



Силви-Мария Тодорова Гюрова

Стохастични числени методи за оценка на собствени стойности

Дисертация

за придобиване на образователна и научна степен “доктор”
по докторска програма: “Математическо моделиране и приложение на
математиката”

научна област: 4. “Природни науки, математика и информатика”
професионално направление: 4.5. “Математика”

научен ръководител: проф. д-р Анета Недева Караиванова

София, 2026

Съдържание

Списък на използваните означения и абривиатури	7
Означения	7
Абривиатури	8
Увод	10
Актуалност на темата	10
Обзор на основните резултати в областта	13
Цели и задачи на дисертацията	14
Методология на изследването	15
Формулиране на задачата	16
Монте Карло методи	16
Квази-Монте Карло методи	23
Стохастични методи с вериги на Марков	34
Структура на дисертацията	35
1 Стохастичен Степенен метод: разработване и изследване на ефективни алгоритми за оценяване на максимална собствена стойност	37
1.1 Постановка на задачата	37
1.2 Методи за изчисляване на максималната собствена стойност	38
1.2.1 Степенен метод	38
1.2.2 Стохастичен Степенен метод	38
1.3 Почти оптимални Степенен Монте Карло и Степенен Квази-Монте Карло методи и алгоритми	42
1.3.1 Прецизиране на конструкцията на $\theta^{(k)}$ и $\theta^{(k-1)}$ с почти оптимални вероятности за веригата на Марков: балансиране на грешките	42
1.3.2 Почти оптимален СМК и почти оптимален СКМК алгоритми: изчислителна сложност	45
1.4 Числени резултати и дискусия	46
1.4.1 Случай (i): Размер на матрицата $n = 100$	48
1.4.2 Случай (ii): Размер на матрицата $n = 500$	55
1.5 Обобщение	62
2 Стохастичен Резолвентен метод: разработване и изследване на ефективни алгоритми за оценяване на минимална собствена стойност	64
2.1 Постановка на задачата	64

2.2	Методи за изчисляване на минималната собствена стойност	64
2.2.1	Резолвентен степенен метод	66
2.2.2	Стохастичен Резолвентен метод	69
2.3	Почти оптимален Резолвентен Монте Карло и Резолвентен Квази-Монте Карло алгоритми	72
2.4	Числени резултати и дискусия	75
2.4.1	Случай (i): Размер на матрицата $n=500$	76
2.4.2	Случай (ii): Размер на матрицата $n=1000$	78
2.5	Обобщение	80
3	Стохастични методи за оценяване на пазарния риск на инвестиционни портфейли	83
3.1	Постановка на задачата	83
3.2	Методи и данни	84
3.2.1	Обща формулировка на ковариационна и корелационна матрица	84
3.2.2	Конструиране на ковариационна и корелационна матрица от емпирични данни от финансови пазари	85
3.2.3	Дял на обяснената вариация	87
3.3	Стохастични алгоритми за оценяване на максималната собствена стойност на корелационна матрица	88
3.4	Числени резултати и дискусия	90
3.4.1	Случай (i): корелационна матрица с размерност $n = 32$	91
3.4.2	Случай (ii): корелационна матрица с размерност $n = 490$	97
3.5	Обобщение	99
	Заклучение	101
	Списък на публикациите по дисертацията	102
	Апробация на резултатите	102
	Основни научни и научно-приложни приноси	104
	Декларация за оригиналност	106
	Благодарности	106
	Приложение	107
	Литература	109

Списък на таблиците

1.1	Резултати от $PMC_{(MT)}$ и $PMC_{(MT)}^{(n)}$ при различни N и k за матрица A с размерност $n = 100$, за която $\lambda_2/\lambda_1 \approx 0.081$, където $\lambda_1 = 50.0408$ и $\lambda_2 = 4.0522$	49
1.2	Резултати от $PMC_{(MS)}$ и $PMC_{(MS)}^{(n)}$ при различни N и k за матрицата A с размерност $n = 100$, за която $\lambda_2/\lambda_1 \approx 0.081$, където $\lambda_1 = 50.0408$ и $\lambda_2 = 4.0522$	50
1.3	Резултати от $PQMC_{(H)}$ и $PQMC_{(H)}^{(n)}$ при различни N и k за матрица A с размерност $n = 100$, за която $\lambda_2/\lambda_1 \approx 0.081$, където $\lambda_1 = 50.0408$ и $\lambda_2 = 4.0522$	52
1.4	Резултати от $PQMC_{(S)}$ и $PQMC_{(S)}^{(n)}$ при различни N и k за матрица A с размерност $n = 100$, за която $\lambda_2/\lambda_1 \approx 0.081$, където $\lambda_1 = 50.0408$ и $\lambda_2 = 4.0522$	53
1.5	Резултати от $PMC_{(MT)}$ и $PMC_{(MT)}^{(n)}$ при различни N и k за матрица A с размерност $n = 500$, за която $\lambda_2/\lambda_1 \approx 0.0363$, където $\lambda_1 = 250.2454$ и $\lambda_2 = 9.0721$	56
1.6	Резултати от $PMC_{(MS)}$ и $PMC_{(MS)}^{(n)}$ при различни N и k за матрица A с размерност $n = 500$, за която $\lambda_2/\lambda_1 \approx 0.0363$, където $\lambda_1 = 250.2454$ и $\lambda_2 = 9.0721$	57
1.7	Резултати от $PQMC_{(H)}$ и $PQMC_{(H)}^{(n)}$ при различни N и k за матрица A с размерност $n = 500$. Абсолютната грешка е за случая $\lambda_2/\lambda_1 \approx 0.0363$	59
1.8	Резултати от $PQMC_{(S)}$ и $PQMC_{(S)}^{(n)}$ при различни N и k за матрица A с размерност $n = 500$. Абсолютната грешка е за случая $\lambda_2/\lambda_1 \approx 0.0363$	60
1.9	Емпирична оценка на дисперсиите получена чрез <i>почти оптималния</i> и <i>класическия</i> СМК алгоритъм с генератори MT и MS при оценяване на максималната собствена стойност.	62
2.1	Числени резултати за матрицата $A \in \mathbb{R}^{500 \times 500}$ при различни стойности на N, k и m с параметъра на ускорение $q = -0.157428$. Точната минимална собствена стойност $\lambda_{\min} = \lambda_{500} = 0.152992$	76
2.2	Числени резултати за матрицата $A \in \mathbb{R}^{1000 \times 1000}$ при различни стойности на N, k и m с параметър на ускорение $q = -0.2302458$. Точната минимална стойност $\lambda_{\min} = \lambda_{1000} = 0.151911$	79
3.1	Резултати за максималните стойности на $\lambda_{max}^{(k)}$ и абсолютната грешка за $PMC_{(MT)}$ и $PMC_{(MS)}$ при различни N и k , когато $\lambda_{max} = 18.14714$	92

