Parabolic Systems. IN: Applications of Mathematics in Engineering and Economics, American Institute of Physics, 1789, 030002 (2016), DOI: 10.1063/1.4968448, **SJR (2016): 0.198.**
5. V. Todorov. Computing High Dimensional Integrals with Monte Carlo Methods, Journal Scientific and Applied Research, Journal Scientific and Applied Research, 10, 2016, 11-16, Konstantin Preslavsky Publishing House, ISSN 1314-6289.

Четири от представените публикации са в издания със SJR фактор.

Втората статия е цитирана в изданието с импакт фактор:

1. Farshid Mirzaee, Nasrin Samadyar. Application of orthonormal Bernstein polynomials to construct a efficient scheme for solving fractional stochastic integro-differential equation, Optik - International Journal for Light and Electron Optics, 132, 2017, pp. 262–273, DOI:10.1016/j.ijleo.2016.12.029, **IF: 0.742.**

Алгоритми Монте Карло за многомерни интеграли

Тази глава е посветена на изследването, реализирането, сравнението и приложението на различни методи Монте Карло и квази-Монте Карло за многомерно интегриране. В дисертацията е направено подробно описание на методът Монте Карло, базиран на извадката латински хиперкуб (LHS). Реализиран е адаптивен метод Монте Карло за многомерни интеграли, който използва само апостериорна информация за порядъка на дисперсията (стандартното отклонение), но не и за гладкостта. Изследвана е изчислителната сложност на алгоритъма, която е съпоставена със сложността на обикновения метод Монте Карло. Направени са числени експерименти с тестови функции на Генц [18], които доказват очаквания ефект на намаляване на дисперсията. Конструиран е и квази-Монте Карло метод от тип решетки, с генериращ вектор обобщената редица на Фибоначи от съответната размерност. Направено е сравнение с квази-случайните редици на Соболев за гладки функции и при оценка на европейски опции във финансите. Най-важното приложение е при оценка на ядрото на Вигнер в квантовата механика. Полученият резултат е оригинален по своя характер и отговаря на въпрос, поставен преди няколко десетилетия от един от най-забележителните физици на 20 век Ричард Файнман за съществуването на алгоритъм с линейна или полиномиална, но не и експоненциална, изчислителна сложност за ядрото на Вигнер. Всички представени методи от тип Монте Карло/квази-Монте Карло дават по-добри резултати от досега използваните методи за оценка на ядрото, като адаптивният алгоритъм се оказва най-ефективен, поради поведението на ядрото.

Резултатите, представени в тази глава, са публикувани в [47, 48].

Описание на алгоритмите

Извадка латински хиперкуб

Подходът на „слоистите“ извадки („stratified sampling“; вж. [46]) се основава само на начално разбиване на областта и подходящ избор на броя на случайните точки в съответната подобласт, което да гарантира намаляване на дисперсията в сравнение с обикновения метод Монте Карло, но не и повишаване на порядъка $\mathcal{O}(\sqrt{1/N})$ на вероятностната грешка. Избират се повече точки в областите,

в които локалната дисперсия е голяма. Подходът не съдържа идеята за адаптивност (рекурсивно разделяне на подобласти) и не използва предварителна информация за гладкостта на подинтегралната функция.

Теорема 1. *Теорема(Соболъ, [46]) За фиксирано разбиване на областта на интегриране и $N_j = N p_j$ е в сила, че*

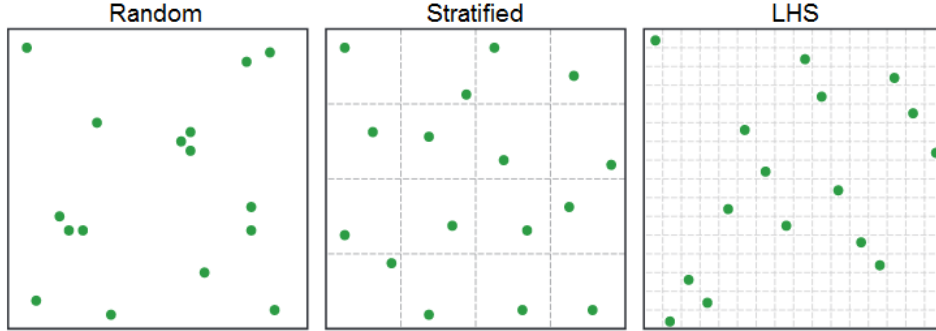
$$D\bar{\theta}_N^* \leq D\bar{\theta}_N,$$

където $D\bar{\theta}_N$ е дисперсията на оценката на интеграла, получена с обикновен метод Монте Карло и N случайни точки в областта на интегриране, а $D\bar{\theta}_N^*$ е дисперсията на оценката на интеграла, получена с метода на „слоистите“ извадки, и p_j е плътността в съответната подобласт.

Съществен недостатък при прилагане на този подход, е разделянето на подобласти. Ако една s -мерна област се разбие на равномерни подобласти чрез разделяне на две по всяко направление, се получават 2^s подобласти. В случая на големи s броят на подобластите рязко нараства, което ограничава силно избора на размерност на случайната извадка. Този проблем се преодолява от извадката "latin hypercube sampling"(LHS) [26]. Извадката LHS е специален вид „слоиста“ извадка, използвана за пръв път през 1979г. в [37] и развита в по скорошни публикации – виж [39]. В двумерния случай квадратна мрежа е латински квадрат, тогава и само тогава, когато във всеки ред и стълб на квадрата попада само една реализация на случайната величина. Латински хиперкуб е обобщение на тази концепция за произволна размерност. Тази техника изисква точно по една случайна точка във всяка подобласт, това свойство е едно от основните предимства на извадката латински хиперкуб. За по-голяма яснота на Фиг. 1 е дадена извадката латински хиперкуб, сравнена със „слоистата“ и случайна извадка с 16 точки ($s = 2$ е размерността, а $M = 4$ са броя подобласти).

Адаптивен алгоритъм Монте Карло

Изследван е адаптивен алгоритъм, базиран на алгоритъма, описан в [8, 20, 28]. Основната идея е концентриране на случайни точки в подобластите, в които дисперсията е голяма (по отношение на предварително зададена точност). Подходът се основава на рекурсивно разделяне на областта, използвайки



Фигура 1: Сравнение на различни типове извадки

апостериорна информация за грешката при текущото разделяне. Така адаптивният алгоритъм дава приближение с грешка $\varepsilon \leq c N^{-1/2}$, където $c \leq 0.6745\sigma(\theta)$ ($\sigma(\theta)$ е стандартното отклонение), но порядъкът е същият, както при обикновения алгоритъм Монте Карло. Реализираният тук адаптивен алгоритъм не използва никаква предварителна информация за гладкостта на подинтегралната функция, но използва апостериорна информация за дисперсията. Адаптивният алгоритъм се оказва най-ефективен за функции с особености в изчислително отношение - при ядрото на Вигнер и тестовите функции на Генц, той дава относителни грешки с няколко порядъка по-добри от останалите алгоритми. Алгоритъмът започва с разделяне на интервалите по всички направления на M подинтервала, като M е зададено като входящ параметър. Във всяка подобласт се пресмята стойността на съответния интеграл $I_{\Omega_j}, j = 1, \dots, M$ и дисперсия. След това получената дисперсия се сравнява с предварително зададена стойност. Получената информация се използва за следващото съгъстяване, тъй като подобластта с най-голямо стандартно отклонение се разделя на 2^s нови подобласти. Изпълнението на алгоритъма спира, когато за стандартното отклонение е достигната предварително зададената точност ε във всички подобласти, получени след делението (или когато е достигнат предварително зададен максимален брой на нивата на разделяне или на подобластите, в които не е изпълнен стоп-критерият). Така се получава едно приближение на интеграла $I = \sum_j I_{\Omega_j}, j = 1, \dots, M^s$ с подход Монте Карло.

Квази-Монте Карло методи с използване на точкови множества от

тип решетка

Да разгледаме квадратурата

$$I_N(f) = \frac{1}{N} \sum_{j=0}^{N-1} f(x_j),$$

където $P_N = \{x_0, x_1, \dots, x_{N-1}\}$, $x_i \in [0, 1]^s$ е множество от точки, задаващо интеграционните възли на формулата. Изборът на възлите е много съществен, тъй като от това зависи дискрепансът и точността на полученото приближение. Практическо приложение е намерил теоретико-числовият метод за интегриране, предложен от Корабов [31], при който възлите на квадратурната формула имат вида:

$$x_k = \left(\left\{ \frac{kz_1}{N} \right\}, \left\{ \frac{kz_2}{N} \right\}, \dots, \left\{ \frac{kz_s}{N} \right\} \right), \quad k = 1, 2, \dots, N, \quad (1)$$

където N е зададен брой точки, z е s -мерен целочислен вектор, а чрез $\{a\}$ означаваме дробната част на числото a , т.е. $\{a\} = a - [a]$. Векторът z се нарича генериращ вектор или генератор на множеството.

Да поставим допълнително изискване към гладкостта на функцията $f(x)$, а именно да предположим, че функцията $f(x)$ принадлежи на класа на Корабов $E_s^\alpha(C)$, зададен по следния начин [45]:

Дефиниция 1. Казваме, че $f(x)$ принадлежи на класа $E_s^\alpha(c)$ за $\alpha > 1$ и $c > 0$, ако f е периодична функция с период единица по всяка от променливите си $x_i, i = 1, 2, \dots, s$, дефинирана върху единичния хиперкуб $[0, 1]^s$ и коефициенти й на Фурие удовлетворяват неравенството:

$$|a(m)| \leq \frac{C}{(\overline{m}_1 \dots \overline{m}_s)^\alpha}, \quad (2)$$

където

$$\overline{m} = \begin{cases} |m|, & \text{ако } |m| \neq 0, \\ 0, & \text{ако } m = 0. \end{cases}$$

и константата c не зависи от $m_1, \dots, m_s$.

Теорема 2. ([5]) Ако $f(x) \in E_s^\alpha(c)$, то съществува оптимален избор на генериращ вектор z , за който за грешката от интегриране е в сила

$$\left| \frac{1}{N} \sum_{k=0}^{N-1} f\left(\left\{\frac{k}{N}z\right\}\right) - \int_{[0,1]^s} f(u)du \right| \leq Cd(s, \alpha) \frac{(\log N)^{\beta(s, \alpha)}}{N^\alpha} \quad (3)$$

за функция $f \in E_s^\alpha(c)$, $\alpha > 1$ и $d(s, \alpha), \beta(s, \alpha)$ не зависят от N . Нещо повече, ако N е просто число, то $\beta(s, \alpha) = \alpha(s - 1)$.

Векторът z , за който неравенството (3) е изпълнено, се нарича добър или оптимален генериращ вектор, а неговите компоненти - оптимални коефициенти в смисъла на Корабов. Точковото множество P_N е множество от добри целочислени точки, а съответният метод за числено интегриране с квадратурата се нарича метод на добрите целочислени точки (Good Lattice Point method - GLP метод). Известни са редица теореми, доказващи съществуването на оптимални генериращи вектори. Трудността е в конструирането на тези оптимални вектори, особено в задачи от голяма размерност. При $s = 2$ оптимална конструкция съществува. Вакхавалов през 1959г. в [5] представя конструкция, основана на числата на Фибоначи. През 1981 г. в [25] е направено обобщение на редицата на Фибоначи за произволна размерност. Ниа и Wang конструират точково множество от тип решетка със следния генериращия вектор [25, 50]:

$$z = (1, F_n(2), \dots, F_n(s)), \quad (4)$$

за някое естествено число n . В сила е $F_n(j) = F_{n+j-1} - F_{n+j-2} - \dots - F_n$, където F_i са съответните обобщени числа на Фибоначи с размерност s , т.е.:

$$F_{l+s} = F_l + F_{l+1} + \dots + F_{l+s-1}, l = 0, 1, \dots \quad (5)$$

с начално условие:

$$F_0 = F_1 = \dots = F_{s-2} = 0, F_{s-1} = 1, \quad (6)$$

за $l = 0, 1, \dots$. След опростяване, може да се види, че генериращият вектор по-горе е:

$$z = (1, F_{n-1} + F_{n-2} + \dots + F_{n-s+1}, \dots, F_{n-1} + F_{n-2}, F_{n-1}) \quad (7)$$

Според [25], имаме следната оценка за дискрепанса на множеството, получено с използването на този вектор и F_n брой точки:

Теорема 3. *Множеството*

$$\left(\left\{ \frac{1}{F_n} k \right\}, \left\{ \frac{F_n(2)}{F_n} k \right\}, \dots, \left\{ \frac{F_n(s)}{F_n} k \right\} \right), \quad 1 \leq k \leq F_n.$$

има дискрепанс

$$D(F_n) = \mathcal{O} \left(F_n^{-\frac{1}{2} - \frac{1}{2^{s+1} \cdot \log 2} - \frac{1}{2^{2s+3}}} \right).$$

Предимството на разглеждания метод е, че броят на операциите, необходими за получаване на генериращия вектор, е асимптотично $\mathcal{O}(\log F_n)$. След като имаме този вектор, генерирането на нова точка изисква константен брой операции. Тъй като трябва да генерираме F_n точки, за да получим точково множество от тип решетка от разглеждания вид с F_n точки, ще бъдат необходими $\mathcal{O}(F_n)$ брой операции.

Приложение в Бейсовската статистика

Фундаментален проблем в Бейсовската статистика е точното пресмятане на многомерни интеграли от следните два вида, описани от Shaowei Lin през 2011 в [33, 34]. Първият вид интеграли имат следния вид: $\int_{\Omega} p_1^{u_1}(x) \dots p_k^{u_k}(x) dx$, където $\Omega \in \mathcal{R}^s$, $x = (x_1, \dots, x_d)$, $p_i(x)$ са полиноми и u_i са цели числа. Вторият вид интеграли имат вида $\int_{\Omega} e^{-Nf(x)} \phi(x) dx$, където N е естествено число и $f(x)$, $\phi(x)$ са многомерни полиноми. Такива са 5-мерния и 15-мерния многомерни интеграли, дадени по-долу.

$$\int_{[0,1]^5} \exp(-100x_1x_2x_3)(\sin(x_4) + \cos(x_5))dx \approx 0.1854297367. \quad (8)$$

$$\int_{[0,1]^{15}} \left(\sum_{i=1}^{10} x_i^2 \right) (x_{11} - x_{12}^2 - x_{13}^3 - x_{14}^4 - x_{15}^5)^2 dx \approx 1.96440666. \quad (9)$$

$$\int_{[0,1]^{30}} \frac{4x_1x_3^2e^{2x_1x_3}}{(1+x_2+x_4)^2} e^{x_5+\dots+x_{20}} x_{21} \dots x_{30} dx \approx 3.244540. \quad (10)$$

Направено е детайлно сравнение между обикновен Монте Карло алгоритъм (Crude), адаптивен алгоритъм Монте Карло (Adapt), алгоритъм с използване на квазислучайната редица на Соболев (Sobol), алгоритъм с използване на точково множество от тип решетки базирано на обобщената редица на Фибоначи (FIBO) и извадка от тип латински хиперкуб (LHS). Резултатите са дадени в Таблицы 1, 2 и 3.