3.2	Резултати за максимални собствени стойности за $\lambda_{max}^{(k)}$ и абсолютна грешка чрез сценарий по <i>подразбиране</i> и сценарий, използващ <i>skip</i> и <i>leap</i> при различни N и k , когато $\lambda_{max} = 18.14714$	94
3.3	Приблизени стойности за дяла на обяснената вариация (FVE_1), получени при $k = 8$ за $PMC_{(MT)}$, $k = 11$ за $PQMC_{(S)}^{(d)}$ и $k = 13$ за $PQMC_{(S)}^{(s,l)}$ за корелационната матрица \hat{R} с размерност $n = 32$. Референтната стойност е $FVE_1 = 56.71\%$	96
3.4	Резултати за максимални собствени стойности за $\lambda_{max}^{(k)}$ и абсолютна грешка при $PMC_{(MT)}$, $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ при различни N и k , когато $\lambda_{max} = 201.090512$	97
3.5	Приблизени стойности за дяла на обяснената вариация (FVE_1), получени при $k = 8$ за $PMC_{(MT)}$ и при $k = 11$ за $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ за корелационната матрица \hat{R} с размерност $n = 490$. Референтната стойност е $FVE_1 = 41.04\%$	99

Списък на фигурите

1	Сравнение на двумерна псевдослучайна редица, получена с МТ PRNG, с двумерни проекции на 16-мерни редици на Собол и Холтън. Генерираните точки са 256, а проекцията е направена по първите две координати.	29
2	Сравнение между двумерна проекция с наличие на корелация при 11-та и 12-та координата на редица на Собол и съответните двумерни проекции на рандомизирани редици на Собол без и с допълнително включени параметри $skip = 1024$ и $leap = 128$	31
3	Сравнение между двумерна проекция с наличие на корелация при 11-та и 12-та координата на редица на Холтън и съответните двумерни проекции на рандомизирани редици на Холтън без и с допълнително включени параметри $skip = 1024$ и $leap = 128$	32
1.1	Графично сравняване на приближените резултати за $\lambda_{max}^{(k)}$ с $\lambda_{max} = 50.0408$ при балансирания случай за грешките, т.е. когато ($k = 9, N = 128$) и ($k = 8, N > 128$). Приложени алгоритми: $PMC_{(MT)}$, $PMC_{(MS)}$, $PMC_{(MT)}^{(n)}$ и $PMC_{(MS)}^{(n)}$	51
1.2	Графично сравняване на приближените резултати за $\lambda_{max}^{(k)}$ с $\lambda_{max} = 50.0408$ при балансирания случай за грешките, т.е. когато ($k = 12, N = 128$) и ($k = 11, N > 128$). Приложени алгоритми: $PQMC_{(S)}$, $PQMC_{(H)}$, $PQMC_{(S)}^{(n)}$ и $PQMC_{(H)}^{(n)}$	54
1.3	Графично сравнение на λ_{max} на симетричната матрица A ($n = 100$) с $\lambda_{max}^{(k)}$ в балансиращите случаи. Използвани са <i>почти оптималните</i> СМК и СКМК алгоритми.	54
1.4	Графично сравняване на λ_{max} за матрицата A ($n = 500$) с приближените резултати за $\lambda_{max}^{(k)}$, получени при балансирания случай на грешките, т.е. когато ($k = 9, N \geq 128$). Приложени алгоритми: $PMC_{(MT)}$, $PMC_{(MS)}$, $PMC_{(MT)}^{(n)}$ и $PMC_{(MS)}^{(n)}$	58
1.5	Графично сравняване на λ_{max} за матрицата A ($n = 500$) с приближените резултати за $\lambda_{max}^{(k)}$, получени при балансирания случай на грешките, т.е. когато ($k = 12, N \geq 128$). Приложени алгоритми: $PQMC_{(S)}$, $PQMC_{(H)}$, $PQMC_{(S)}^{(n)}$ и $PQMC_{(H)}^{(n)}$	61
1.6	Графично сравнение на λ_{max} на симетричната матрица A ($n = 500$) с $\lambda_{max}^{(k)}$ в балансиращите случаи. Използвани са <i>почти оптималните</i> СМК и СКМК алгоритми.	61

2.1	Сходимость към точната минимална собствена стойност на матрицата A при използване на Алгоритъм 2 и Алгоритъм 2а с параметри $k = 5, m = 10$ за различни на брой N	77
2.2	Сходимость към точната минимална собствена стойност на матрицата A при използване на Алгоритъм 2 и Алгоритъм 2а с параметри $k = 5, m = 5$ за различни на брой N	80
3.1	Сравнение между точната максимална собствена стойност λ_1 и приближена собствена стойност $\lambda_1^{(k)}$ за $PMC_{(MT)}, PMC_{(MS)}$, получени в балансиращите случаи при различни стойности на N и $k = 8$	93
3.2	$(PQMC_{(H)}^{(d)}, PQMC_{(S)}^{(d)})$ и $(PQMC_{(H)}^{(s,l)}, PQMC_{(S)}^{(s,l)})$. Сравнение между максималната собствена стойност λ_{\max} на корелационната матрица и оценените стойности $\lambda_1^{(k)}$, получени при $k = 11$ в случая по <i>подразбиране</i> и $k = 13$ в случая с пропускане и прескачане и различни стойности на N	95
3.3	Сравнение между максимална собствена стойност λ_{\max} и приближените собствени стойности $\lambda_1^{(k)}$, получени при $k = 8$ за $PMC_{(MT)}$, $k = 11$ за $PQMC_{(S)}^{(d)}$ и $k = 13$ за $PQMC_{(S)}^{(s,l)}$ при различни стойности на N	96
3.4	Сравнение между максимална собствена стойност λ_{\max} и оценените собствени стойности $\lambda_1^{(k)}$, получени при $k = 8$ за $PMC_{(MT)}$ и $k = 11$ за $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ за различни стойности на N	98

Списък на използваните означения и абревиатури

Означения

\mathbb{R}^n - n -мерно евклидово пространство

N - брой реализации (опити, наблюдения, експерименти, проби)

$\epsilon \in \mathbb{R}$ - достатъчно малко положително реално число

n - размерност на матрицата A

m - степен на резолвентната матрица R_q

k - брой скокове във веригата на Марков

$q \in \mathbb{R}$ - параметър на ускорение на сходимостта

$C_n^k = \binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\dots(n-k+1)}{k(k-1)\dots 1}$ - биномен коефициент

$A = \{a_{ij}\} \in \mathbb{R}^{n \times n}$ е квадратна матрица с a_{ij} - елемент в i -тия ред и j -тия стълб

$\det(A)$ - детерминанта на матрицата A

I - единична матрица

$\mathbf{x} = \{x_i\}_{i=1}^n = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ е n -мерен вектор в \mathbb{R}^n

$\mathbf{x}^T = \{x_i\}_{i=1}^n{}^T = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$ е n -мерен вектор в \mathbb{R}^n

$\mathbf{f} = \{f_i\}_{i=1}^n$, $\mathbf{h} = \{h_i\}_{i=1}^n \in \mathbb{R}^n$ са произволни n -мерни вектори в \mathbb{R}^n

$(h, f) = \sum_{i=1}^n h_i f_i$ е скалярно произведение на произволни n -мерни вектори в \mathbb{R}^n

$\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}$ е p -норма или \mathcal{L}_p -норма на вектора $\mathbf{x} = \{x_i\}_{i=1}^n \in \mathbb{R}^n$ при $p \geq 1$

$\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$ е първа норма на вектора $\mathbf{x} = \{x_i\}_{i=1}^n \in \mathbb{R}^n$ при $p = 1$

$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}$ е евклидова норма на вектора $\mathbf{x} = \{x_i\}_{i=1}^n \in \mathbb{R}^n$ при $p = 2$

$\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$ е максимална норма или \mathcal{L}_∞ -норма на вектора $\mathbf{x} = \{x_i\}_{i=1}^n \in \mathbb{R}^n$ при $p = \infty$

$\|A\|_p = \sup\{\|A\mathbf{x}\| : \mathbf{x} = \{x_i\}_{i=1}^n \in \mathbb{R}^n, \|\mathbf{x}\|_p \leq 1\}$ е p -норма или \mathcal{L}_p матрична норма

$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$ е \mathcal{L}_1 матрична норма при $p = 1$

$\|A\|_2 = \sup_{\|\mathbf{x}\|_2=1} \|A\mathbf{x}\| = \sup_{\|\mathbf{x}\|_2=1} \sqrt{(A^T A \mathbf{x}, \mathbf{x})}$ е \mathcal{L}_2 матрична норма при $p = 2$

$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$ е \mathcal{L}_∞ матрична норма при $p = \infty$

λ, μ са скаларни величини (реални числа за симетрична матрица), наречени собствени стойности или собствени числа

θ е случайна величина, дефинирана във вероятностното пространство $(\Omega, \mathfrak{A}, \mathbb{P})$

$E[\theta]$ е математическо очакване на случайната величина θ

$Var[\theta]$ е дисперсия на случайната величина θ

$\sigma(\theta)$ е стандартно отклонение на случайната величина θ

γ е случайно число, равномерно разпределено в $[0, 1]$ с $E(\gamma) = 1/2, Var(\gamma) = 1/12$

$p = \{p_i\}_{i=1}^n \in \mathbb{R}^n, \sum_{i=1}^n p_i = 1$ е n -мерен вектор на плътността

$\mathbb{P}(l_0 = i) = p_i \geq 0, i = 1 \dots n$ е вероятност

$L = (l_0 \rightarrow l_1 \rightarrow \dots \rightarrow l_t \rightarrow \dots \rightarrow l_k), 1 \leq l_t \leq n, 0 \leq t \leq k$ е крайна верига на Марков

$P = \{p_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}$ е матрицата на плътността на преходите, за която

$$\mathbb{P}(l_t = j | l_{t-1} = i) = p_{ij}, \quad i, j = 1 \dots, n, \quad t = 1, \dots, k$$

Абревиатури

Абревиатура	Пълно наименование
CLT	Central Limit Theorem
ЦГТ	Централна гранична теорема
CPU	Central Processing Unit Централен процесор
GPU	Graphics Processing Unit Графичен процесор
LDS	Low Discrepancy Sequences Редици с малък дискрепанс
LLN	Law of Large Numbers
ЗГЧ	Закон за големите числа
MPI	Message Passing Interface Интерфейс за предаване на съобщения
MC	Monte Carlo
МК	Монте Карло
ODE	Ordinary Differential Equations
ОДУ	Обикновени диференциални уравнения
PMC	Power Monte Carlo
СМК	Степенен Монте Карло
PQMC	Power Quasi-Monte Carlo
СКМК	Степенен Квази-Монте Карло
RMC	Resolvent Monte Carlo
РМК	Резолвентен Монте Карло
RQMC	Resolvent Quasi-Monte Carlo
РКМК	Резолвентен Квази-Монте Карло
SDE	Stochastic Differential Equations
СДУ	Стохастични диференциални уравнения
SLAE	System of Linear Algebraic Equations
СЛАУ	Система от линейни алгебрични уравнения
QMC	Quasi-Monte Carlo
КМК	Квази-Монте Карло

Увод

Методите Монте Карло (МК) и рандомизираните Квази-Монте Карло (КМК), разглеждани в научната литература като стохастични числени методи, са мощни инструменти за решаване на сложни задачи с големи размерности, възникващи в много области на математиката, физиката, финансите и инженерните науки. Те обхващат широк клас изчислителни алгоритми, при които търсената величина се оценява чрез усредняване върху случайни (или рандомизирани квазислучайни) извадки, така че полученото приближение да е с висока вероятност близко до точната стойност. Съществен проблем при разработването на стохастични алгоритми е постигането на висока точност при относително малък брой проби спрямо размерността на задачата. Алгоритъм, който осигурява надеждна точност при ограничен обем симулации, се определя като ефективен МК/КМК алгоритъм. В много случаи стохастичните методи са единствената приложима алтернатива за оценяване на решенията на сложни задачи, дефинирани в пространства с голяма размерност, където детерминистичните подходи са изчислително неприложими.

В настоящата дисертация акцентът е поставен върху реализацията и численото изследване на нови ефективни стохастични алгоритми за оценяване на екстремалните собствени стойности на квадратни матрици с голяма размерност, в частност на симетрични матрици. Алгоритмите за оценяване на максималната собствена стойност са разработени на базата на Степенния Монте Карло (СМК) и Степенния Квази-Монте Карло (СКМК) подход. За оценяване на минималната собствена стойност са използвани Резолвентният Монте Карло (РМК) и Резолвентният Квази-Монте Карло (РКМК) подход.

Предложените алгоритми могат да бъдат приложени за реални задачи, възникващи в биологията, медицината, физиката, финансите и други научни области. В дисертационния труд е разглеждана и конкретна задача от финансовата математика, свързана с оценяване на риска и анализ на конструиран портфейл от активи.

Актуалност на темата

Задачата за намирането на собствени стойности на квадратни матрици е една от основните задачи на линейната алгебра. В частност, оценяването на екстремалните собствени стойности на симетрични матрици е от особено значение поради широкия спектър от приложения в различни научни и приложни области, като например:

- В квантовата механика, при уравнението на Шрьодингер [81], собствените стойности на Хамилтоновия оператор описват енергийния спектър на квантовата система.

Познаването на най-голямата и най-малката собствена стойност дава информация съответно за енергията на основното състояние и за възбудените състояния [80, 93].

- Във финансовата математика, особено при моделиране на риска и оптимизация на портфейл от активи, ковариационната/корелационната матрица на възвръщаемост е от съществено значение. Най-голямата собствена стойност характеризира доминиращата вариация в портфейла и често се интерпретира като индикатор за систематичен пазарен риск. Анализът ѝ е съществен за идентифициране на възможности за диверсификация [61, 92].
- В спектралната теория на графите, екстремалните собствени стойности на матрицата на съседство са свързани с важни структурни свойства на графа като разширение, свързаност и устойчивост. Тези характеристики влияят върху ефективността на алгоритми в машинното обучение и мрежовия анализ [13, 17, 32].
- В обработката на изображения и компютърното зрение (например, разпознаване на лица и компресия), собствените стойности определят как е разпределена „енергията“ на изображението. В рамките на анализа на главните компоненти (англ. *Principal Component Analysis*, PCA) те подпомагат редуцията на размерността и потискането на шума [55, 108]. При спектрални методи за сегментация (например, *normalized cuts*) и клъстериране, разликите между екстремалните собствени стойности служат като индикатор за степента на разделимост на сегментите [40].
- В тензорния анализ и обработката на многомерни данни, спектралните норми са пряко свързани с най-голямата собствена стойност. Те управляват поведението на итеративните решения и техниките за регуляризация, които от своя страна се използват при обработка на сигнали и модели на изкуствен интелект. Изучаването на максималната собствена стойност в тензорни декомпозиции е важно за разбирането на концентрацията на енергия, устойчивостта и интерпретируемостта на извлечените компоненти [59, 67, 102].