За сравнително малка размерност за 5-мерния интеграл методът FIBO постига най-малка относителна грешка при предварително зададен брой реализации - виж Таблица 1. За размерност 15 Sobol постига по-малки относителни грешки от FIBO - виж Таблица 2. Анализите показват, че за висока размерност 30 извадката LHS превъзхожда FIBO и Sobol - виж Таблица 3. От таблиците може да се направи извода, че методът FIBO е ефективен за пресмятане на многомерни интеграли от гладки подинтегрални функции при сравнително малки размерности, заради високата точност и ниската изчислителна сложност. За по-големи размерности за предпочитане е методът LHS. Методът на Собол е значително по-бавен от FIBO и LHS, затова дори в случаите на сходни относителни грешки FIBO е по-ефективен, защото трудоемкостта му е по-малка. Адаптивният метод е най-бавен, заради рекурсивното делене на подобласти и е най-подходящ за интегриране на функции с особености, което ще бъде показано в следващите параграфи.

Таблица 1: Относителна грешка за 5-мерния интеграл

N	Crude	време(s)	Adapt	време(s)	FIBO	време(s)	Sobol	време(s)	LHS	време(s)
10^3	2.10e-2	0.007	2.15e-3	0.27	1.75e-4	0.007	5.29e-4	0.03	9.38e-3	0.007
10^4	4.52e-3	0.07	2.01e-3	2.43	1.28e-5	0.06	1.43e-4	0.3	3.44e-3	0.07
10^5	1.19e-3	0.64	6.91e-4	22.2	9.50e-6	0.61	2.36e-5	2.77	2.01e-3	0.69
10^6	8.47e-4	6.06	2.92e-4	219.5	5.47e-7	5.98	6.07e-6	24.2	1.80e-4	6.17
10^7	2.38e-4	59.9	8.21e-5	2043	8.71e-9	58.4	2.30e-6	245	2.46e-5	60.5

Таблица 2: Относителна грешка за 15-мерния интеграл

N	Crude	време(s)	Adapt	време(s)	FIBO	време(s)	Sobol	време(s)	LHS	време(s)
10^3	6.31e-2	0.09	3.16e-3	8.24	5.34e-2	0.08	2.04e-3	0.98	1.06e-2	0.12
10^4	4.30e-2	0.95	1.49e-3	68	1.22e-3	0.93	2.89e-4	9.3	7.33e-3	1.07
10^5	2.77e-2	9.70	5.76e-4	547	1.08e-4	9.65	1.13e-5	93.8	1.54e-4	10.11
10^6	7.13e-3	95.8	1.29e-4	5235	6.37e-6	96.9	5.93e-6	735	1.14e-5	99.6

Таблица 3: Относителна грешка за 30-мерния интеграл

N	Crude	време(s)	Adapt	време(s)	FIBO	време(s)	Sobol	време(s)	LHS	време(s)
10^3	8.56e-1	0.02	1.56e-1	2.27	8.73e-1	0.02	1.29e-1	0.27	2.31e-2	0.02
10^4	7.13e-1	0.1	6.91e-2	20.1	1.19e-2	0.18	8.56e-2	2.5	6.89e-3	0.19
10^5	4.21e-1	1.12	3.76e-2	229	2.78e-2	1.56	1.91e-2	20.2	1.65e-3	1.54
10^6	1.73e-1	11.07	6.29e-3	2271	9.56e-3	13.61	9.47e-3	208	9.61e-5	13.9

Тестови функции на Генц

Методът на решетките не е приложим за функции с особености, както се вижда от числените експерименти в тази секция. Нека е дадена следната моделна функция:

$$f(x) = (1 + \sum_{i=1}^s a_i x_i)^{-(s+1)}. \quad (11)$$

Разглежданият клас от тестови функции принадлежи на пакет, предложен от Genz [18]. Избраното множество от функции имат единствен локален максимум в близост до един от върховете на многомерния единичен куб, подобно на някои моделни функции, описващи изменението в концентрациите на замърсители във въздуха.

Таблица 4: Относителна грешка и изчислително време в секунди за размерност $s = 5$, $I[f] = 2.12\text{e-}06$, $a = (5, 5, 5, 5, 4)$.

Адаптивен алг. МК			LHS			FIBO		
N	$I_N[f]$	(s)	N	$I_N[f]$	(s)	N	$I_N[f]$	(s)
10^2	$3.7735\text{e-}03$	0.33	10^5	$7.2274\text{e-}02$	0.27	1346269	$9.7135\text{e-}02$	0.38
10^3	$1.2877\text{e-}03$	1.44	10^6	$3.2518\text{e-}02\text{e-}02$	1.22	3524578	$6.7594\text{e-}02$	1.32
10^4	$4.2452\text{e-}04$	10.75	10^7	$2.5207\text{e-}03$	12.3	14930352	$1.5377\text{e-}02$	15.07
10^5	$4.7169\text{e-}05$	142.18	10^8	$1.6646\text{e-}03$	124.2	102334155	$2.9245\text{e-}03$	134.58

Резултатите, получени след прилагането на описаните алгоритми в тази глава - латински хиперкуб, точково множество от тип решетки, базирано на обобщената редица на Фибоначи и адаптивен алгоритъм Монте Карло, за интеграли

Таблица 5: Относителна грешка и изчислително време в секунди за размерност $s = 18$, $I[f] = 9.919\text{e-}06$, $a = \left(\frac{1}{9}, \frac{2}{27}, \frac{2}{27}, \frac{1}{9}, \frac{2}{27}, \frac{1}{9}, \frac{1}{9}, \frac{4}{27}, \frac{2}{27}, \frac{1}{9}, \frac{1}{9}, \frac{2}{27}, \frac{2}{27}, \frac{1}{9}, \frac{1}{9}, \frac{4}{27}, \frac{1}{9}, \frac{1}{9}\right)$.

Адаптивен алг. МК			LHS			FIBO		
N	отн. гр.	вр.(s)	N	отн. гр.	вр.(s)	N	отн. гр.	вр(s)
10	$9.2341\text{e-}04$	15.7	10^7	$8.6285\text{e-}03$	13.6	14930352	$7.1579\text{e-}02$	14.7
10^2	$8.0653\text{e-}05$	142	10^8	$5.1195\text{e-}03$	140	102334155	$5.1096\text{e-}02$	144.1
10^3	$1.0081\text{e-}05$	1408	10^9	$1.6283\text{e-}03$	1353.5	1134903170	$2.8883\text{e-}02$	1344.3

с размерност 5 и 18, са дадени съответно в Таблица 4 и Таблица 5. Изследвана е ефективността на реализирания адаптивен алгоритъм Монте Карло по отношение на относителна грешка и изчислително време. Резултатите показват, че за почти едно и също време адаптивният алгоритъм дава по-малки относителни грешки, дори в порядъци.

Приложение за ядрото на Вигнер

Един от най-известните физици на нашето време Ричард Файнман поставя преди няколко десетилетия въпроса за съществуване на ефективен алгоритъм с линейна или полиномиална изчислителна сложност за ядрото на Вигнер в многомерния случай [17]. Досега са използвани само детерминистични алгоритми с експоненциална изчислителна сложност, които страдат от „проклятието на размерността“, за което беше вече споменато. Трите постулата, които описват напълно новата математическа формулировка на квантовата механика са дадени в термините на частици със знак и тези три постулата са достатъчни да отговорят на резултатите на по сложни квантови теории [43]. Тук даваме само втория постулат:

Постулат II. Частица със знак, отделяща се в потенциал $V = V(x)$, се държи като класическа точкова частица без поле, която създава нови частици със знак

с вероятност $\gamma(x(t))dt$ с интервал dt , за времето dt , където

$$\gamma(x) = \int_{-\infty}^{+\infty} Dp' V_W^+(x; p') \equiv \lim_{\Delta p' \rightarrow 0^+} \sum_{M=-\infty}^{+\infty} V_W^+(x; M \triangle p'), \quad (12)$$

и $V_W^+(x; p)$ е положителната част на величината

$$V_W(x; p) = \frac{i}{\pi^s \hbar^{s+1}} \int_{-\infty}^{+\infty} dx' e^{-\frac{2i}{\hbar} x' p} [V(x+x') - V(x-x')], \quad (13)$$

известна като ядрото на Вигнер в s -мерното пространство [51]. Ако по време на създаването, създаващата частица има знак z , позиция x и момент p , новите частици имат същата позиция x , имат знаци $+z$ и $-z$, и моменти $p+p'$ и $p-p'$ съответно, избрани случайно в съответствие с нормализираната вероятност $\frac{V_W^+(x; p)}{\gamma(x)}$.

Използваме следния Вигнеров потенциал:

$$V = V(x) = x_1 \dots x_n, \quad x', x, p, x+x', x-x' \in [0, 1].$$

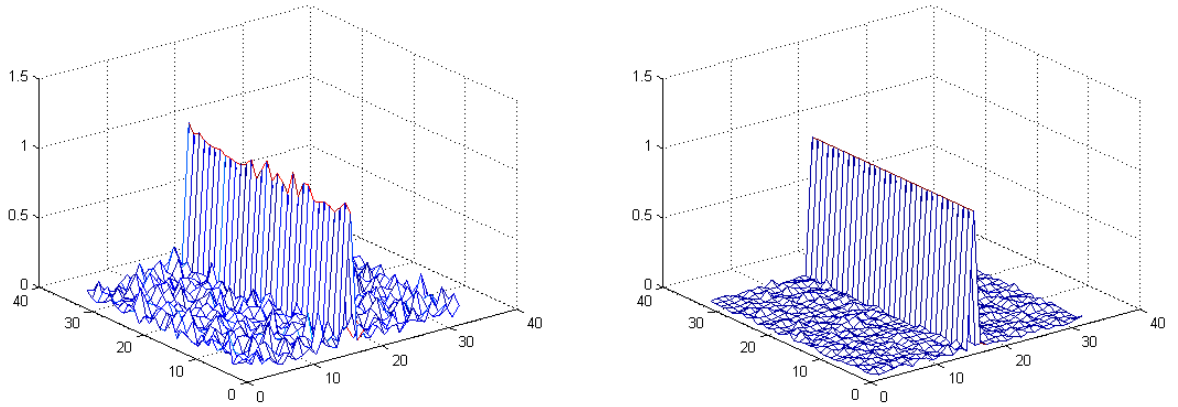
Целта на настоящето изследване е да се пресметне многомерния интеграл (13), представляващ ядрото на Вигнер.

Най-напред прилагаме детерминистичен метод на средните правоъгълници, обикновен метод Монте Карло и квази-Монте Карло с използване на квазислучайни редици на Соболев. Резултатите са дадени в Таблица 6. Детерминистичният метод е най-бавен и започва да губи точност с увеличаване на размерността на интеграла. Методът на Соболев е по-точен от обикновения метод Монте Карло, но е по-бавен - виж Таблица 6. Обикновеният алгоритъм Монте Карло е по-неточен от квази-Монте Карло алгоритъма с редици на Соболев, поради свойствата на добре разпределените редици.

Ядрото на Вигнер, получено с обикновен и адаптивен алгоритъм Монте Карло, е илюстрирано на Фиг. 2. На Фиг. 3 са показани позицията на частиците със знак в ядрото на Вигнер и върхът на ядрото на Вигнер, получени с адаптивния алгоритъм. От поведението на ядрото се вижда, че най-ефективен за пресмятането му би бил адаптивният алгоритъм Монте Карло, което се потвърждава и от проведените експерименти. Числените резултати показват, че методът от

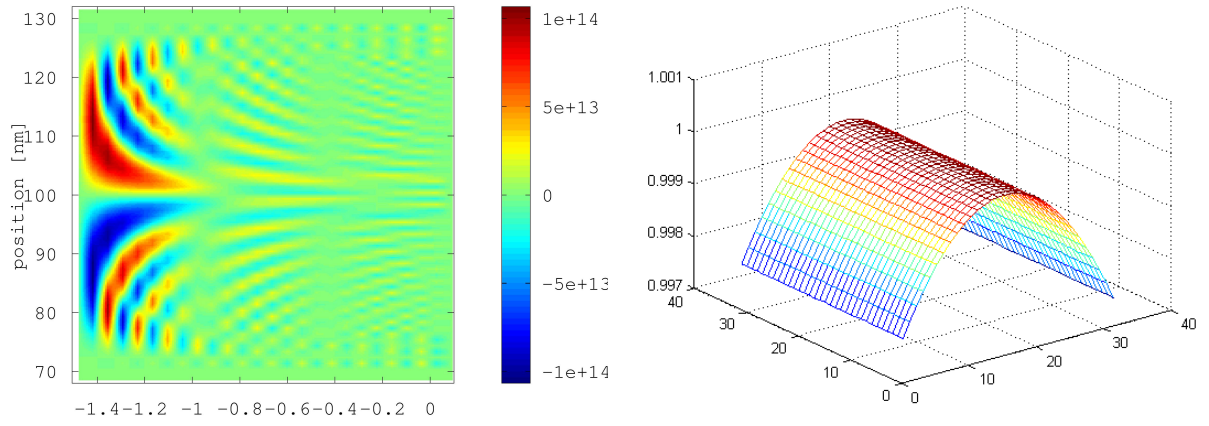
Таблица 6: Относителна грешка за ядрото на Вигнер с детерминистични и стохастични методи

s	N	Средни правоъг.	време (s)	Собол алг.	време (s)	Обикновен алг.	време (s)
3	$32^2 \times 50$	8.51e-03	0.2	8.47e-03	0.1	8.37e-03	0.003
	$32^2 \times 100$	8.21e-03	0.5	8.11e-03	0.21	8.19e-03	0.008
	$64^2 \times 50$	5.76e-03	1	5.26e-03	0.5	5.11e-03	0.1
	$64^2 \times 100$	4.89e-03	1.9	4.55e-03	1.1	4.76e-03	0.3
6	$8^4 \times 50^2$	1.16e-02	41.2	8.64e-04	19.5	9.09e-04	4.1
	$8^4 \times 100^2$	9.75e-03	160.6	5.73e-04	53.4	6.44e-04	17.9
	$16^4 \times 50^2$	7.84e-03	635.2	1.90e-04	149	4.37e-04	51.5
	$16^4 \times 100^2$	2.12e-03	2469.1	1.29e-04	601.6	3.80e-04	132.1
9	$6^6 \times 16^3$	1.75e-03	835.5	5.45e-04	188.5	9.35e-04	46.5
	$6^6 \times 32^3$	1.35e-03	5544.1	2.36e-04	1067.6	8.29e-04	225.1
	$6^6 \times 40^3$	1.12e-03	10684.4	1.07e-04	2190.5	5.12e-04	491.5



Фигура 2: Ядрото на Вигнер, получено с обикновен и адаптивен алгоритъм Монте Карло

тип решетки и този на Собол произвеждат подобни относителни грешки за един и същ брой реализации. За 9-мерния интеграл FIBO и Sobol дават сходни ре-



Фигура 3: Позиция на частиците в ядрото на Вигнер и върхът на ядрото на Вигнер с адаптивен алгоритъм

зултати, които са по-неточни от адаптивния алгоритъм и извадката LHS – виж Таблица 7. Адаптивният алгоритъм дава най-добри резултати в сравнение с останалите стохастични алгоритми с до два порядъка по-добри от LHS, но е значително по-бавен – виж Таблица 7. Разработеният адаптивен алгоритъм е с линейна изчислителна сложност, което е един възможен подход за много точно пресмятане на многомерния интеграл, описващ ядрото на Вигнер, и е решение на въпроса, поставен от Ричард Файнман, за който беше споменато в началото. Прави впечатление и добрият резултат на метода LHS, който заради бързината си също е един възможен подход за оценяване на ядрото. Предимството на адаптивния Монте Карло метод следва от вида на ядрото, илюстриран на Фиг. 2. С увеличаване на размерността предимствата на адаптивния алгоритъм пред останалите методи се засилват. Всички използвани методи дават значително по-добри резултати от досега използвания детерминистичен подход, както по отношение на точност, така и по отношение на изчислително време. Изследваният адаптивен алгоритъм позволява да се постигне висока точност със сравнително малка изчислителна сложност, което е един възможен подход за пресмятане на многомерния интеграл, описващ ядрото на Вигнер, и е решение на въпроса, поставен от Ричард Файнман [17].

Таблица 7: Относителна грешка за 3, 6 и 9 мерния интеграл

	N	Latin	t,s	адаптивен	t,s	FIBO	t,s	Sobol	t,s
$s = 3$	10^3	4.38e-03	0.01	6.75e-04	0.4	3.72e-02	0.02	1.07e-02	0.05
	10^4	7.94e-04	0.06	8.15e-05	3.3	7.06e-03	0.07	8.77e-03	0.54
	10^5	2.51e-04	0.41	5.01e-06	32.6	3.40e-03	0.43	8.57e-04	5.74
	10^6	8.20e-05	3.52	4.38e-07	302	1.01e-03	4.4	6.73e-04	45.6
$s = 6$	10^3	1.54e-03	0.01	2.23e-04	0.5	7.82e-03	0.01	2.42e-02	0.09
	10^4	6.34e-04	0.06	4.74e-05	4.1	5.01e-03	0.07	5.02e-03	0.78
	10^5	4.22e-04	0.44	5.43e-06	37	6.88e-03	0.43	4.60e-04	6.19
	10^6	8.57e-05	3.7	5.04e-07	351	7.68e-04	5.97	3.59e-04	53
$s = 9$	10^3	6.11e-03	0.04	8.23e-04	0.5	2.03e-02	0.06	5.42e-02	0.11
	10^4	1.02e-03	0.06	2.02e-05	4.7	2.02e-03	0.07	6.02e-03	0.88
	10^5	4.69e-04	0.43	1.08e-06	40	9.16e-04	0.53	3.57e-03	6.56
	10^6	8.08e-05	3.8	4.14e-07	381	7.13e-04	3.7	8.02e-04	57

Приложения за европейски опции

Опцията е договор (ценна книга), даваща на собственика правото, но не и задължението да продаде (пут опция) или да купи (кол опция) определено количество ценни книжа или валута на фиксирана в договора цена, в рамките на даден период от време. При опциите притежателя може да упражни или да не упражни своето право в съответния срок. При европейския вид опции собственикът им може да упражни правото си само в края на периода. Монте Карло и квази-Монте Карло методите могат директно да се приложат към финансови проблеми, включващи многомерни интеграли. По-долу ще илюстрираме как изследваните алгоритми могат ефективно да бъдат приложени към пресмятане на Европейски опции. Идеята се състои в следното: стойността на опцията V се формулира в термините на математическо очакване на случайна величина, след това средно аритметичното на независими реализации на случайната величина се използва за оценяване на опцията. Използва се рисково-неутралната

формула за Европейска опция:

$$V(S, t) = E(e^{-r(T-t)}h(S(T)) | S(t) = S, \mu = r), \quad (14)$$

където $E(\cdot)$ е математическото очакване, $h(S)$ е платежната функция, $h(S) = \max(S - E, 0)$ за кол опцията, $h(S) = \max(E - S, 0)$ за пут опцията.

Да разгледаме европейска опция, чиято платежна функция зависи от $k > 1$ актива с цени $S_i, i = 1, \dots, k$. Всеки актив следва случайното движение (така наречената случайна разходка): $dS_i = \mu_i S_i dt + \sigma_i S_i dX_i$, където σ_i е годишното стандартно отклонение на i -тия актив и dX_i е Брауновото движение. Да предположим, че при падежа T платежната функция се дава с: $h(S'_1, \dots, S'_k)$, (където S' означава стойността на i -тия актив при изтичане). Тогава текущата стойност на опцията, V , отхвърляйки възможността за арбитраж (получаване на безрискова печалба) ще бъде

$$V = e^{-r(T-t)}(2\pi(T-t))^{-k/2}(\det \Sigma)^{-1/2}(\sigma_1, \dots, \sigma_k)^{-1} \times \\ \times \int_0^\infty \dots \int_0^\infty \frac{h(S'_1, \dots, S'_k)}{S'_1, \dots, S'_k} \exp(-0.5 \alpha^T \Sigma^{-1} \alpha) dS'_1, \dots, dS'_k,$$

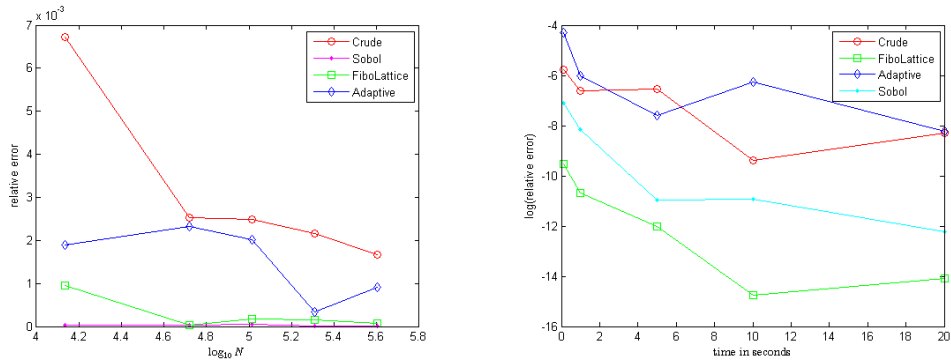
където $\alpha_i = (\sigma_i(T-t)^{1/2})^{-1}(\log \frac{S'_i}{S_i} - (r - \frac{\sigma_i^2}{2})(T-t))$, r е лихвеният процент и Σ е матрицата на ковариация, където елемента (i, j) е ковариацията на dX_i и dX_j за k актива. За задачи като тази, многомерните интеграли могат да се оценят чрез Монте Карло и квази-Монте Карло методи. За да получим търсената формулировка, безкрайната област на интегриране може да се изобрази в s -размерния единичен куб по различни начини. Например, $\frac{2}{\pi} \arctan(x)$ изобразява $(0, \infty)$ в $(0, 1)$. Може да се използва и функция на разпределение на различни случайни величини. Такава трансформация преобразува проблема в такъв, при който търсим стойността на многомерния интеграл $\int_0^1 \dots \int_0^1 g(x_1, \dots, x_s) dx_1 \dots dx_s$ над единичния хиперкуб. Когато g е функцията експонента, с подходящ избор на константите включени в горното уравнение [32] получаваме s -мерния интеграл $\int_{[0,1]^k} \exp(x_1, \dots, x_s) dx_1 \dots dx_s$ Числените експерименти включват пресмятането на 5 и 20-мерни интеграли:

$$\int_{[0,1]^5} \exp(\sum_{i=1}^5 0.5 a_i x_i^2 (2 + \sin \sum_{j=1, j \neq i}^5 x_j)) dx \approx 2.923651, \quad a_i = (1, 0.5, 0.2, 0.2, 0.2), \quad (15)$$

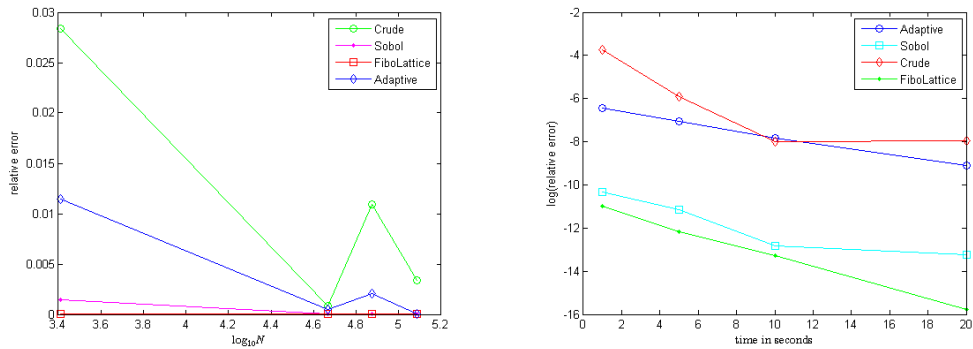
Таблица 8: Относителна грешка за 5-мерния интеграл

време в сек.	обикновен МК	адаптивен МК	FIBO	Собол
0.1	3.07e-3	1.34e-2	7.26e-5	8.22e-4
1	1.32e-3	2.44e-3	2.28e-5	2.91e-4
5	1.43e-3	4.93e-4	5.94e-6	1.71e-5
10	8.47e-5	1.88e-3	3.85e-7	1.79e-5
20	2.52e-4	2.71e-4	7.49e-7	4.96e-6

$$\int_{[0,1]^{20}} \exp\left(\prod_{i=1}^{20} x_i\right) dx \approx 1.00000949634. \quad (16)$$



Фигура 4: Относителна грешка и изчислително време за 5-мерния интеграл.



Фигура 5: Относителна грешка и изчислително време за 20-мерния интеграл.

Таблица 9: Относителна грешка за 20-мерния интеграл

време в сек.	обикновен МК	адаптивен МК	FIBO	Sobol
1	9.14e-3	1.58e-3	1.48e-5	3.25e-5
2	3.68e-3	1.028e-3	9.17e-6	3.97e-5
5	2.67e-3	8.58e-4	5.19e-6	1.45e-5
10	3.34e-4	4.02e-4	1.73e-6	2.71e-6
20	1.53e-4	1.13e-4	1.38e-7	1.76e-6

За всеки интеграл е пресметната относителната грешка при отнапред зададен брой реализации и получената относителна грешка при отнапред зададено изчислително време. Прави впечатление, че при малки размерности предимствата на FIBO са безспорни – виж Таблица 8 и Фиг. 4. При голяма размерност за 20-мерния интеграл методът FIBO и този с редици на Соболев дават близки резултати, но пак FIBO има предимство – виж Таблица 9. Адаптивния алгоритъм има предимства пред обикновения метод Монте Карло с увеличаване на размерността, но както видяхме той е най-подходящ за функции с особености в изчислително отношение. Най-бърз алгоритъм е обикновеният метод Монте Карло, като методът FIBO има сходна бързина. Може да се направи извода, че FIBO е най-добър избор за малки размерности, като дори за 20-мерния интеграл, има предимство по отношение на бързина и точност спрямо квази-Монте Карло алгоритъма с редици на Соболев – виж Фиг. 5.

Алгоритми Монте Карло за интегрални уравнения

В първа секция на втора глава е описан модифициран алгоритъм Монте Карло за приближено пресмятане на линеен функционал от решението на интегрално уравнение на Фредхолм от втори род, базиран на балансиране на систематичната и стохастичната грешка.

Резултатите, представени в тази глава, са публикувани в [9, 14] и са цитирани в статията с IF: [40].

Разглежда се задачата за приближено пресмятане на линеен функционал

(φ, u) от решението на интегралното уравнение на Фредхолм от втори род:

$$u(x) = \int_{\Omega} k(x, x')u(x')dx' + f(x), \quad x, x' \in \Omega \subset \mathbb{R}^d, \quad \text{или} \quad u = \mathcal{K}u + f,$$

където $\mathcal{K}$ е интегрален оператор, $k(x, x') \in L_2(\Omega \times \Omega)$, $f(x) \in L_2(\Omega)$ са дадени функции, $u(x) \in L_2(\Omega)$ е неизвестна функция, Ω е ограничена област. В съответствие с началната $\pi(x)$ и преходните $p(x, x')$ вероятности в Ω се конструира случайна траектория (верига на Марков) T_k с дължина k , стартираща със състояние x_0 : $T_k: x_0 \longrightarrow x_1 \longrightarrow \dots \longrightarrow x_k$. В [46] е показано, че

$$E\theta_k[\varphi] = (\varphi, u^{(k)}), \quad \text{където} \quad \theta_k[\varphi] = \frac{\varphi(x_0)}{\pi(x_0)} \sum_{j=0}^k W_j f(x_j),$$

$$W_0 = 1, \quad W_j = W_{j-1} \frac{k(x_{j-1}, x_j)}{p(x_{j-1}, x_j)}, \quad j = 1, \dots, k,$$

и съответната оценка по метод МК за $(\varphi, u^{(k)})$ е: $(\varphi, u^{(k)}) \approx \frac{1}{N} \sum_{n=1}^N \theta_k[\varphi]_n$. Следователно случайната величина $\theta_k[\varphi]$ може да се разгледа като оценка за търсената стойност (φ, u) при достатъчно голямо k с вероятностна грешка с порядък $\mathcal{O}(N^{-1/2})$, където N е броят на реализациите на веригата на Марков, а $\theta_k[\varphi]_n$ е стойността на $\theta_k[\varphi]$, получена върху n -тата траектория. Ако предположим [46]

$$p(x, x') = \frac{|k(x, x')|}{\int_{\Omega} |k(x, x')|dx'}, \quad \pi(x) = \frac{|\varphi(x)|}{\int_{\Omega} |\varphi(x)|dx},$$

тогава алгоритъмът се нарича почти оптимален Монте Карло алгоритъм (МАО) и това гарантира намаляване на дисперсията. За да се повиши изчислителната ефективност на алгоритъма трябва стохастичната грешка r_N да бъде приблизително равна на систематичната грешка r_k или да е изпълнено [7] $r_N = \mathcal{O}(r_k)$. Задачата за балансиране на грешките се трансформира в задача за получаване на оптимално съотношение между броя на реализациите N на случайната величина и средния брой стъпки k във всяка случайна траектория.

За оценка на вероятностната грешка е получено:

$$r_N \leq \frac{0.6745 \|f\|_{L_2} \|\varphi\|_{L_2}}{\sqrt{N} (1 - \|\mathcal{K}\|_{L_2})}.$$

За систематичната грешка получаваме следната оценка чрез неравенството на Коши-Шварц:

$$r_k = |(\varphi, u) - (\varphi, u^{(k)})| \leq \|\varphi\|_{L_2} \|u - u^{(k)}\|_{L_2} \leq \frac{\|\varphi\|_{L_2} \|f\|_{L_2} \|\mathcal{K}\|_{L_2}^{k+1}}{1 - \|\mathcal{K}\|_{L_2}}.$$

Подобреният алгоритъм е базиран на по-долните твърдения, които играят важна роля за изчислителната му сложност. Да допуснем, че

$$r_N \leq \frac{0.6745 \|\varphi\|_{L_2} \|f\|_{L_2}}{\sqrt{N} (1 - \|\mathcal{K}\|_{L_2})} \leq \frac{\delta}{2}, \quad r_k \leq \frac{\|\varphi\|_{L_2} \|f\|_{L_2} \|\mathcal{K}\|_{L_2}^{k+1}}{1 - \|\mathcal{K}\|_{L_2}} \leq \frac{\delta}{2}.$$

Приближаваме $k = k(N)$, като предполагаме, че $0.6745 \|f\|_{L_2} \frac{1}{\sqrt{N}} = \frac{\|\mathcal{K}\|_{L_2}^{k+1} \|f\|_{L_2}}{1 - \|\mathcal{K}\|_{L_2}}$

Теорема 4. (Теорема за балансираност) В Монте Карло алгоритъма за интегрални уравнения, базиран на балансиране на систематичната и стохастичната грешка, долните граници за N и k са:

$$N \geq \left(\frac{1.349 \|\varphi\|_{L_2} \|f\|_{L_2}}{\delta (1 - \|\mathcal{K}\|_{L_2})} \right)^2, \quad k \geq \frac{\ln \frac{\delta(1 - \|\mathcal{K}\|_{L_2})}{2 \|\varphi\|_{L_2} \|f\|_{L_2} \|\mathcal{K}\|_{L_2}}}{\ln \|\mathcal{K}\|_{L_2}}.$$

Можем да получим оптимално съотношение между двете грешки по два различни начина:

В Монте Карло алгоритъма за интегрални уравнения базиран на балансиране на систематичната и стохастичната грешки, ако N е близо до своята долна граница, тогава долната граница за k е:

$$k \geq \frac{\ln \frac{0.6745}{\|\mathcal{K}\|_{L_2} \sqrt{N}}}{\ln \|\mathcal{K}\|_{L_2}}.$$

Ако първо изберем k да бъде близо до своята долна граница, по аналогичен начин получаваме долната граница за N :

$$N \geq \frac{0.455}{\|\mathcal{K}\|_{L_2}^{2k+2}}.$$

Горните твърдения са от особено важно значение за изчислителната сложност на алгоритъма.