За ефективно изчисляване на собствени стойности и собствени вектори се използват различни подходи като точни, итерационни и стохастични, всеки от които е подходящ в зависимост от характеристиките на разглежданата квадратна матрица. Точни (директни) се наричат методите, при които при точно зададени начални данни и при идеално точни аритметични операции (т.е. без закръгляване) търсеното решение се получава след краен брой аритметични стъпки. Като примери могат да се посочат решаването на характеристичния полином при матрици с малка размерност, QR декомпозицията и други [114, 117]. Изчислителната сложност на този клас методи обикновено е от порядък $\mathcal{O}(n^3)$, където n е размерността на матрицата. Основен практически недостатък на точните методи е, че при матрици с голяма размерност аритметичните операции се извършват с крайна точност. Това неизбежно води до закръгляване и натрупване на числени грешки. В резултат може да се наруши точността на крайния резултат при пресмятането на собствените стойности.

При итерационните методи на всяка стъпка се получава едно следващо приближение на решението като точното решение се явява граница на редицата от последовател-

ните приближения. Примери за итерационни методи са: Степенният метод (англ. *Power method*), Обратният Степенен метод (англ. *Inverse method*), методът на Якоби (англ. *Jacobi method*), методът на Арнолди (англ. *Arnoldi method*) и други [39, 51]. Изчислителната сложност на този клас методи обикновено е от порядък $\mathcal{O}(n^2k)$, където n е размерността на матрицата, а k е броят на итерациите. Итерационните методи при матрици с големи размерности имат недостатък, породен от бавна сходимост, особено когато собствените стойности са близки една до друга или задачата е лошо обусловена. Резултатът зависи силно от началното приближение и може да изисква много итерации, което води до натрупване на числени грешки. При методите на Ланцош/Арнолди (Lanczos/Arnoldi) допълнителни проблеми са нуждата от памет и „скъпата“ ортогонализация. Те са изчислително интензивни и по-малко подходящи за паралелизация, което на практика ограничава размерността на матриците.

Стохастичните числени методи преодоляват значителна част от посочените по-горе ограничения. Поради стохастичната си природа те притежават редица предимства спрямо детерминистичните числени методи, а именно:

- Зависимостта между размерността на задачата (например, размерността на матрицата) и необходимата компютърна памет обикновено е линейна. Това позволява прилагането им към матрици с много големи размерности, съдържащи хиляди или милиони елементи.
- Съответните алгоритми са относително лесни за реализация и притежават вътрешно присъщ паралелизъм, тъй като изчисленията по различни траектории могат да се извършват независимо.

Поради по-бавната сходимост в сравнение с детерминистичните итерационни методи, стохастичните подходи често ”жертват” част от прецизността за сметка на по-висока изчислителна и паметова ефективност, както и възможност за решаване на задачи с голяма размерност. Това ги прави подходящи за оценяване на екстремални собствени стойности на симетрични матрици с големи размерности.

С развитието на високопроизводителните изчислителни системи (англ. *High-Performance Computing*, HPC) и с непрекъснатото въвеждане на нови хардуерни решения в тях, нараства необходимостта от разработване на софтуерни пакети, които интегрират ефективни стохастични алгоритми за решаване на трудоемки задачи, дефинирани в пространства с висока размерност. В частност, това се отнася до задачи за оценяване на екстремални собствени стойности на матрици с големи размерности [23, 29, 73].

В този контекст изследователите в областта на стохастичните методи и алгоритми за задачи от линейната алгебра продължават да се фокусират върху повишаване на точността, ускоряване на сходимостта и разширяване на приложимостта на алгоритмите за намиране на собствени стойности, особено при сложни задачи, възникващи в научните изчисления, машинното обучение и анализа на данни. Съобразно технологичния напредък на изчислителните системи нараства и необходимостта да се разработват и изследват по-бързи стохастични алгоритми, които осигуряват необходимата точност при решаване

на приложни задачи, свързани с оценяване на собствени стойности на матрици с големи размерности.

Обзор на основните резултати в областта

Първите публикации, свързани с Монте Карло (МК) методите в линейната алгебра, датират от 50-те и 60-те години на ХХ век, като първоначално не е разглеждан пряко проблемът за приближено изчисляване на собствени стойности. Задачата за намиране на собствените стойности и собствените вектори на матрици обикновено възниква при решаване на системи от линейни алгебрични уравнения (СЛАУ) и стохастични диференциални уравнения (СДУ). Ако матрицата е с голяма размерност, не е необходимо да се интересуваме от всичките ѝ собствени стойности. В повечето случаи свойствата на системата, описани чрез матрицата, се определят от най-голямата и най-малката ѝ собствена стойност. Двете екстремални стойности определят числото на обусловеност на матрицата [79, 114, 117]. Много задачи от реалния свят, описани в научните области, като обработка на изображения и сигнали, системи за управление, машинно обучение и квантова механика, изискват оценка на екстремалните собствени стойности на дадени матрици. Например, енергийните нива на квантовите системи се определят от собствените стойности на Хамилтоновия оператор. Ако са известни стойностите на най-големите и най-малките собствени числа, могат да бъдат представени енергията на основното състояние и възбудените състояния на системата [80, 93]. Някои от популярните детерминистични методи, които се използват за оценка на собствените стойности са: Степенният метод (англ. *Power method*), итерацията на коефициента на Релей (англ. *coefficient Rayleigh*), Резолвентният Степенен метод (англ. *Resolvent Power method*) и алгоритъмът на Ланцош (англ. *Lanczos algorithm*) [39, 51].

Монте Карло методът (МКМ) за намиране на доминиращата собствена стойност е предложен от Собол (Sobol) през 1973 година [119], а за най-малката от Михайлов през 1987 година [66]. Степенният Монте Карло (СМК) метод е предложен от Димов и Караиванова в поредица от публикации през периода 1996-2000 година [20, 23, 26, 28, 29, 30]. През последните десетилетия итеративните алгоритми на Монте Карло (МК) и Квази-Монте Карло (КМК) за намиране на екстремалните собствени стойности, базирани на Степенния метод и Резолвентния Степенен метод, са разработени от учени, работещи в тази сфера [3, 24, 26, 30, 74]. Използването на резолвентната матрица за намиране на най-малката собствена стойност е предложено от Димов и Караиванова през 1998 година [29]. Първите публикации в литературата, представящи Квази-Монте Карло (КМК) методи за намиране на екстремални собствени стойности, са на Караиванова и Маскани в периода 2001-2003 година [72, 73, 74, 75]. В публикациите на Александров, Караиванова, Ивановска, Чи, Маскани и други [2, 3, 15, 16, 52, 57] са разглеждани използването на различни редици с малък дискрепанс (генерализирани, рандомизирани) в Монте Карло (МК) методите за намирането на екстремни собствени стойности. Квази-Монте Карло (КМК) методите използват редици с малък дискрепанс като редици на Собол (Sobol, [96, 97, 98]), Холтън (Halton, [45, 46]) или Фор (Faure, [34, 35, 36]), които предлагат по-равномерно и детерминистично покритие на пробното пространство, вместо да използват случайни проби.

В сравнение със стандартните Монте Карло (МК), Квази-Монте Карло (КМК) методите притежават някои предимства като гъвкавост и обобщеност, приложимост към матрици с големи размерности, достъпност, простота и лесно паралелизиране на алгоритъма [21, 27, 75].