Числени експерименти

Първият пример има приложно значение в биологията. Разглеждаме уравнението, описващо популационен модел в биологията [15]:

$$u(x) = \int_{\Omega} k(x, x') u(x') dx' + f(x),$$

Таблица 10: Относителна грешка и изчислително време за първия пример с балансиране на грешката (различни преходни плътности).

δ	N	k	очаквана отн. грешка	експериментална(const) отн. грешка	време сек.	експериментална(МАО) отн. грешка	време сек.
0.037	5101	4	2.24e-02	4.12e-03	11	4.06e-03	7
0.025	11172	5	1.52e-02	1.45e-03	16	1.21e-03	9
0.014	35623	6	8.49e-03	4.572e-04	56	4.001e-04	34
0.0055	230809	7	3.34e-03	1.524e-04	424	9.881e-05	346

където $\Omega \equiv [0, 1]$, $k(x, x') = \frac{1}{3}e^x$, $f(x) = \frac{2}{3}e^x$, $\varphi(x) = \delta(x)$.

Точното решение е $u(x) = e^x$. Искаме да намерим решението в средата на интервала. Изчисляваме L_2 нормите: $\|\varphi\|_{L_2} = 1$, $\|\mathcal{K}\|_{L_2} = 0.3917$, $\|f\|_{L_2} = 1.1915$. Монте Карло алгоритъмът започва от $x_0 = 0.5$, така че точното решение е 1.6487, $\pi(x) = \delta(x)$. Направени са 20 алгоритмични стъпки.

В Таблица 10 са представени броя на реализациите N и броя на итерациите k в зависимост от предварително зададената точност δ според получените условия за балансираност на систематичната и стохастичната грешка. Очакваната или теоретична грешка се получава като разделим δ на точната стойност. Седмата и осма колона са за експерименталната относителна грешка с МАО алгоритъма, който сме конструирали и времето за пресмятане на функционала от решението, а пета и шеста колона са съответно за алгоритъма Монте Карло за интегрални уравнения с константни плътности. Веднага се вижда, че очакваната относителна грешка се потвърждава от получената експериментална грешка. Може да се отчете предимство на алгоритъма МАО в сравнение с алгоритъма с константи плътности.

Разглежданият втори пример има важно значение в невронните мрежи и изкуствения интелект. Разглеждам следното интегрално уравнение на Фредхолм от втори род, описващо процеса на обучение на невронни мрежи [20]:

$$u(x) = \int_{\Omega} k(x, x')u(x') dx' + f(x),$$

където $\Omega \equiv [-2, 2]$, $k(x, x') = \frac{0.055}{1+e^{-3x}} + 0.07$, $f(x) = 0.02(3x^2 + e^{-0.35x})$, $\varphi(x) = 0.7((x+1)^2 \cos(5x) + 20)$. Точното решение е 8.98635750518, като $\|\varphi\|_{L_2} = 27.7782$,

Таблица 11: Относителна грешка и изчислително време за втория пример с балансиране на грешката (различни преходни плътности).

δ	N	k	очаквана отн. грешка	експериментална(const) отн. грешка	време сек.	експериментална(МАО) отн. грешка	време сек.
0.2	3457	4	2.23e-02	9.45e-03	9	1.21e-02	23
0.1	13827	4	1.11e-02	1.13e-02	28	8.63e-03	46
0.05	55306	5	5.56e-03	1.77e-02	132	3.24e-03	222
0.028	176357	5	3.12e-03	1.76e-02	448	3.11e-03	540

$\|\mathcal{K}\|_{L_2} = 0.2001$, $\|f\|_{L_2} = 0.2510$. Резултатите са изложени в Таблица 11.

Лесно може да се види, че балансираният МАО алгоритъм дава много по-добри резултати от Монте Карло алгоритма с константни плътности за големи стойности на N и k . При малък брой реализации метода Монте Карло с константни плътности дава по-малка относителна грешка, но резултатите получени с МАО алгоритъма са по-близки до очакваната теоретична относителна грешка. Единствено при балансирания МАО алгоритъм експерименталната грешка потвърждава очакваната относителна грешка. При МАО алгоритъма изчислителното време е по-голямо, защото използваме метода на селекцията за моделиране на началната плътност. В този случай това не е необходимо за преходните плътности, тъй като ядрото на интегралното уравнение е функция само на една променлива и когато поискаме преходните плътности да са пропорционални на ядрото, получаваме, че те са константи. Разработеният балансиран МАО алгоритъм дава много по-добри резултати от балансирания метод Монте Карло с константни плътности при условие, че преходната плътност е различна от δ -функцията. Когато преходната плътност е δ -функцията двата метода дават близки резултати, но отново балансирания МАО е малко по-точен.

Интересно е да се види дали предложения алгоритъм може да се прилага за нелинейни интегрални уравнения с полиномиална нелинейност. Може да се очаква, че ако нелинейността не е много строга вместо да се използват функционали върху разклоняващи се вериги на Марков ([10]), може да се използва конструирания балансиран алгоритъм. Следващия пример е от основно значение при моделирането на процеси във физиката [7]. Разглеждаме следното

Таблица 12: Относителна грешка за третия пример с балансиране на грешката.

δ	N	k	очаквана отн. грешка	експериментална отн. грешка
0.03	1487	2	6.00e-02	8.56e-02
0.01	13381	3	2.00e-02	7.10e-02
0.005	53522	4	1.00e-02	6.02e-02
0.002	334512	5	4.00e-03	4.56e-02

интегрално уравнение с полиномиална нелинейност, описващо процеса на взаимодействие между две твърди физични тела:

$$u(x) = \int_{\Omega} \int_{\Omega} k(x, y, z) u(y) u(z) dy dz + f(x),$$

където $\Omega = \mathbf{E} \equiv [0, 1]$, $f(x) = c - \frac{x}{a_3}$, $\varphi(x) = \delta(x - x_0)$,

$$p(y, z) = \frac{6}{2a_2^2 - 3a_2 + 2} (a_2 y - z)^2, \quad k(x, y, z) = \frac{x(a_2 y - z)^2}{a_1}.$$

Едно достатъчно условие за сходимост на процеса е:

$$K_2 = \max_{x \in [0, 1]} \int_0^1 \int_0^1 |k(x, y, z)| dy dz = \frac{2a_2^2 - 3a_2 + 2}{6a_1} < \frac{1}{2}.$$

Интегралното уравнение има единствено регулярно решение $u(x) = c$, когато

$c = \pm \left(\frac{6a_1}{a_3(2a_2^2 - 3a_2 + 2)} \right)^{\frac{1}{2}}$. Избираме $a_1 = 11, a_2 = 4, a_3 = 12, c = 0.5$ и точното решение е $u(x) = 0.5$, $\|\mathcal{K}\|_{L_2} = 0.408$, $\|f\|_{L_2} = 0.459$, $\|\varphi\|_{L_2} = 1$. Резултатите са изложени в Таблица 12. Пропускаме метода Монте Карло с константи плътности, тъй като дава много лоши резултати за нелинейни интегрални уравнения.

За проблеми като този учените са доволни да имат грешка от порядъка на 5% или 10%. Оттук може да заключим, че конструираният почти оптимален метод Монте Карло за интегрални уравнения, базиран на балансиране на систематичната и стохастичната грешка, е приложим и за нелинейни интегрални уравнения.

Методи Монте Карло за линейни системи

Във втора секция на втора глава е конструиран и изследван подобрен метод Монте Карло за линейни системи (IWE), конструиран на базата на метода Монте Карло за линейни системи "случайно блуждаене по уравненията" (WE), описан в [13]. Направено е сравнение с рафинирания метод Монте Карло (RIMC), описан в [7]. Проведени са числени експерименти с различни примери на матрици от ниска и висока размерност. Методът може ефективно да се конкурира по бързина с вградените алгоритми в MATLAB и може да се приложи и при решаване на линейните алгебрични системи, получени след дискретизация на частните диференциални уравнения, описани в следващата глава.

Разглеждаме следната система от линейни алгебрични уравнения (СЛАУ): $Bx = f$, където $B = \{b_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}$ е дадена матрица; $f = (f_1, \dots, f_n)^t \in \mathbb{R}^{n \times 1}$ и $v = (v_1, \dots, v_n) \in \mathbb{R}^{1 \times n}$ са дадени вектори. Използваме матрицата $A = \{a_{ij}\}_{i,j=1}^n$, такава че $A = I - DB$, където D е диагоналната матрица $D = \text{diag}(d_1, \dots, d_n)$ и $d_i = \frac{\gamma}{b_{ii}}$, $i = 1, \dots, n$, и $\gamma \in (0, 1]$ е релаксационен параметър, чийто избор влияе върху скоростта на итерационния процес. Горната система може да се представи като система във вида $x = Ax + b$, където $b = Df$ и a_{ij} са реални числа. Да предположим, че B има диагонално преобладаване. Очевидно, ако B е матрица с преобладаващ главен диагонал, тогава елементите на матрицата A трябва да удовлетворяват следните условия: $\sum_{j=1}^n |a_{ij}| \leq 1$, $i = 1, \dots, n$.

Дефинираме верига на Марков T_k с $n + 1$ състояния $\alpha_1, \dots, \alpha_n, n + 1$, такава че $P(\alpha_{k+1} = j | \alpha_k = i) = |a_{i,j}|$, ако $i \neq n + 1$ и $P(\alpha_{k+1} = n + 1 | \alpha_k = n + 1) = 1$. Дефинираме вектора c , такъв че $c_i = b_i$, ако $1 \leq i \leq n$ и $c(n + 1) = 0$. Означаваме с $\tau = (\alpha_0, \alpha_1, \dots, \alpha_k, n + 1)$ случайната траектория, която започва в началното състояние $\alpha_0 < n + 1$ и минава през $(\alpha_1, \dots, \alpha_k)$, докато стигне до състоянието на поглъщане $\alpha_{k+1} = n + 1$. Вероятността да бъде осъществена траекторията τ е $P(\tau) = p_{\alpha_0} p_{\alpha_0 \alpha_1} \dots p_{\alpha_{k-1}, k} p_{\alpha_k}$. Използваме МАО алгоритъма [7] за вектора на началната вероятност $p = \{p_\alpha\}_{\alpha=1}^n$ и за матрицата на преходните вероятности $P = \{p_{\alpha\beta}\}_{\alpha,\beta=1}^n$, съответно. Теглата Q_α се определят чрез:

$$Q_m = Q_{m-1} \frac{a_{\alpha_{m-1}, \alpha_m}}{p_{\alpha_{m-1}, \alpha_m}}, \quad m = 1, \dots, k, \quad Q_0 = \frac{c_{\alpha_0}}{p_{\alpha_0}}. \quad (17)$$

Величината $X_\alpha(\tau)$ може да се представи като $X_\alpha(\tau) = c_\alpha + Q_k \frac{a_{\alpha k} \alpha}{p_{\alpha k}}$, $\alpha = 1, \dots, n$, взета с вероятност $P(\tau) = p_{\alpha_0} p_{\alpha_0 \alpha_1} \dots p_{\alpha_{k-1}, k} p_{\alpha_k}$. Използваме, че случайната величина $X_\alpha(\tau)$ е неизместена оценка за x_α , т.е. $E\{X_\alpha(\tau)\} = x_\alpha$ [13].

Описание на алгоритъма

Идеята на алгоритъма е да се пресметнат приближено стойностите на всички компоненти на решението (разгледани като начални състояния). Първоначалното уравнение се избира случайно и равномерно сред първите n уравнения. След това, за всяко състояние i дефинираме общ резултат $S(i)$ с общ брой посещения $NV(i)$, които се модифицират, веднага щом състоянието i е посетено по време на случайното блуждаене. За дадена траектория запазваме посетените състояния в списък l , за да пресметнем по-лесно приноса им към резултата на посетените състояния. Накрая решението x_i се приближава с общия резултат $S(i)$, разделен на общия брой посещения $NV(i)$. Ако състоянието никога не е посетено, то $NV(i) = 1$. Ако състоянието е посетено повече от веднъж в рамките на една траектория, обикновено пазим приноса на първото посещение, за да се редуцира дисперсията. Общият брой траектории N се избира близко до размерността n на системата.

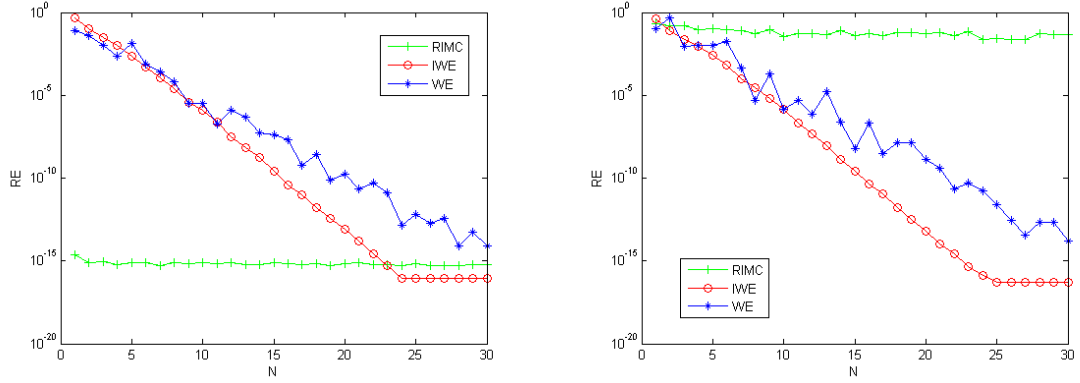
При ускоряване на сходимостта на алгоритмите се използва последователният метод Монте Карло (Sequential Monte Carlo, SMC) за линейни системи, който е реализиран от John Halton през 1960-те [22] и е развит в по-скорошни публикации [24, 23].

Подобреният алгоритъм за намиране на една компонента на решението се състои в избиране на специален избор на релаксационния параметър, което води до балансиране на итерационната матрица и до по-малък брой изчислителни операции спрямо WE, което води до намаляване на изчислителното време.

Числени експерименти

По-долу ще направим сравнения за линейни системи с различна размерност между стандартния рафиниран Монте Карло метод (RIMC) [7], оригиналният Монте Карло метод „случайно блуждаене по уравненията“ (WE) [13] и конструираният подобрен метод Монте Карло за линейни системи, означен с IWE, базиран на метода „случайно блуждаене по уравненията“.

Пример 1. Трябва да намерим решенията x_1 и x_2 дефинирани чрез линейната система $Bx = f$, където матриците B и векторите f_1 и f_2 са:



Фигура 6: Теглови резидуал на решението за матрицата 7×7 : (a) за $\mathbf{x}_1$; (b) за $\mathbf{x}_2$.

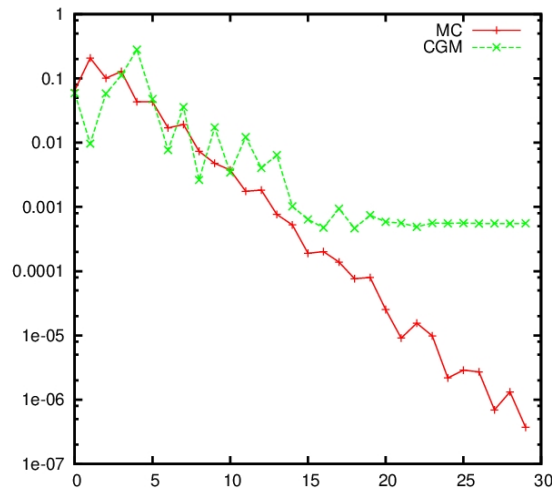
$$A = \begin{pmatrix} 5 & -1 & -1 & 0 & 0 & -1 & -1 \\ -1 & 5 & -1 & -1 & 0 & 0 & -1 \\ -1 & -1 & 5 & -1 & -1 & 0 & 0 \\ 0 & -1 & -1 & 5 & -1 & -1 & 0 \\ 0 & 0 & -1 & -1 & 5 & -1 & -1 \\ -1 & 0 & 0 & -1 & -1 & 5 & -1 \\ -1 & -1 & 0 & 0 & -1 & -1 & 5 \end{pmatrix}, \quad f_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad f_2 = \begin{pmatrix} 4 \\ -2 \\ -1 \\ 0 \\ -1 \\ -2 \\ 4 \end{pmatrix}. \quad (18)$$

Решенията са

$$x_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (19)$$

Пример 2. Нека B е симетричната положително определена матрицата NOS4 от колекцията на Harwell-Boeing [55], а $f \in \mathbb{R}^{100}$, $b_i = 1$, $i = 1, \dots, 100$. Тази матрица е взета от приложения, свързани с апроксимация по метода на крайните елементи на модел, описващ гредова структура в конструктивната механика. Матрицата NOS4 има точно 597 ненулеви елемента, 100 ненулеви елемента по

главния диагонал и 247 под и над главния диагонал, или средно 5.9 ненулеви елемента във всяка колона и всеки ред. За матрицата NOS4 е показано, че методът Монте Карло WE е по-добър от метода на спрегнатия градиент PCG – виж Фиг. 7. Успехът на PCG се дължи на факта, че той е приложим за оптимални подпространства на Крилов. Методът Монте Карло също е приложим за оптимални подпространства на Крилов. Разликата е, че докато при PCG трябва да се решава оптимизационната задача, и ако има локален минимум до глобалния минимум, сходимостта на процеса може да клони към локалния минимум, като точно такъв е и случая за матрицата NOS4. Методът Монте Карло за линейни системи е независим от такава оптимизационна процедура, и затова резултатите с Монте Карло са по-добри. Този факт не може да се обобщи, понеже успехът на метода зависи от конкретния функционал, който трябва да се минимизира при PCG.



Фигура 7: Сравнение на метода Монте Карло IWE и метода PCG за матрицата NOS4 от колекцията Harwell-Boeing.

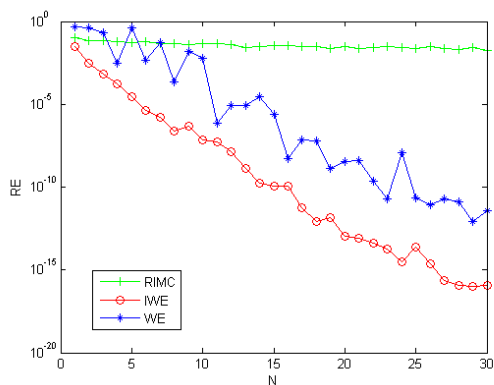
Пример 3. Нека B е плътна положителна матрица 5000×5000 с елементи в $[0,1]$, и $f \in \mathbb{R}^{5000}$, $b_i = 1$, $i = 1, \dots, 5000$.

Матриците B и дясната част f са нормирани, за да се ускори сходимостта на стохастичния процес. Избрани са специални стойности на релаксационния параметър γ . Числените експерименти показват, че това води до балансиране на итерационната матрица A . Може да се види, че в 7-мерния случай, разликата

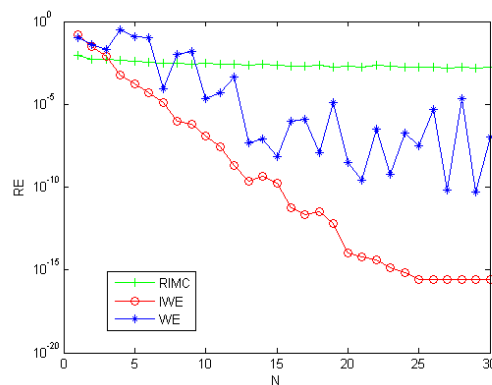
Таблица 13: Теглови резидуал на решението за матриците $B \in \mathbb{R}^{100 \times 100}$ и $B \in \mathbb{R}^{5000 \times 5000}$

	100 × 100						5000 × 5000					
N	RIMC	t	WE	t	IWE	t	RIMC	t	WE	t	IWE	t
2	7.25e-02	0.05	4.18e-01	0.84	3.02e-03	0.08	5.43e-03	10.05	4.30e-02	3.95	2.93e-02	0.15
5	5.45e-02	0.22	4.14e-01	2.37	3.07e-05	0.24	3.87e-03	60.2	1.21e-01	13.3	1.82e-04	0.9
10	4.32e-02	0.56	5.94e-03	5.31	7.46e-08	0.61	2.86e-03	130.5	2.30e-05	32.3	1.23e-07	2.4
15	3.52e-02	0.78	2.41e-06	9.1	1.21e-10	0.89	2.36e-03	310.7	6.48e-09	67.8	1.83e-10	5.1
20	3.19e-02	1.11	3.33e-09	13.5	1.02e-13	1.13	1.94e-03	811	3.20e-09	171.5	1.05e-14	11.1
30	1.83e-02	2.15	3.66e-12	24.6	1.11e-16	1.92	1.70e-03	2135	1.12e-07	418.6	2.48e-16	25.2

между WE и IWE при зададен брой итерации е 2-3 порядъка за $N > 15$. Трябва да се спомене, че сходимостта на рафинирания метод Монте Карло е много бавна, освен в случая на тривиалното решение за x_1 от Пример 1 – виж Фиг. 6. За 100-мерния случай на разредената матрица NOS4, подобреният алгоритъм IWE постига много по-добри резултати от WE и е почти 5-6 пъти по-бърз – виж Таблица 4.2. Поведението на Монте Карло алгоритъма не зависи от плътността на матрицата. Предимствата на алгоритъма са в сила и за разредени матрици. Разликата в относителната грешка за фиксиран брой итерации е 3-5 порядъка – виж Фиг. 7. Следователно IWE има предимство по отношение на по-добра точност с увеличаване на размерността. Също предимството на новия метод може ясно да се види по отношение на по-малко изчислително време. За по-голяма размерност, предимствата на IWE в сравнение с WE е още по-добре изразено. След 30 итерации WE постига точност близо до 10^{-10} , докато IWE постига точност 10^{-16} . Изчислителното време за IWE е 15 пъти по-добре в сравнение с WE – виж Таблица 20. Специалният избор на релаксационния параметър води до балансиране на итерационната матрица и това и това обяснява защо IWE е по-бърз. Предимството на подобрения Монте Карло алгоритъм е особено явно за големи размерности. Експериментите показват, че за големи размерности подобрението води до по-малка относителна грешка за по-малък брой SMC итерации за IWE.



(a)



(b)

Фигура 8: Теглови резидуал на решението за матрицата: (a) 100×100 ; (b) 5000×5000 .

Нови числени методи с висок ред на точност за модели в екологията

В последната глава се конструират нови компактни диференчни схеми (схеми с минимален шаблон) за слабо свързани системи от параболични частни диференциални уравнения (ЧДУ) с нелинейни химични реакции. Осъществено е приложение в едномерния и двумерния случай, както и при модел на далечен пренос на замърсители във въздуха на базата на Unified Danish Eulerian Model, UNI-DEM [52, 53]. Разглежда се подмодел, в който участват 10 химични замърсителя. Разглежда се случая на точно аналитично решение и когато няма точно решение, което съответства на реалната задача. Разглежда се правоъгълна област с размери от 500 км за едно денонощие. Разглежда се също и опростен модел на химични процеси в атмосферата по цикъла на Чапман на базата на три вещества. Направено е сравнение между два различни подхода за получаване на схеми от четвърти ред на точност- компактна диференчна схема и стандартна схема с повишен порядък с екстраполация по Ричардсон. Получена е схема с шести ред на точност по пространството като върху компактната схема е приложена екстраполация по Ричардсон.

Резултатите, представени в тази секция, са публикувани в [12].

Постановка на задачата

Тази глава е посветена на конструирането на нови компактни диференчни схеми (CFDS) с висок ред на точност за полулинейни параболични системи. Задачи от модели на преноси на замърсители във въздуха със свързани нелинейни реакции, са от основен интерес, а именно:

$$\frac{\partial u_l}{\partial t} - K \Delta u_l + \mathbf{b}_l \nabla u_l = R_l(x, y, u_1, \dots, u_L), \quad (x, y, t) \in \Omega \times (0, T], \quad (20)$$

$$\mathbf{u} = 0, \quad (x, y, t) \in \partial\Omega \times (0, T], \quad (21)$$

$$\mathbf{u} = \mathbf{u}_0(x, y), \quad (x, y) \in \Omega, \quad (22)$$

където $\mathbf{u} = (u_1, u_2, \dots, u_L)$, $u_l = u_l(x, y, t)$, $l = 1, \dots, L$ са концентрации на L химични вещества (замърсители) и $K > 0$ са коефициенти на дифузията и $\Omega \in R^2$ е ограничена област. Предположението за константите $K := K_x = K_y$ не е ограничение за развитието на разработения числен подход. Това съответства на физически модел, описан в [19, 29].

Основната цел е приложението и числените експерименти на по-горе споменатите диференчни апроксимации за следната реално физична параболична транспортна система, описана в [19]. Следвайки [19, 29, 54] адвекцията в (20) е

$$\mathbf{b}_l \cdot \nabla u_l = \mu(y - y_c) \frac{\partial u_l}{\partial x} + \mu(x_c - x) \frac{\partial u_l}{\partial y},$$

където $x \in (0, X)$, $y \in (0, Y)$, $x_c = X/2$, $y_c = Y/2$. Нелинейната химична част на модела е следната:

$$\begin{aligned} R_1(u_1, \dots, u_{10}) &= k_5 u_2 - (k_6 u_5 + k_4 u_7 + k_3 u_8) u_1, \\ R_2(u_1, \dots, u_{10}) &= (k_6 u_5 + k_4 u_7 + k_3 u_8) u_1 - (k_5 + k_9 u_9) u_2, \\ R_3(u_1, \dots, u_{10}) &= -k_1 u_3 u_9, \\ R_4(u_1, \dots, u_{10}) &= 2k_1 u_3 u_9 + k_3 u_1 u_8 - k_2 u_4, \\ R_5(u_1, \dots, u_{10}) &= k_2 u_5 \\ R_6(u_1, \dots, u_{10}) &= k_9 u_2 u_9, \\ R_7(u_1, \dots, u_{10}) &= 2k_2 u_4 + k_3 u_1 u_8 + k_{10} u_9 - k_4 u_1 u_7, \\ R_8(u_1, \dots, u_{10}) &= 4k_1 u_3 u_9 - k_3 u_1 u_8, \\ R_9(u_1, \dots, u_{10}) &= k_4 u_1 u_7 + 2k_8 u_{10} - (k_1 u_3 - k_9 u_2 + k_{10}) u_9, \\ R_{10}(u_1, \dots, u_{10}) &= k_7 u_5 - k_8 u_{10}. \end{aligned} \quad (23)$$

Таблица 14: Химични реакции на модела

1	$HC + OH \rightarrow 4RO_2 + 2ALD$	6	$NO + O_3 \rightarrow NO_2 + O_2$
2	$ALD + h\nu \rightarrow 2HO_2 + CO$	7	$O_3 + h\nu \rightarrow O_2 + O(^1D)$
3	$RO_2 + NO \rightarrow NO_2 + ALD + HO_2$	8	$O(^1D) + H_2O \rightarrow 2OH$
4	$NO + HO_2 \rightarrow NO_2 + OH$	9	$NO_2 + OH \rightarrow HNO_3$
5	$NO_2 + h\nu \rightarrow NO + O_3$	10	$CO + OH \rightarrow CO_2 + HO_2$

Таблица 15: Скоростни константи на химичните реакции

k_1	$6.0e - 12$	k_6	$1.6e - 14$
k_2	$7.8e - 05. \exp(-0.87/\cos \theta)$	k_7	$1.6e - 04. \exp(-1.9/\cos \theta)$
k_3	$8.0e - 12$	k_8	$2.3e - 10$
k_4	$8.0e - 12$	k_9	$1.0e - 11$
k_5	$1.0e - 02. \exp(-0.39/\cos \theta)$	k_{10}	$2.9e - 13$

Химичната част на модела е дадена в Таблица 14 за пълнота. Скоростните константи на химичните реакции са дадени в Таблица 15. Някои от коефициентите принадлежат на фотохимични реакции (тези с член $h\nu$), което значи, че тези реакции зависят от светлината, по точно от позицията на Слънцето спрямо хоризонта: в k_2 , k_5 и k_7 ъгълът θ означава слънчевия зенитен ъгъл, който е ъгъла на Слънцето, измерен по вертикала. Химичните вещества, включени в опростените реакции, са дадени в Таблица 16.

Таблица 16: Химичните вещества на модела

u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}
NO	NO_2	HC	ALD	O_3	HNO_3	HO_2	RO_2	OH	$O(^1D)$

От практическа и математическа гледна точка, интерес представлява съществуването и качественото поведение (неотрицателността) на решението на задачата (20)-(22). Коректността на начално-граничната задача за по-обща система

от (20) е получена в [41]. Ще допускаме до края съществуване и единственост на класическо решение на задачата (20)-(22), което означава, че функцията принадлежи на $C([0, T] \times \bar{\Omega}) \cap C^1((0, T); C(\bar{\Omega})) \cap (C(0, T); C^2(\bar{\Omega}))$ и удовлетворява (20)-(22) поточно.

Разглеждаме система, описваща химични концентрации, неотрицателността на решението трябва да се запази. Доказано е [41], че ако

1. $\mathbf{u}_0(x, y) \geq 0$;
2. $R_l(x, y, \mathbf{u})$, $l = 1, \dots, L$ е непрекъсната по Липшиц по отношение на концентрациите $u_1, u_2, \dots, u_L$ и удовлетворява неравенството $R_l(x, y, \mathbf{u}) \geq 0$, където $u_l = 0$, и $\mathbf{u} \in R_+^L \equiv \{u_k \geq 0, k = 1, \dots, L\}$,

тогава $\mathbf{u} \geq 0$ за всяко $(x, y) \in \Omega$ и $t \in [0, T]$.

Лесно е да се провери, че химичните реакции $R_l(u_1, u_2, \dots, u_{10})$, $l = 1, \dots, 10$ зададени от (23) удовлетворяват точка 2. и решението на задачата (20)-(22) с (23) е неотрицателно по времето $t > 0$, ако началните данни $\mathbf{u}_0(x, y) \geq 0$.

В последните години се наблюдава нов и все по-нарастващ интерес към компактните диференчни схеми с висок ред на точност за решаване на ЧДУ [42]. Напоследък, голямо усилие се съсредоточава в развитието на компактни диференчни схеми с висок ред на точност, които използват само възлите на мрежата, съседни на централния възел. Компактните схеми, предложени от Крайс и Олигер [21] използват подобен шаблон, но изискват тридиагонално или петдиагонално обръщане. Друга идея, която се използва, е да се работи върху диференциалните уравнения, така че да се изразят производните от по-висок ред (трети и четвърти) в локалната грешка на апроксимация (LTE) [49].