През последните 20 години са разработени различни паралелни методи относно Монте Карло алгоритмите. Стратегии за паралелизиране могат да бъдат открити в следните научни трудове [4, 5, 11, 100, 101, 109]. Паралелни Монте Карло методи за СЛАУ, базирани на итерация на Якоби чрез Монте Карло алгоритъм, са представени от Димов [21]. Паралелни Монте Карло методи, използващи минимални вериги на Марков и минимални комуникации, са представени в [33, 104]. Повечето от горните подходи се основават на идеята за балансиране на стохастичната и систематичната грешка [21]. Монте Карло алгоритмите, които се използват при решаването на системи линейни алгебрични уравнения (СЛАУ), могат да бъдат паралелизирани [22, 23]. Хибридни Монте Карло алгоритми за матрични пресмятания са представени от Ваджарга [110], Александров, Атанасов и Димов [1], както и от други автори. За специфични задачи от линейната алгебра, като намирането на екстремалните собствени стойности на дадена матрица, могат да се използват както директни методи за пресмятане на решението, така и Монте Карло алгоритми. Двата типа подходи могат да бъдат паралелизирани. Статии, описващи паралелизацията на Монте Карло алгоритми за намиране на най-голямото и най-малкото собствено число на матрица, са представени от Димов, Маскани и Караиванова [25, 29, 73]. С развитието на високопроизводителните хетерогенни изчислителни системи възникна необходимост от разработване на нови подходи за паралелно генериране на квазислучайни редици. В отговор на тази потребност бяха създадени ефективни алгоритми за генериране на квазислучайни редици (редиците на Соболев и Холтън), реализирани върху сървърни системи от платформата HP Cluster SL250S Gen8 с процесори Intel Xeon E2650v2 и копроцесори Intel Xeon Phi 7120P, които формират хетерогенни изчислителни архитектури [7, 10].

Цели и задачи на дисертация

Основната цел на дисертационния труд е да се предложат и изследват стохастични числени методи за оценяване на собствени стойности като се разработят и анализират Монте Карло и рандомизирани Квази-Монте Карло алгоритми за приближено пресмятане на екстремалните собствени стойности на симетрични квадратни матрици. За постигане на основната цел са формулирани следните задачи:

1. Да се разгледат и изучат Степенен (Резолвентен) Монте Карло метод и рандомизиран Степенен (Резолвентен) Квази-Монте Карло метод за оценяване на екстремалните собствени стойности на симетрични матрици, както и да се анализира баланса между възникващите грешки - стохастична и систематична.
2. Да се разработят и изследват ефективни Монте Карло и рандомизирани Квази-Монте Карло алгоритми, базирани на тези стохастични методи за числено пресмятане на екстремалните собствени стойности на симетрични квадратни матрици, като

се използват генератори на псевдослучайни числа и генератори на редици с малък дискрепанс.

3. Да се проведат числени експерименти с разработените алгоритми, като се обоснове изборът на специализиран софтуер за генериране на редици с малък дискрепанс, в частност редици на Соболев и Холтън, и числено да се изследва балансът между стохастичната и систематичната грешка в зависимост от определящите ги параметри.
4. Да се формулира и изследва задача от областта на финансовата математика, свързана с управлението на пазарния риск на портфейл от активи, при който оценката на риска се основава на определянето на максималната собствена стойност на корелационна матрица. Да се приложат разработените в Глава 1 Степенен Монте Карло и рандомизиран Степенен Квази-Монте Карло алгоритми за числено оценяване на най-голямата собствена стойност и да се анализира тяхната ефективност при оценяване на пазарния риск.

Методология на изследването

Методологията на изследването в настоящия дисертационен труд се основава на фундаментални научни резултати от следните области:

- Теория на вероятностите и математическа статистика - вероятностни пространства, случайни величини и техни характеристики като функции на разпределение, функция на плътност, математическо очакване, дисперсия, закон за големите числа, централна гранична теорема; вероятностна грешка, доверителен интервал, верига на Марков, извадки на случайни величини, статистически тестове за изследване на генератори на псевдослучайни величини.
- Математически анализ - редици и редове, сходимост на редове, критерии за сходимост, систематична грешка, норми и метрични пространства.
- Числени методи, линейна и изчислителна алгебра - итерационни методи за СЛАУ и собствени стойности, матрични норми, оптимизационни методи, алгоритми за генериране на псевдослучайни числа, стохастични числени методи (Монте Карло и Квази-Монте Карло методи).
- Теория на числата и по-точно раздела за дискретна и аналитична теория на числата - равномерно разпределени редици, дискрепанс (англ. *discrepancy*), радикално-инверсни функции, разлагане по прости основи.
- Информатика/компютърни науки - алгоритми и структури от данни, програмни кодове написани на MATLAB, C++ с използване на генератори на псевдослучайни числа като MT и MS; библиотеки за рандомизирани редици на Соболев и Холтън.

Част от числените експерименти са извършени на високопроизводителния клъстър, който се състои от 12 сървъра Fujitsu Primergy RX 2540 M4. Всеки сървър е конфигуриран с GPU карта от вида NVIDIA Tesla V100 32GB и CPU 2 x Intel Xeon Gold 5118 2.30GHz 24 cores, всеки с 128 GB RAM. Редиците на Соболев са получени чрез BRODA's Sobol Randomized Sequence Generator (RSG) [12].

В следващите секции сме формулирали основната задача за оценяване на собствени стойности на симетрична матрица, описали сме МК и КМК подходите за числено решаване на многомерни задачи и грешките, с които се характеризират тези методи.

Формулиране на задачата

Нека разгледаме матрицата $A = \{a_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}$, $a_{ij} = a_{ji}$, $i, j = 1, \dots, n$. Задачата за намиране на собствени стойности $\lambda(A)$ на несингулярна симетрична матрица се формулира по следния начин:

$$A\mathbf{x} = \lambda(A)\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n \text{ е произволен ненулев вектор.} \quad (1)$$

Ще отбележим, че всички собствени стойности на симетричните матрици са реални числа. Без ограничение на общността предполагаме следната наредба на собствените стойности на A :

$$\lambda_{min} = |\lambda_n| \leq |\lambda_{n-1}| \leq \dots \leq |\lambda_2| < |\lambda_1| = \lambda_{max}. \quad (2)$$

Матрицата A може да се представи векторно и по следния начин:

$$A = \{a_{ij}\}_{i,j=1}^n = (\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_n)^T, \text{ където } \mathbf{a}_i = (a_{i1}, \dots, a_{in}), \quad i = 1, \dots, n.$$

Символът T означава операцията *транспониране*. За нуждите на изследването се използват следните векторни и матрични норми:

$$\|\mathbf{h}\| = \|\mathbf{h}\|_1 = \sum_{i=1}^n |h_i|, \quad \|\mathbf{a}_i\| = \|\mathbf{a}_i\|_1 = \sum_{j=1}^n |a_{ij}|, \quad \|A\| = \|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|, \quad j = 1, \dots, n.$$

Прилагането на стохастични числени методи за оценка на екстремалните собствени стойности λ_{min} и λ_{max} на така дефинираната задача включва разработването и изследването на Степенни (Резолвентни) Монте Карло и рандомизирани Степенни (Резолвентни) Квази-Монте Карло алгоритми за тяхното приближено пресмятане. Под стохастични числени методи по-нататък в дисертационния труд ще разбираме Монте Карло методите и тяхната разновидност, рандомизираните Квази-Монте Карло методи.

Монте Карло методи

При методите МК се конструира случайна величина θ , която всъщност представлява измерима функция $\theta = \theta(\omega)$, дефинирана във вероятностното пространство $(\Omega, \mathfrak{A}, \mathbb{P})$ [115] с дефиниционна област Ω и с област на стойностите \mathbb{R} (множеството на реалните числа). В теорията на вероятностите, множеството Ω се нарича *пространство на елементарните*

събития, \mathfrak{A} – алгебра на събитията или булова алгебра, (Ω, \mathfrak{A}) – измеримо пространство, $(\mathfrak{A}, \mathbb{P})$ – поле на вероятностите. Вероятността \mathbb{P} се дефинира като числова функция $\mathbb{P}(\cdot)$ в измеримото пространство с дефиниционна област в \mathfrak{A} и функционални стойности в числовия интервал $[0, 1]$ [115, 118].

Условието при конструирането на случайна величина θ е свързано с математическото ѝ очакване, като изискването е математическото очакване да съществува и да е равно на търсеното решение на дадена задача или на функционал от решението на изходната задача, т.е.:

$$E[\theta] = I, \quad (3)$$

където I е търсеното решение (или стойността на функционал от решението). По определение $E[\theta]$ съществува тогава и само тогава, когато съществува $E|\theta|$. В зависимост от вида на математическата задача, случайната величина θ може да бъде непрекъснатата или дискретна случайна величина.

Математическото очакване на дискретна случайна величина θ се дефинира с равенството:

$$E[\theta(\omega)] = \sum_i \mathbb{P}\{\theta(\omega) = x_i\}x_i = \sum_i p_i x_i = I, \text{ където } \sum_i p_i = 1.$$

Тук p_i означава вероятността, с която случайна величина получава стойността x_i .

В непрекъснатия случай, математическото очакване на случайна величина се изразява с равенството:

$$E[\theta] = \int_{-\infty}^{\infty} xp(x) dx, \text{ където } \int_{-\infty}^{\infty} p(x) dx = 1,$$

Неотрицателната функция $p(x) \geq 0$ се нарича плътност на разпределение (англ. *probability density function*, pdf).

Друга характеристика на случайните величини е *функцията на разпределение (ф.р.)*. Функцията на разпределение на случайна величина θ се дефинира с равенството: $F_\theta(x) = \mathbb{P}\{\theta < x\}$. Връзката между $p(x)$ и $F(x)$ за непрекъснати случайни величини е:

$$F(x) = \int_{-\infty}^x p(y)dy \text{ и } p(x) = \frac{dF(x)}{dx}.$$

Стойността на функцията на разпределение показва натрупаната вероятност до дадената точка, докато с плътността на разпределение се изразява скоростта, с която се натрупва тази вероятност.

Нека $\theta_1, \dots, \theta_i, \dots, \theta_N$ са независими наблюдения (извадка с размерност N) на случайната величина θ . За приближена стойност на I при МК методите разбираме средната стойност:

$$\bar{\theta}_N = \frac{1}{N} \sum_{i=1}^N \theta_i. \quad (4)$$

Оценка за грешката при Монте Карло методите

За да се оцени грешката при МК методите се въвеждат следните дефиниции и теореми.

Дефиниция 0.0.1. (*Сходимость по вероятности*) (англ. *in probability (p.)*)

Казваме, че редица от случайни величини $\theta_1, \theta_2, \dots, \theta_N, \dots$ клони по вероятност към случайна величина θ , ако за всяко $\varepsilon > 0$ е изпълнено:

$$\mathbb{P}\{|\theta_N - \theta| > \varepsilon\} \rightarrow 0, \quad N \rightarrow \infty.$$

Означава се $\theta_N \xrightarrow{p.} \theta$ или $\lim_{N \rightarrow \infty} \theta_N = \theta (p.)$.

Дефиниция 0.0.2. (*Почти сигурна сходимость*) (англ. *almost sure (a.s.)*)

Казваме, че редица от случайни величини $\theta_1, \theta_2, \dots, \theta_N, \dots$ е сходяща с вероятност единица или почти сигурно към случайна величина θ , ако е изпълнено:

$$\mathbb{P}\{\lim_{N \rightarrow \infty} \theta_N = \theta\} = 1.$$

Означава се $\theta_N \xrightarrow{a.s.} \theta$ или $\lim_{N \rightarrow \infty} \theta_N = \theta (a.s.)$.

Дефиниция 0.0.3. (*Закон за големите числа (ЗГЧ)*)

Нека $\theta_1, \dots, \theta_i, \dots, \theta_N, \dots$ е редица от случайни величини дефинирана във фиксирано вероятностно пространство $(\Omega, \mathfrak{A}, \mathbb{P})$. Въвеждаме означенията $E[\theta_i] = I_i$, $\sum_{i=1}^N I_i = NI$ и $\sum_{i=1}^N \theta_i = N\bar{\theta}_N$. Ще казваме, че тази редица удовлетворява слабия закон за големите числа (СЗГЧ), ако за всяко $\varepsilon > 0$ е изпълнено:

$$\lim_{N \rightarrow \infty} \mathbb{P}\{|\bar{\theta}_N - I| < \varepsilon\} = 1,$$

т.е. при $N \rightarrow \infty$, следва $\bar{\theta}_N \xrightarrow{p.} I$. Ще казваме, че тази редица удовлетворява усиления закон за големите числа (УЗГЧ), ако е изпълнено:

$$\mathbb{P}\{\lim_{N \rightarrow \infty} (\bar{\theta}_N - I) = 0\} = 1,$$

т.е. при $N \rightarrow \infty$, следва $\bar{\theta}_N \xrightarrow{a.s.} I$.

Теорема 0.0.1. (*Теорема на Хинчин*) [19, 115]

Ако $\theta_1, \dots, \theta_i, \dots, \theta_N, \dots$, е редица от независими и еднакво разпределени случайни величини, за които математическото очакване съществува, т.е. $E[\theta_i] < \infty$, то тази редица удовлетворява СЗГЧ.

Теорема 0.0.2. (*Теорема на Колмогоров*) [19, 115]

Нека $\theta_1, \dots, \theta_i, \dots, \theta_N, \dots$, е редица от независими и еднакво разпределени случайни величини. Съществуването на математическото очакване $E[\theta_i]$ е необходимо и достатъчно условие тази редица да удовлетворява УЗГЧ.

Отново се разглежда равенство (4). Извадката $\theta_1, \dots, \theta_i, \dots, \theta_N$ на случайната величина θ може да се разгледа като редица от независими и еднакво разпределени случайни величини, за които математическото очакване съществува. Теоремата на Хинчин се прилага като

се получава, че средната стойност от тези случайни величини клони към математическото очакване на θ , т.е.:

$$\bar{\theta}_N \xrightarrow{p} E[\theta] = I, \quad \text{когато } N \rightarrow \infty,$$

т.е. при достатъчно голямо N следва, че:

$$\bar{\theta}_N \approx I. \quad (5)$$

По този начин, ако се конструира случайна величина с математическо очакване $E\theta = I$, може да се приложи равенството (4), за да се оцени приближено стойността I на решението или на функционал от решението на дадена задача.