Компактни диференчни схеми (CFDS) в двумерния случай

В тази секция се описва построяването на компактни диференчни схеми (CFDS) в двумерния случай за системата от две уравнения:

$$\frac{\partial u}{\partial t} - a(x, y) \frac{\partial^2 u}{\partial x^2} - b(x, y) \frac{\partial^2 u}{\partial y^2} + c(x, y) \frac{\partial u}{\partial x} + d(x, y) \frac{\partial u}{\partial y} = r(x, y, t, u, v), \quad (24a)$$

$$\frac{\partial v}{\partial t} - e(x, y) \frac{\partial^2 v}{\partial x^2} - f(x, y) \frac{\partial^2 v}{\partial y^2} + g(x, y) \frac{\partial v}{\partial x} + h(x, y) \frac{\partial v}{\partial y} = s(x, y, t, u, v), \quad (24b)$$

дефинирана в областта $Q_T = \Omega \times (0, T)$, където $\Omega \subset R^2$ е ограничена област с Липшицова граница. Нелинейните функции r и s са достатъчно гладки по тех-

ните аргументи. Коефициентите $a(x, y)$, $b(x, y)$, $e(x, y)$ и $f(x, y)$ са положителни в Ω . Предполагаме гранични условия на Дирихле

$$u(x, y, t) = \bar{\phi}(x, y, t), \quad v(x, y, t) = \bar{\bar{\phi}}(x, y, t), \quad (x, y, t) \in \partial\Omega \times (0, T] \quad (25)$$

и начални условия

$$u(x, y, 0) = \bar{\psi}(x, y), \quad v(x, y, 0) = \bar{\bar{\psi}}(x, y), \quad (x, y) \in \Omega, \quad (26)$$

където $\bar{\phi}$, $\bar{\bar{\phi}}$, $\bar{\psi}$ и $\bar{\bar{\psi}}$ са дадени и гладки и съвместимост с началните и гранични условия е осигурена. Нека за простота областта Ω е правоъгълник $\Omega = [0, X] \times [0, Y]$. Да въведем равномерните мрежи $\bar{\omega}_{h,x} = \{x_i = ih_x, \quad i = 0, 1, \dots, N_x, \quad h_x = X/N_x\}$, $\bar{\omega}_{h,y} = \{y_j = jh_y, \quad j = 0, 1, \dots, N_y, \quad h_y = Y/N_y\}$ и тогава $\bar{\Omega}_h = \omega_{h,x} \times \omega_{h,y}$, $\bar{\Omega}_h = \Omega_h \cup \partial\Omega_h$, където Ω_h се състои от всички вътрешни мрежови точки, а $\partial\Omega_h$ - от всички гранични мрежови точки.

Дискретизация по пространството. Идеята е, че в локалната грешка на апроксимация членовете от втори порядък съдържат трета и четвърта производна на решението, и за да останем на същия шаблон, трябва да ги изразим чрез производни от по-нисък ред, като диференцираме двукратно диференциалните уравнения. Следователно, за да елиминираме членовете от ред $\mathcal{O}(h_x^2 + h_y^2)$, диференцираме уравнението (24a) два пъти по x и получаваме израз за $\frac{\partial^3 u}{\partial x^3}$, $\frac{\partial^4 u}{\partial x^4}$, и два пъти по y за $\frac{\partial^3 u}{\partial y^3}$, $\frac{\partial^4 u}{\partial y^4}$. В подробности тази процедура е описана в [12, 49].

Получаваме следната система ОДУ

$$\bar{Q} \frac{d}{dt} U^h + \bar{P} U^h = \bar{Q} R + \bar{\Phi}, \quad t \in (0, T], \quad (27)$$

$$\bar{\bar{Q}} \frac{d}{dt} V^h + \bar{\bar{P}} V^h = \bar{\bar{Q}} S + \bar{\bar{\Phi}}, \quad (28)$$

с начални условия $U^h(0)$ и $V^h(0)$, получени от $\bar{\psi}$ и $\bar{\bar{\psi}}$ за $(i, j) \in \Omega_h$ след пренареждането. В системата (27), (28) матрицата $\bar{P}$ (аналогично $\bar{\bar{P}}$) е $(N_y - 1) \times (N_y - 1)$ блочно тридиагонална матрица $\bar{P} = \text{tridiag}(\bar{P}_{k,k-1}, \bar{P}_{k,k}, \bar{P}_{k,k+1})$ и $\bar{P}_{k,l}$, $l = k-1, k, k+1$ са също тридиагонални матрици от ред $(N_x - 1) \times (N_x - 1)$. Тогава елементите на $\bar{P}_{k,l}$ са

$$\bar{P}_{k,l} = \text{tridiag}(p_{k,2:N_x-1}^{(-1,\varepsilon)}, p_{k,2:N_x}^{(0,\varepsilon)}, p_{k,1:N_x-2}^{(1,\varepsilon)}) \quad l = k + \varepsilon, \quad \varepsilon = 0, \pm 1. \quad (29)$$

Елементите на $\bar{Q}_{k,l}$ (аналогично $\bar{\bar{Q}}$) са

$$\bar{Q}_{k,l} = \text{tridiag}(q_{k,2:N_x-1}^{(-1,\varepsilon)}, q_{k,2:N_x}^{(0,\varepsilon)}, q_{k,1:N_x-2}^{(1,\varepsilon)}) \quad l = k + \varepsilon, \quad \varepsilon = 0, \pm 1 \quad (30)$$

със забележката, че за $\varepsilon = \pm 1$ матриците $\overline{Q}_{k,l}$ са диагонални (вместо тридиагонални) матрици. Векторите $\overline{\Phi}$ и $\overline{\overline{\Phi}}$ са свързани с граничните функции и също зависят от времето t .

Дискретизация по времето. За дискретизация по времето е използвана схема с тегла с $\theta = 1/2$. Тогава пълната дискретизация по Кранк-Никълсън за (27), (28), е както следва:

$$\begin{aligned}\overline{Q}\frac{U^{n+1} - U^n}{\tau} + \overline{P}U^{n,\theta} &= \overline{Q}R^{n,\theta} + \overline{\Phi}^{n,\theta}, \quad n = 1, \dots, N-1, \\ \overline{\overline{Q}}\frac{V^{n+1} - V^n}{\tau} + \overline{\overline{P}}V^{n,\theta} &= \overline{\overline{Q}}S^{n,\theta} + \overline{\overline{\Phi}}^{n,\theta}, \quad n = 1, \dots, N-1.\end{aligned}\quad (31)$$

Системата (31) е пренаписана във формата $\Upsilon(W) = 0$, където $W = [U^T, V^T]^T$ е вектор с дължина $2(N_x - 1)(N_y - 1)$. Поставяме W^{n+1}_0 като начално приближение на новия слой по времето $t = t_{n+1}$ да бъде решението на стария слой по времето $t = t_n$. Тогава за да намерим решението на слоя $t = t_{n+1}$ итерационен процес с подходящо стоп правило е използван:

$$\begin{cases} \Upsilon'(W^{n+1}) \overset{k}{\Delta} = -\Upsilon(W^{n+1}), \\ W^{n+1}_{k+1} = W^{n+1}_k + \overset{k}{\Delta}. \end{cases}\quad (32)$$

Тук $\overset{k}{\Delta}$ е вектора на нарастванията и матрицата на Якобиана $\Upsilon'(W^{n+1})$ за $\theta = 1/2$ е

$$\Upsilon'(W^{n+1}) = \begin{pmatrix} \frac{1}{\tau}\overline{Q}I + \frac{1}{2}\overline{P} - \frac{1}{2}\overline{Q}\frac{\partial R}{\partial U} & \frac{1}{2}\overline{Q}\frac{\partial R}{\partial V} \\ \frac{1}{2}\overline{\overline{Q}}\frac{\partial S}{\partial U} & \frac{1}{\tau}\overline{\overline{Q}}I + \frac{1}{2}\overline{\overline{P}} - \frac{1}{2}\overline{\overline{Q}}\frac{\partial S}{\partial V} \end{pmatrix} \Big|_{(U,V)=(\overset{k}{U}, \overset{k}{V})}. \quad (33)$$

В числените експерименти, за да решим първото уравнение в (32), който е линейна система от $2(N_x - 1)(N_y - 1)$ уравнения, използваме така наречения „неточен метод“ на Нютон (inexact Newton) [6], т.е. решаваме системата приближено с вградената MatLab функция bicgstab(1) (biconjugate gradients stabilized (1) method), който дава най-добрите резултати за числените експерименти по отношение на брой вътрешни операции и изчислително време. Може да се използва и методът Монте Карло за линейни системи, но е значително по-бавен за разглежданите задачи.

Екстраполация на Ричардсон

Друг начин за получаване на диференчни схеми с висок ред на точност, е да се използва метода на екстраполация на Ричардсон (RE). Главната идея [36] е да се реши диференчната схема на две или повече вложени мрежи и след това да се комбинират получените числени решения с подходящи тегла. Да допуснем, че $h_x = h_y = h$ и за численото решение на n -тия слой по времето, следният израз е верен:

$$U_h^\tau = U_{(i,j)}^n = u(x_i, y_j, t^n) + C_1 h^\sigma + \chi(h, \tau), \quad (x_i, y_j, t_n) \in \Omega_{h,\tau}, \quad (34)$$

където функцията $\chi(h, \tau)$ е остатъчен член и константата C_1 не зависи от h_x , h_y и τ . Ако искаме да елиминираме члена $C_1 h^\sigma$, се правят следните стъпки:

- решаваме диференчната схема на две последователни мрежи: груба $\Omega_{h,\tau}$ и фина $\Omega_{h/2,\tau}$ и нека съответстващите числени решения да бъдат U_h^τ и $U_{h/2}^\tau$;
- намираме теглата γ_1 и γ_2 от системата

$$\begin{aligned} \gamma_1 + \gamma_2 &= 1 \\ \gamma_1 + \frac{\gamma_2}{2^\sigma} &= 0 \end{aligned} \quad (35)$$

- получаваме ново числено решение на грубата мрежа

$$U_{extr} = \gamma_1 U_h^\tau + \gamma_2 U_{h/2}^\tau \quad (x_i, y_j, t_n) \in \Omega_{h,\tau}.$$

От (35) имаме за случая на стандартна схема (CDS) на Кранк-Никълсън ($\sigma = 2$), че коефициентите на екстраполацията по Ричардсон са

$$\gamma_1 = -1/3 \quad \gamma_2 = 4/3. \quad (36)$$

За случая на компактна схема (CFDS) и екстраполация по Ричардсон ($\sigma = 4$) съответните теглови коефициенти са

$$\gamma_1 = -1/15 \quad \gamma_2 = 16/15. \quad (37)$$

Числени експерименти за модел в екологията - системата ЧДУ от 10 уравнения

В тази секция се разглеждат два числени експеримента за потвърждение на теоретичните резултати. Първата задача (Пример1) е с точно решение, а втората (Пример 2) е опростена двумерна моделна задача с пренос на замърсители, описан в [19].

Пример 1 (известно аналитично решение)

Тук разглеждаме проблем, малко по различен от проблема (20):

$$\frac{\partial u_l}{\partial t} - K \Delta u_l + \mathbf{b}_l \cdot \nabla u_l = R_l(x, y, \mathbf{u}) + \xi_l(x, y, t), (x, y, t) \in \Omega \times (0, T]. \quad (38)$$

Функциите ξ_l , $l = 1, \dots, 10$, и началните и гранични условия са избрани, така че точното решение да бъде

$$u_l = \exp(-t/T) \sin\left(\frac{\pi x}{X}\right) \sin\left(\frac{\pi y}{Y}\right), \quad l = 1, \dots, 10, \quad (x, y, t) \in \partial\Omega \times [0, T]. \quad (39)$$

Другите параметри са, както следва: $X = Y = 500$, $T = 1440$, $\mu = 2\pi/(60T)$, $K = 1.8$.

За l -тата субстанция с $error_{M,l}$ означаваме грешката (разликата между точното и приближеното решение) в равномерна норма, получена на последния слой по времето $t_N = T$ за брой подинтервали по времето $M_x = M_y = M$:

$$error_{M,l} = \max_{i,j \in \bar{\Omega}_h} \|u_l(x_i, y_j, t_N) - u_l^h(i, j, N)\|.$$

Отношението между грешките, получени на две последователни мрежи (обикновено удвоени) е означено с *ratio*:

$$ratio = ratio_{M,l/2M,l} =: error_{M,l}/error_{2M,l}.$$

В таблиците са представени анализи със съгъстяване на мрежата с използването на CDS и CFDS, CDS и CFDS с (RE) по пространството и CDS и CFDS с (RE) по пространството и времето. Резултатите потвърждават теоретичния ред на сходимост за двата числени метода. От сравняването на CPU времето за Таблицы 17, 18 и 19 следва предимство на използването на RE едновременно по времето и пространството за получаването на по-малки грешки за по-кратко време. Предимството на CFDSRE е също очевидно. Фиг. 9 показва грешката в

Таблица 17: Сравнение на грешката за CDS и CFDS за **Пример 1**

CDS, $O(h^2 + \tau^2)$						CFDS, $O(h^4 + \tau^2)$					
M_x	M_y	N	$error_M$	$ratio$	CPU	M_x	M_y	N	$error_M$	$ratio$	CPU
4	4	4	5.702 e-03	-	0.58	4	4	4	5.875 e-03	-	0.72
8	8	8	1.449 e-03	3.94	1.82	8	8	16	3.595 e-04	16.34	3.04
16	16	16	3.637 e-04	3.99	14.42	16	16	64	2.232 e-05	16.11	29.74
32	32	32	9.102 e-05	4.001	143.7	32	32	256	1.392 e-06	16.03	1076
64	64	64	2.276 e-05	4.00	3959	64	64	1024	8.698 e-08	16.003	60907
128	128	128	5.691 e-06	4.00	32709	128	128	4096	5.436 e-09	16.0001	720477

Таблица 18: Грешка в равномерна норма за **Пример 1** за CDS и CFDS в RE по пространството

CDS с RE по пространството, $O(h^4 + \tau^2)$						CFDS с RE по пространството, $O(h^6 + \tau^2)$					
M_x	M_y	N	err_N	ratio	CPU	M_x	M_y	N	err_N	ratio	CPU
4	4	4	5.677 e-03	-	1.34	4	4	4	5.711 e-03	-	1.38
8	8	16	3.545 e-04	16.014	16.17	8	8	32	8.912 e-05	64.087	17.45
16	16	64	2.216 e-05	15.997	544	16	16	256	1.392 e-06	64.022	1497
32	32	256	1.385 e-06	16.001	3055	32	32	2048	2.1757 e-08	63.989	23390

максимална норма за *Пример 1* (а) със CDS и RE по пространството и времето за $M_x = M_y = 16$, $N = 16$; (б) с CFDS и RE по пространството и времето за $M_x = M_y = 16$, $N = 64$ и е в съответствие с резултатите в Таблица 19.