Каква е грешката на МК методите? Нека се предположи, че случайната величина θ има крайна дисперсия (англ. *variance*), т.е.:

$$\text{Var}[\theta] = E[\theta - E\theta]^2 = E[\theta^2] - (E[\theta])^2 < \infty.$$

Ако вторият момент на случайната величина съществува, т.е. $E[\theta^2] < \infty$, то дисперсията е крайна.

Теорема 0.0.3. (Централна гранична теорема (ЦГТ)) [37, 115]

Нека $\theta_1, \dots, \theta_i, \dots, \theta_N$ е редица от независими и еднакво разпределени случайни величини с $E[\theta_i] = I$ и $\text{Var}[\theta_i] = \sigma^2 < \infty$. Тогава за всяко $x \in \mathbb{R}$ е в сила равенството:

$$\lim_{N \rightarrow \infty} \mathbb{P} \left\{ \frac{\sum_{i=1}^N (\theta_i - I)}{\sigma\sqrt{N}} < x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt. \quad (6)$$

При прилагане на ЦГТ за произволни x_1 и x_2 се получава равенството:

$$\lim_{N \rightarrow \infty} \mathbb{P} \left\{ x_1 < \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (\theta_i - I) < x_2 \right\} = \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{t^2}{2}} dt. \quad (7)$$

След полагане $x = x_2 = -x_1$ и при заместване в (7) следва:

$$\lim_{N \rightarrow \infty} \mathbb{P} \left\{ \left| \frac{1}{N} \sum_{i=1}^N (\theta_i - I) \right| < x\sigma N^{-1/2} \right\} = \Phi(x),$$

където $\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$ се нарича вероятностен интеграл и се определя от функцията на разпределение на стандартна нормална случайна величина в интервала $(0, 1)$ [119].

Следователно, когато N е достатъчно голямо се получава:

$$\mathbb{P} \{ |\bar{\theta}_N - I| < x\sigma N^{-1/2} \} \approx \Phi(x). \quad (8)$$

Формула (8) определя цяла фамилия от оценки, които зависят от параметъра x . Ако вероятността β (т. нар. *коэффициент на доверие*) е зададена, тогава може да се намери

решението $x = x_\beta$ (т. нар. *ниво на значимост*) на уравнението:

$$\Phi(x) = \beta, \quad (9)$$

използвайки статистически таблици [119]. От израза (9) се получава, че с вероятност равна на β е изпълнено неравенството:

$$|\bar{\theta}_N - I| < x_\beta \sigma N^{-1/2} = r_N. \quad (10)$$

Величината r_N в неравенството (10) се нарича *вероятностна грешка*. Така задавайки стойност на β в $(0, 1)$ се определя с каква вероятност разликата $(\bar{\theta}_N - I)$ попада в доверителния интервал $(-r_N, r_N)$. Най-често използваните коефициенти на доверие са: (1) при $\beta = 0.95$, получаваме $x_\beta = 1.96$; (2) при $\beta = 0.997$, $x_\beta = 3$. Този случай е известен като (*правило на трите сигми*). Важна характеристика за грешката се получава, когато $\beta = 0.5$. В този случай r_N се нарича *вероятната грешка*, а $x_{0.5} \approx 0.6745$. За тези стойности получаваме:

$$\mathbb{P}\{|\bar{\theta}_N - I| \leq r_N\} = \frac{1}{2} = \mathbb{P}\{|\bar{\theta}_N - I| \geq r_N\}, \quad (11)$$

където $r_N = 0.6745\sigma N^{-1/2}$. При методите Монте Карло чрез вероятната грешка се определя скоростта на сходимост на метода. Математическото обяснение е, че с вероятност 0.5 разликата между приближеното решение и точното решение на задачата $(\bar{\theta}_N - I)$ попада в доверителния интервал $(-r_N, r_N)$. Със същата вероятност тази разлика може да попадне извън интервала.

Оценката получена в (10) показва, че вероятностната грешка зависи от *стандартното отклонение* $\sigma = \sqrt{Var[\theta]}$, т.е. от дисперсията и от броя на реализациите N на случайна величина (големината на извадката).

При оценяване на вероятностната грешка, рядко се получава теоретична оценка за дисперсията $Var[\theta]$. В случаите, когато дисперсията на случайната величина θ е ограничена, може да се оцени емпирично в хода на получаване на извадката от N реализации на случайната величина. Като се има в предвид формулите (3), (4) и (5), за втория момент на случайната величина θ се получава:

$$E[\theta^2] \approx \frac{1}{N} \sum_{i=1}^N (\theta_i)^2,$$

От тук за приближена стойност на дисперсията се получава:

$$Var[\theta] \approx \frac{1}{N} \sum_{i=1}^N (\theta_i)^2 - (\bar{\theta}_N)^2. \quad (12)$$

При неголяма размерност на извадката N от стойности на случайната величина, обикно-

вено се използва по-точната формула [119]:

$$\text{Var}[\theta] \approx \frac{1}{(N-1)} \sum_{i=1}^N (\theta_i)^2 - \frac{N}{N-1} (\bar{\theta}_N)^2. \quad (13)$$

Различават се два типа МК методи, които се използват за намаляване на грешка.

- **Ефективни МК методи** [27, 116, 119]: Целта тук е да се намери случайна величина θ , принадлежаща на класа от случайни величини удовлетворяващи равенство (3), но да е с минимална дисперсия или с дисперсия близка до минималната. Вероятната грешка продължава да е от порядъка $r_N = \mathcal{O}(N^{-1/2})$. При МК методите за оценяване на многомерни интеграли са познати разнообразни подходи за намаляване на дисперсията (англ. *variance reduction approaches*), като: отделяне на главна част, симетризация на подинтегралната функция, метод на съществената извадка, метод на разделяне по важност и други.
- **Суперсходящи МК методи** [119]: При тях порядъка на вероятната грешка се подобрява и се записва във вида $r_N = \mathcal{O}(N^{-1/2-\varepsilon(k)})$, където $\varepsilon(k) > 0$, а k е размерността на задачата (например, при оценяване на k -мерни определени интеграли). При този клас Монте Карло (МК) методи се извършва симулиране на стойности на случайна величина в равномерни, взаимно неприпокриващи се подобласти, на които е разделена областта на интегриране, като допълнително се налага условие за гладкост на подинтегралната функция. Този подход е известен като стратификация (англ. *stratified sampling*).

Дефиниция 0.0.4. *Процесът на пресмятане стойностите $\theta_1, \theta_2, \dots, \theta_i, \dots$ на случайна величина θ (т.е. получаването на извадка с определена размерност), чрез използване на вероятностите p_i в дискретния случай или на плътността на разпределение $p(x)$ в непрекъснатия случай, се нарича симулация (или моделиране).*

Съществуват различни подходи за симулиране на стойностите на случайни величини, като най-прилагани са: методът на обратната функция (англ. *inverse transform method*), методът на селекцията (англ. *acceptance-rejection method*), методът на съществената извадка (англ. *importance sampling*) и други [56, 60, 116, 119]. При практическата реализация на тези методи основна роля играят генераторите на псевдослучайни числа, които се използват за симулиране стойностите на дадена случайна величина.