Пример 2 (без точно решение)

В този случай е разгледан по-реалистичен вариант на задачата (20)-(22) със следните параметри на областта: областта е квадрат $\Omega = [0, 500]^2$ с дължина 500 km, дължината на интервала по времето $[0, T]$ е 1440 min и броят на уравненията е $L = 10$. Началните условия на първия слой по времето $t = 0$ са константни функции

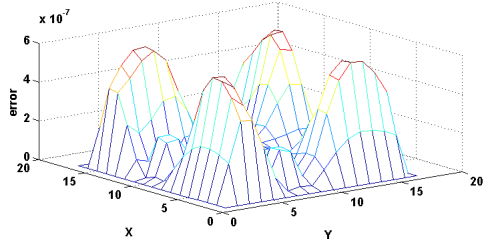
$$\mathbf{u}_0(x, y) = (10^3, 10^3, 10^3, 5.10^3, 5.10^3, 10^2, 10^{-2}, 10^{-2}, 10^{-3}, 10^{-11}),$$

измерени в mol/km^3 и граничните условия са избрани периодични функции: γ_i които имат вида

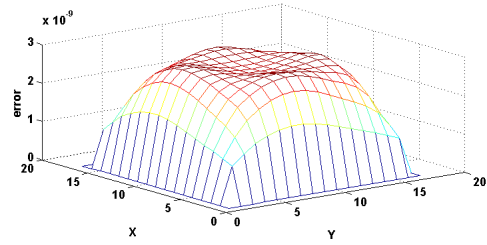
$$\gamma_l(t) = const_l(\sin(t/C) + 2),$$

Таблица 19: Грешка в равномерна норма за **Пример 1** за CDS и CFDS в RE по пространството и времето

CDS с RE, $O(h^4 + \tau^4)$						CFDS с RE, $O(h^6 + \tau^4)$					
M_x	M_y	N	err_N	ratio	CPU	M_1	M_y	N	err_N	ratio	CPU
4	4	4	5.649 e-05	-	6.73	4	4	4	8.476 e-06	-	3.36
8	8	8	9.722 e-06	5.81	18.71	8	8	16	1.748 e-07	48.49	30.26
16	16	16	5.989 e-07	16.23	194.81	16	16	64	2.847 e-09	61.39	1276
32	32	32	3.715 e-08	16.12	4594	32	32	256	4.529 e-11	62.86	66991
64	64	64	2.171 e-09	16.03	37101	64	64	1024	7.086 e-13	63.91	790800



(a)



(b)

Фигура 9: Грешката в максимална норма за *Пример 1*: (a) CDS с RE по пространството и времето за $M_x = M_y = 16$, $N = 16$; (b) CFDS с RE по пространството и времето за $M_x = M_y = 16$, $N = 64$

където C е константа и константите $const_l$, $l = 1, \dots, L$ са избрани по такъв начин, че да е осигурена съгласуваност на началните и граничните условия. Коефициентът на дифузия е избран да бъде $K = 1.8 km^2/min$ и коефициентът $\mu = 2\pi/(60 * T)$. В този случай няма аналитично решение. Един възможен начин за пресмятане на реда на сходимост е методът на Рунге на три вложени мрежи. Тук използваме друга идея. Като „точно” решение вземаме решението, получено с най-малките размери на мрежата по пространството. В таблицата това решение е означено с удебелен шрифт. Следим реда на сходимост означен с $order$ и пресметнат с помощта по формулата: $order = \log_2(ratio)$ когато удвояваме броя на мрежовите точки и в другия случай $order = \log(error_{M',l}/error_{M'',l})/\log(M''/M')$, където M' и M'' са два последователни броя мрежови точки по пространството в анализите на съгъстванията на мре-

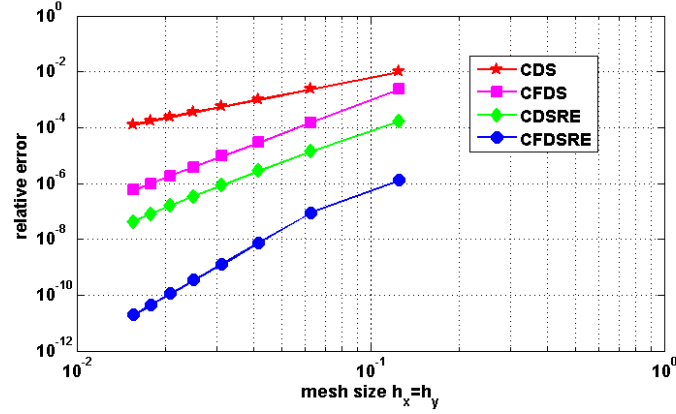
Таблица 20: Числените стойности, относителната грешка и реда на сходимост за **Пример 2** за CFDS за централния възел $(x, y) = (X/2, Y/2)$ със слоеве по времето $N = 256$ за u_1 и u_5

U_1					U_5				
M_x	M_y	числ. ст.	отн. гр.	ред	M_x	M_y	числ. ст.	отн. гр.	ред
8	8	2000.6332968	2.273 e-03	-	8	8	4580.1540334	2.417 e-03	-
16	16	1996.1958272	1.495 e-04	3.988	16	16	4569.7951222	1.495 e-04	4.014
24	24	1995.9567369	2.972 e-05	3.984	24	24	4569.2477794	2.972 e-05	3.984
32	32	1995.9162197	9.419 e-06	3.994	32	32	4569.1550256	9.420 e-06	3.994
40	40	1995.9051188	3.858 e-06	4.000	40	40	4569.1296128	3.858 e-06	4.000
48	48	1995.9011265	1.858 e-06	4.008	48	48	4569.1204736	1.858 e-06	4.008
56	56	1995.8994140	9.997 e-07	4.019	56	56	4569.1165532	9.997 e-07	4.019
64	64	1995.8985824	5.831 e-07	4.037	64	64	4569.1146497	5.831 e-07	4.037
192	192	1995.8974186			192	192	4569.1121206		

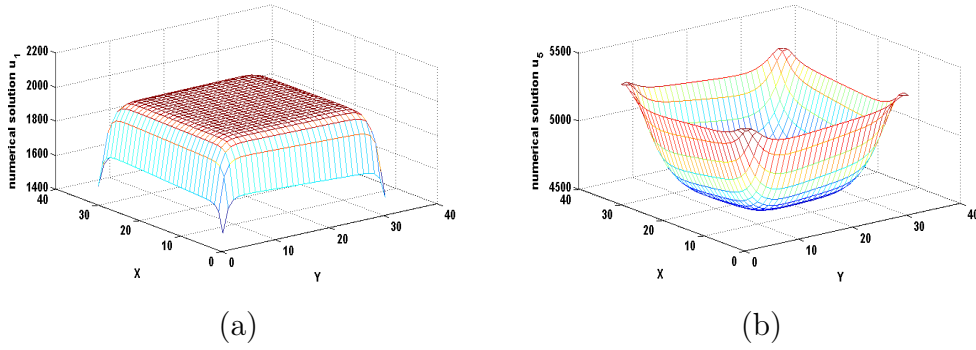
жата.

В Таблица 20 са представени резултатите от CDFS брой слоеве по времето $N = 256$ за първото и петото вещество u_1 и u_5 в централния възел с координати $(x_{M/2}, y_{M/2}) = (X/2, Y/2) = (250, 250)$. Със същите параметри експериментите са повторени с използването на CDS с RE. Четвъртият ред на сходимост в двата случая (централен възел $(x, y) = (X/2, Y/2)$ и възел $(x, y) = (X/6, Y/6)$) за двете вещества u_1 и u_5 е потвърден. Отново в централния възел относителните грешки са почти еднакви. Потвърден е и шестия ред на сходимост пространството на CFDS с RE.

На Фиг. 10 Логаритмична графика на грешките и пространствените мрежови възли за *Пример 2* е представена, получена с: CDS - червена линия; CFDS - виненочервена линия; CDSRE - зелена линия; CFDSRE - синя линия. Нарастването на наклона на линиите съответства на нарастването на реда на сходимост. Най-наклонената линия съответства на предимството на CFDS в комбинация с екстраполация по Ричардсон. Проведени са и други експерименти. Интересно е да се види поведението на решението, ако коефициентите μ на конвекцията



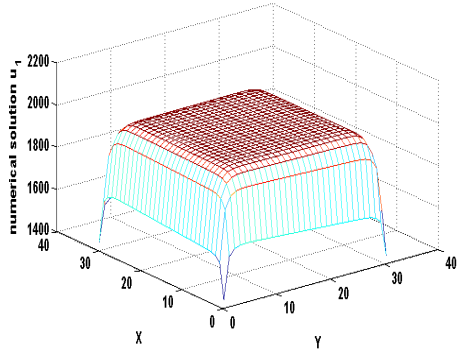
Фигура 10: Логаритмична графика на грешките и пространствените мрежови възли за *Пример 2*, получен с: CDS - червена линия, —★—; CFDS - виненочервена линия, —■—; CDSRE - зелена линия, —◆—; CFDSRE - синя линия, —●—.



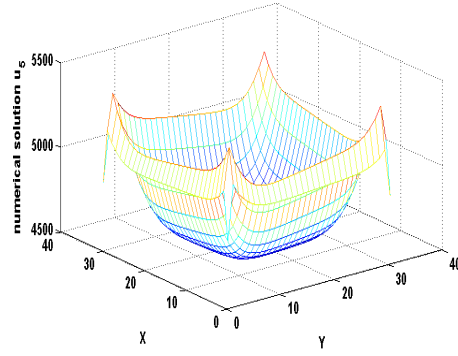
Фигура 11: Численото решение получено с CDS за $\mu = 2\pi/(60T)$ с мрежови параметри $M_x = M_y = 32$, $N = 256$ за *Пример 2*: (a) за u_1 ; (b) за u_5

са $\mu = 2\pi/(X)$, както е в [29], вместо $\mu = 2\pi/(60T)$, както е в [19]. На Фиг. 11 е показано численото решение получено с CDS за $\mu = 2\pi/(60T)$ с мрежови параметри $M_x = M_y = 32$, на последния слой по времето $N = 256$ (a) за u_1 ; (b) за u_5 . Аналогично, на Фиг. 12 е показано численото решение, получено с CFDS. Нарастването на конвективните членове води до съществено изменение на численото решение близо до границата – виж Фиг. 13.

Въпреки всички предимства на CFDS по отношение на точност и CPU време, има някои недостатъци. Шаблонът на CFDS е 9 точков и условието на знаците на дискретния принцип на максимума не е изпълнено. В резултат на

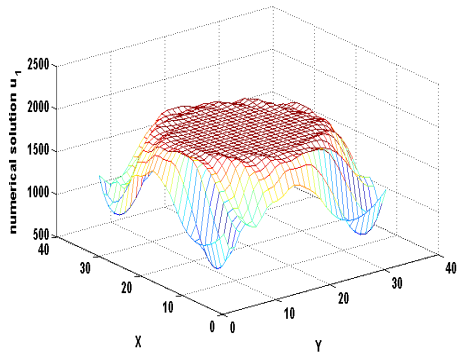


(a)

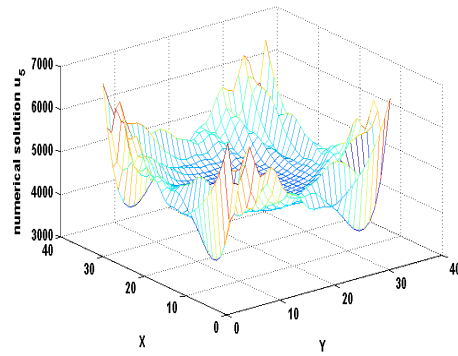


(b)

Фигура 12: Численото решение получено с CFDS за $\mu = 2\pi/(60T)$ с мрежови параметри $M_x = M_y = 32$, $N = 256$ за *Пример 2*: (a) за u_1 ; (b) за u_5

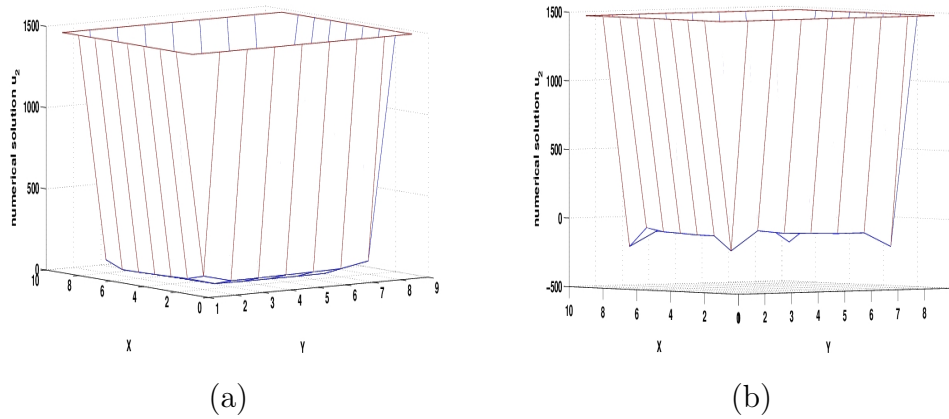


(a)



(b)

Фигура 13: Численото решение за *Пример 2*, получено с CFDS за $\mu = 2\pi/(500)$ и $M_x = M_y = 32$, $N = 256$: (a) за u_1 ; (b) за u_5



Фигура 14: Численото решение за замърсителя NO_2 - u_2 за $\mu = 2\pi/(60T)$ с параметри на мрежата $M_x = M_y = 8$, $N = 256$ Пример 2, получено с: (a) CDS; (b) CFDS

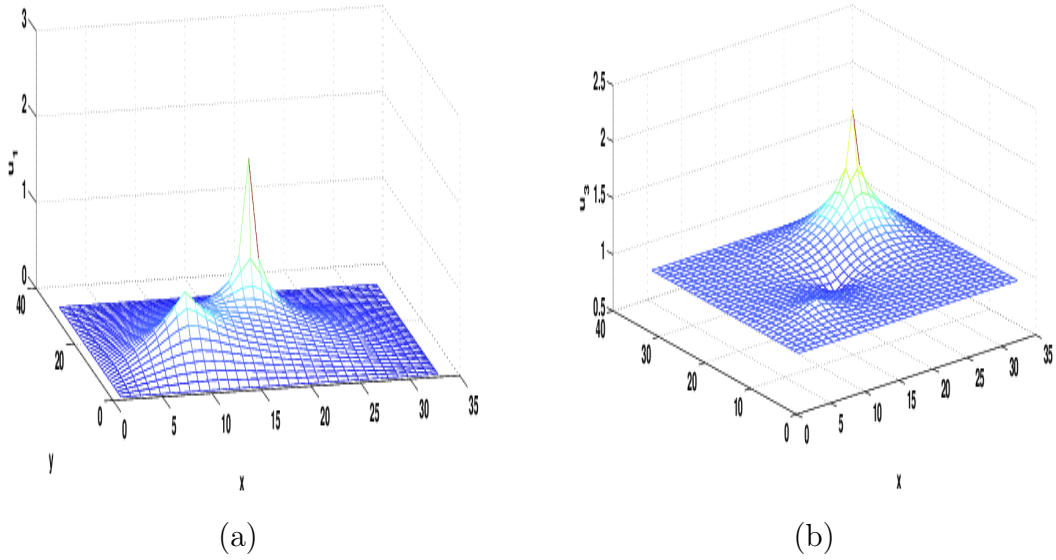
това положителността на численото решение е нарушена за някои стойности на параметрите на мрежата по времето и пространството. На Фиг. 14 численото решение за замърсителя NO_2 (u_2) за Пример 2 когато $\mu = 2\pi/(60T)$ и $M_x = M_y = 8$, $N = 256$, получено със (a) CDS и с (b) CFDS е представено. Схемата CDS запазва положителността на численото решение, докато CFDS не го запазва - близо до ъглите численото решение е отрицателно и няма теоретично обяснение.

Числени експерименти за атмосферен модел на базата на цикъла на Чапман

Използваме опростен, но реалистичен, модел на три компонентна химична система, която моделира химични процеси, протичащи в атмосферата. В тази секция се концентрираме на системата (20)-(22) за $L = 3$ с коефициенти, реакции и източници, които отговарят на модел в атмосферата, базиран на цикъла на Чапман [30, 35]. Докато в един реалистичен атмосферен модел в екологията могат да участват концентрации на много реагиращи вещества [53, 19], опростеният модел тук притежава основните свойства на големите практически модели. Разглеждаме системата:

$$\frac{\partial u_l}{\partial t} - K \Delta u_l + \mathbf{b}_l \cdot \nabla u_l = R_l(x, y, \mathbf{u}) + \xi_l(x, y, t), (x, y, t) \in \Omega \times (0, T]. \quad (40)$$

Компонентите на системата са азотен оксид (NO), азотен диоксид (NO_2) и озон



Фигура 15: Численото решение с параметри на мрежата $M_x = M_y = 32$, $N = 256$ за *Пример 2*: (a) за $NO - u_1$; (b) за $O_3 - u_3$

(O_3) означени с u_1, u_2, u_3 съответно:

$$R_l(\mathbf{u}) = -r(\mathbf{u}), \quad l = 1, 3, \quad R_2(\mathbf{u}) = r(\mathbf{u}), \quad r(\mathbf{u}) = k_1 u_1 u_3 - k_2 u_2,$$

където k_1, k_2 са скоростите на реакциите.

В изследвания пример се фокусираме върху локални източници, които се описват от Делта - функции. Разглеждаме задачата (40), където функциите ξ_l са точкови източници от вида

$$\xi_l(x, y, t) = f_l(t) \delta(x - \bar{x}_l, y - \bar{y}_l), \quad l = 1, 2, 3.$$

Параметрите са, както следва: $\bar{\Omega} = [0, 500] \times [0, 500]$, $T = 1440$, $b_l = (-0.1, 0)$, за $l = 1, 2, 3$, $K_1 = 1$, $K_2 = K_3 = 5$, $k_1 = 1000$, $k_2 = 2000$, $(\bar{x}_1, \bar{y}_1) = (0.5, 0.5)$, $(\bar{x}_2, \bar{y}_2) = (0.25, 0.25)$, $(\bar{x}_3, \bar{y}_3) = (0.75, 0.75)$, $f_1(t) = 7$, $f_2(t) = 11$, $f_3(t) = 13$. На Фиг. 15 е представено численото решение получено с CFDS с параметри $M_x = M_y = 32$, $N = 256$: (a) за $NO - u_1$; (b) за $O_3 - u_3$. Влиянието на точковите източници се вижда ясно.

Авторска справка

Основните научни приноси на настоящата дисертация са:

1. Изследвани са различни методи Монте Карло и квази-Монте Карло и подходи за генериране на случайни извадки - адаптивен метод Монте Карло, специална решетка, базирана на обобщената редица на Фибоначи, извадка латински хиперкуб. Адаптивният алгоритъм има предимство пред разглежданите алгоритми за функции с изчислителни особености. За гладки функции с по-ниска размерност най-ефективна е решетката с числа на Фибоначи, а при високи размерности - извадката латински хиперкуб.
2. Конструирани са нов почти оптимален алгоритъм Монте Карло за интегрални уравнения, базиран на балансиране на систематичната и стохастичната грешка. Изведени са оценки за избор на броя реализации на случайната величина и броя итерации във веригата на Марков, които са от съществено значение за неговата ефективност и точност.
3. Разработен е нов метод Монте Карло за линейни системи, който е подобрило на метода „случайно блуждаене по уравненията на СЛАУ“. Методът успешно се прилага при решаване на СЛАУ, получени след дискретизация на ЧДУ.
4. Разработени са и са изследвани нови компактни диференчни схеми с четвърти ред на точност по пространството за системи от параболични ЧДУ със свързани нелинейни реакции.

Основните научно-приложни приноси на настоящата дисертация са:

1. Приложени са различни ефективни Монте Карло и квази-Монте Карло алгоритми за ядрото на Вигнер, които са по-ефективни от досега използваните детерминистични алгоритми по отношение на точността и изчислителната сложност.
2. Приложени са различни ефективни Монте Карло и квази-Монте Карло алгоритми за гладки подинтегрални функции при оценки на Европейски опции с важно значение във финансите и за многомерни интеграли в Бейсовската статистика с приложение при машинното обучение.

3. Приложена е схема с шести ред на точност по пространствената променлива за реална физична параболична транспортна система, описваща далечен пренос на замърсители във въздуха, от Датския Ойлеров модел и за атмосферен модел, базиран на цикъла на Чапман.

Благодарности

Издавам своята искрена признателност и благодарност на научния си ръководител проф. д-н Иван Димов за неговите ценни напътствия, професионална компетентност и съдействие при провеждане на настоящите изследвания и при подготовката на дисертацията. Изключително благодаря и за неговата неоценима морална подкрепа и за проявеното търпение.

Благодаря изключително много на доц. д-р Райна Георгиева за нейните ценни напътствия, готовността ѝ за съдействие във всеки момент и неоценимата помощ при подготовката на дисертацията.

Благодаря на колегите от Русенския Университет проф. д-р Любен Вълков и доц. д-р Юрий Кандиларов за полезните дискусии и предоставянето на интересни задачи.

Авторът изказва и благодарности на екипа от секция "Паралелни алгоритми" проф. Стефка Фиданова, доц. Пенчо Маринов, доц. Цветан Остромски и доц. Михаил Недялков за приобщаването на автора към колектива и за ценните напътствия и оказаното съдействие. Благодарности дължа и за целия екип на Института по информационни и комуникационни технологии и към неговия Директор за осигурените качествени условия за работа и обучение.

Благодаря на моите колеги Стоян Димитров от ФМИ-СУ и д-р Николай Икономов от ИМИ-БАН за съвместната работа.

Благодаря специално на баща ми и брат ми за тяхната морална и финансова подкрепа, търпение и стимулиращи напътствия.

Благодаря за финансовата подкрепа по проектите: #ФНИ И-02-20, #ДФНП 91-A1, #ДМУ 03/61.

Библиография

- [1] Боянов, Б.: *Лекции по числени методи*. Дарба, София, 1998.
- [2] Караиванова, А.: *Стохастични числени методи и симулации*, София, 2012.
- [3] Черногорова, Т.: Теория на диференчните схеми, София, 2005.
<https://www.fmi.uni-sofia.bg/econtent/tds.pdf>
- [4] Янев, Н., Димитров, Б.: *Вероятности и статистика*. Наука и изкуство, София, 1990.
- [5] Bahvalov, N.: On the Approximate Computation of Multiple Integrals. In *Vestnik Moscow State University, Ser. Mat., Mech.*, volume 4, pages 3–18, 1959.
- [6] Dembo R.S., Eisenstat, S.C., Steihaug, T.: Inexact Newton methods. *SIAM Journal on Numerical Analysis*, **19**(2), 400-408 (1982)
- [7] Dimov, I.: *Monte Carlo Methods for Applied Scientists*, New Jersey, London, Singapore, **World Scientific** (2008), 291 p., ISBN-10 981-02-2329-3.
- [8] Dimov, I., R. Georgieva: Monte Carlo Algorithms for Evaluating Sobol' Sensitivity Indices. *Math. Comput. Simul.* **81** (2010) (3), 506-514.
- [9] Dimov, I.T., Georgieva, R., Todorov, V.: Balancing of Systematic and Stochastic Errors in Monte Carlo Algorithms for Integral Equations, 8th NMA 2014, Borovets, Bulgaria, August 20-24, 2014, LNCS 8962, Springer, 2015, 44–51.
- [10] Dimov, I., Gurov, T.: Monte Carlo Algorithm for Solving Integral Equations with Polynomial Non-linearity. Parallel Implementation, Pliska (Studia Mathematica Bulgarica), Vol. 13 2000, Proceedings of the 9th International Summer School on Probability Theory and Mathematical Statistics, Sozopol, 1997, pp. 117-132.
- [11] Dimov, I.T., Gurov, T.: *A New Iterative Monte Carlo Approach for Inverse Matrix Problem*, **Journal of Computational and Applied Mathematics**, Vol. **92** (1998), pp. 15-35.
- [12] Dimov I., Kandilarov, J., Todorov, V., Vulkov, L.: High-Order Compact Difference Schemes with Richardson Extrapolation for Semilinear Parabolic Systems. IN:

- Applications of Mathematics in Engineering and Economics, American Institute of Physics, 1789, 030002 (2016).
- [13] Dimov, I.T., Maire, S., Sellier, J.M.: A New Walk on Equations Monte Carlo Method for Linear Algebraic Problems, *Applied Mathematical Modelling*, Volume 39, Issue 15, 2015, Pages 4494–4510.
 - [14] Dimov, I., Todorov, V.: Error Analysis of Biased Stochastic Algorithms for the Second Kind Fredholm Integral Equation, *Studies in Computational Intelligence* 648, 2016, pp. 3-16.
 - [15] Doucet, A., Johansen, A.M., Tadic, V.B.: On solving integral equations using Markov chain Monte Carlo methods. *Applied Mathematics and Computations*, **216** 2869–2880, (2010).
 - [16] Ermakov, S.M.: *Monte Carlo Methods and Mixed Problems*, **Nauka**, Moscow, (1985).
 - [17] Feynman, R.P.: Space-time approach to non-relativistic quantum mechanics, *Rev. Mod. Phys.* 20 (1948).
 - [18] Genz, A.: Testing multidimensional integration routines. *Tools, Methods and Languages for Scientific and Engineering Computation* (1984) 81–94.
 - [19] Georgiev, K., Zlatev, Z.: Implementation of sparse matrix algorithms in an advection-diffusion-chemistry model. *J. of Comp. Appl. Math.*, **236** (3), 342-353 (2011)
 - [20] Georgieva, R.: PhD Thesis: Computational complexity of Monte Carlo algorithms for multidimensional integrals and integral equations, Sofia, 2003
 - [21] Gustafsson, B., Kreiss, H., Oliger, J.: *Time Dependent Problems and Difference Methods*, Wiley, New York (1995)
 - [22] Halton, J.: *Sequential Monte Carlo, Proceedings of the Cambridge Philosophical Society*, **Vol. 58** (1962) pp. 57-78.
 - [23] Halton, J.: *Sequential Monte Carlo for linear systems - a practical summary*, **Monte Carlo Methods & Applications**, **14** (2008) pp. 1–27.
 - [24] Halton, J.: *Sequential Monte Carlo. University of Wisconsin*, Madison, Mathematics Research Center Technical Summary Report No. 816 (1967) 38 pp.
 - [25] Hua, L.K., Wang, Y.: *Applications of Number Theory to Numerical analysis*, 1981
 - [26] Jarosz, Wojciech: Efficient Monte Carlo Methods for Light Transport in Scattering Media, PhD dissertation, UCSD, 2008
 - [27] Kantorovich, L., Akilov, G.: *Functional Analysis*. Nauka, Moscow, 1977.

- [28] Karaivanova, A., Dimov, I.: Error analysis of an adaptive Monte Carlo method for numerical integration, *Mathematics and Computers in Simulation* **47** (1998) 201–213.
- [29] Karatson, J., Kurics. T.: A preconditioned iterative solution scheme for nonlinear parabolic systems arising in air pollution modeling. *Math. Modell. Anal.* **18** (5), 641-653 (2013)
- [30] Kim, J., Cho, S.: Computation accuracy and efficiency of the time splitting method. *J. Atmosph. Envir.* **31**(15), pp. 2215-2224 (1997).
- [31] Korobov, N.M.: *Number-Theoretical Methods in Approximate Analysis*, Fizmatgiz, Moscow, 1963.
- [32] Lai, Y., Spanier, J.: *Applications of Monte Carlo/Quasi-Monte Carlo Methods in Finance: Option Pricing*, Proceedings of a conference held at the Claremont Graduate Univ, 1998.
- [33] Lin, S.: “Algebraic Methods for Evaluating Integrals in Bayesian Statistics,” Ph.D. dissertation, UC Berkeley, May 2011.
- [34] Lin, S., Sturmfels B., Xu Z.: Marginal Likelihood Integrals for Mixtures of Independence Models, *Journal of Machine Learning Research*, Vol. 10 (2009), pp. 1611-1631.
- [35] Mamonov, A. V., Tsai, Y.-H. R.: Point source identification in non-linear advection-diffusion-reaction systems. *Inverse Problems* **29**(3), pp. 035009 (2012)
- [36] Marchuk, G.I., Shaidurov, V.V.: *Difference Methods and Their Extrapolations* (Springer-Verlag, New York Inc. 1983).
- [37] McKay, M.D., Beckman, R.J., Conover, W.J.: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21(2), 239–245 (1979), doi:10.2307/1268522
- [38] Metropolis, N., Ulam, S.: *The Monte Carlo Method*, **J. of Amer. Statistical Assoc.**, **44**, (1949), No. 247, pp. 335–341.
- [39] Minasny B., McBratney B.: A conditioned Latin hypercube method for sampling in the presence of ancillary information *Journal Computers and Geosciences archive*, Volume 32 Issue 9, November, 2006, Pages 1378-1388.
- [40] Mirzaee F., Samadyar N.: Application of orthonormal Bernstein polynomials to construct a efficient scheme for solving fractional stochastic integro-differential equation, *Optik - International Journal for Light and Electron Optics*, Volume 132, March 2017, Pages 262–273.
- [41] Pao C. V.: *Nonlinear Parabolic and Elliptic Equations*, Springer, US, 1992.

- [42] Richards, S.: Completed Richardson extrapolation in space and time. Commun. Numer. Meth. Engineering **13**, 573–582 (1997).
- [43] Sellier, J.M.: A signed particle formulation of non-relativistic quantum mechanics, Journal of Computational Physics 297(2015) 254–265.
- [44] Sloan, I.H., Joe, S.: Lattice Methods for Multiple Integration, Oxford University Press, Oxford (1994).
- [45] Sloan, I.H., Kachoyan, P.J.: Lattice methods for multiple integration: Theory, error analysis and examples, SIAM J. Numer. Anal. 24, (1987) 116-128.
- [46] Sobol, I.M.: *Monte Carlo Numerical Methods*, Nauka, Moscow, 1973, (in Russian).
- [47] Todorov, V., Dimov, I.: Monte Carlo methods for multidimensional integration for European option pricing, AIP Conf. Proc. 1773, 100009, (2016).
- [48] Todorov, V.: Computing high dimensional integrals with Monte Carlo methods, Journal Scientific and Applied Research, Vol.10, 2016, 11-16, ISSN 1314-6289.
- [49] Wang, Y.M., Guo, B.Y., Wu, W.J.: Fourth-order compact finite difference methods and monotone iterative algorithms for semilinear elliptic boundary value problems, Computers and Math. with Appl., **68**, 1671-1688 (2014)
- [50] Wang, Y., Hickernell, F.J.: *An historical overview of lattice point sets*, (2002).
- [51] Wigner, E.: On the quantum correction for thermodynamic equilibrium, Phys. Rev. 40 (1932) 749.
- [52] Zlatev, Z.: *Computer Treatment of Large Air Pollution Models* (Kluwer Academic Publishers, 1995).
- [53] Zlatev, Z., Dimov, I.: *Computational and Numerical Challenges in Environmental Modelling* (Elsevier, Amsterdam, 2006).
- [54] Zlatev, Z., Dimov, I., Farago, I., Georgiev, K., Havasi. A.: Application of Richardson extrapolation for multi-dimensional advection equations, Comp. Math. Appl., **67**, 2279-2293 (2014)
- [55] Website: Matrix market, NOS4: Lanczos with partial reorthogonalization. Finite element approximation to a beam structure, <http://math.nist.gov/MatrixMarket/data/Harwell-Boeing/lanpro/nos4.html>
- [56] MATLAB - MathWorks: <http://www.mathworks.com/>