Генератори на псевдослучайни числа

За получаването на извадката (англ. *sampling*) от стойности на случайна величина θ при МК методите се използват генератори на случайни числа. Тези генератори произвеждат независими и равномерно разпределени числа в $(0, 1)$, които са стойности на случайна величина η , с $E\eta = 1/2$ и $\text{Var}(\eta) = 1/12$. В практиката, най-често се използват генератори на псевдослучайни числа (англ. *pseudorandom number generators*, PRNGs). Те представляват компютърни програми, използващи математически формули за генериране на дълги поредици (редици) от статистически некорелирани числа в интервала $(0, 1)$ [64, 103].

Обикновено при стартиране на PRNG се изисква да се въведе начално състояние (число), с което да се започне генерирането на редицата от псевдослучайни числа. Това първоначално състояние се нарича „зародиш“ или „семе“ (англ. *seed*). Добрите PRNGs трябва да притежават определени свойства [60]. Ето някои от тях:

- **Случайност, равномерно разпределеност в $(0, 1)$, независимост**

Тези свойства се проверяват с различни статистически тестове. Най-често използваните са: (1) χ^2 (хи-квадрат тест) и ω^2 (омега-квадрат тест) [115, 119]; (2) тестове за празни интервали (англ. *gap test*), за липса на корелация; (3) тест на пробези (англ. *runs test*), т.е. анализира се дължината и броя на нарастващи/намаляващи или бинарни последователности; (4) тест за най-дълга поредица (англ. *longest run test*), т.е. проверява се дължината на най-дългата поредица от еднакви стойности; (5) спектрален тест, при който се анализира геометричната структура на точките в многомерно пространство.

- **Възпроизводимост, бързина и ефективност**

Генерирането на числата трябва да става с малък разход на изчислителни ресурси (време) и компютърна памет. При едно и също начално състояние генераторът трябва да произвежда една и съща редица от числа.

- **Дълъг период на повтаряемост, без генериране на 0 и 1**

Колкото по-дълъг е периодът на повтаряемост, толкова генераторът е по-добър. Не трябва да се допуска деление на 0 или други числени усложнения.

- **Възможност за паралелно генериране на подредици без препокриване (англ. *no overlapping*)**

При съвременното развитие на високопроизводителните компютърни системи, това свойство дава редица предимства при решаване на трудоемки задачи с методите МК.

В статията на Марсалия [69] са описани дванадесет теста за оценка на качеството на генератори на псевдослучайни числа, които са включени в пакет *“Diehard tests”*. Чрез различни статистически критерии се проверяват дали генерираните числа се държат като истински случайни числа равномерно разпределени в $(0, 1)$. През последните десетилетия са разработени разнообразни генератори на псевдослучайни числа, които Марсалия класифицира в 6 категории [70]. Най-разпространените и често използвани генератори на псевдослучайни числа са:

- **Конгруентни генератори, основани на модулната аритметика**

Този клас генератори е най-изследваният в последните няколко десетилетия, като най-разпространените за употреба са: (1) Линеен конгруентен генератор (англ. *Linear congruential generator*) с параметри препоръчани от Fishman (*LCG-F*) [38]; (2) Линеен конгруентен генератор с параметри препоръчани от Park и Miller (*LCG-PM*) [91]; (3) Комбиниран линеен конгруентен генератор (англ. *Combined linear congruential generator*) с параметри от P. L’Ecuyer (*CLCG-PL*) [65]; (4) Инверсивен конгруентен

генератор (ICG) (англ. *Inversive congruential generator*, ICG) [48]; и (5) Експлицитен Инверсивен конгруентен генератор (англ. *Explicit inversive congruential generator*, EICG) [31].

- **Генератори с изместени регистри**

Два от най-използваните генератора са: (1) *Shuffled nested Weyl sequences (SNWS)*, генератор създаден на основата на разбърканите вложени редици на Weyl, успешно използван при симулации в областта на молекулярната динамика [50, 113] и (2) *Mersenne Twister PRNG*, създаден от Matsumoto и Nishimura [78] и използван в разнообразни софтуери като MATLAB, MAPLE, MICROSOFT EXCEL. Използва се като генератор по подразбиране (*default*) в програмни езици като PHP, PYTHON, LISP, JULIA и други.

- **PRNG генератори, използващи други математически формули** различни от описаните по-горе

Към този клас генератори можем да причислим генератора на Фибоначи (англ. *Lagged Fibonacci Generator*, LFG) [71, 76] и генератора на Джон фон Нойман (*Middle Square PRNG*) [82], основаващ се на метода на средните квадрати (*Middle-Square method*) [90].

Създадена е библиотека, съдържаща значително разнообразие от генератори на псевдослучайни числа, наречена *Scalable Parallel Random Number Generator Library (SPRNG)* [71], която учените я използват при разработване на паралелни МК алгоритми.

В нашите изследвания за оценяване на екстремалните собствени стойности чрез Степенния МК метод и Резолвентния МК метод сме използвали *Middle Square PRNG* [90] и *Mersenne Twister PRNG* [78] за симулиране стойностите на съответните случайни величини.

Квази-Монте Карло методи

Монте Карло методите, макар широко приложими, зависят от псевдослучайни числа с бавни скорости на сходимост, обикновено от порядъка на $O(N^{-1/2})$, където N е броят на пробите (извадката на конструираната случайна величина).

Квази-Монте Карло (КМК) методите са разработени, за да се повиши скоростта на сходимост. Ключовата разлика с МК методите е, че при КМК методите вместо случайни точки (генерирани с PRNGs) се използват детерминистични множества от точки, известни като редици с малък дискрепанс (англ. *low discrepancy sequences*, LDS). Тези редици са създадени, за да осигурят по-равномерно покритие в единичния хиперкуб $[0, 1]^k$ [45, 83, 96]. Теоретично LDS редиците са изследвани за оценяване на многомерни определени интеграли и е доказано, че чрез тях се достига до по-добри скорости на сходимост от порядъка $O(N^{-1}(\log N)^k)$ [14, 62, 83]. При многомерни определени интеграли приближеното решение получено с КМК метод се дефинира чрез формулата:

$$\int_{[0,1]^k} f(\mathbf{x})d\mathbf{x} \approx \frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_i). \quad (14)$$

Тук $\mathbf{x}_i \in (0, 1]^k$, $i = 1, \dots, N$ са елементи от редица с малък дискрепанс, в която всеки елемент е вектор с k на брой координати. Ако координатите на тези вектори се генерират чрез PRNGs, получаваме приближено решение по метода МК като статистическата грешка се оценява с ЦГТ.

Редици с малък дискрепанс

Дефиниция 0.0.5. (*равномерна разпределеност*) (англ. *equidistributed*)

Казваме, че редицата от точки $S = \{\mathbf{x}_i\}_{i=1}^{\infty} \in [0, 1]^k$ се нарича равномерно разпределена, ако за всеки осово ориентиран паралелепипед в единичния k -мерен хиперкуб,

$$\mathbf{B} = \prod_{j=1}^k [a_j, b_j] \subset [0, 1]^k,$$

относителната честота на точките попадащи в \mathbf{B} , клони към неговия k -мерен обем $Vol(\mathbf{B})$:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{(\mathbf{B})}(\mathbf{x}_i) = Vol(\mathbf{B}). \quad (15)$$

Тук $\mathbf{1}_{(\mathbf{B})}(\mathbf{x}_i)$ е индикаторна функция за \mathbf{B} , която приема стойност 1, ако $\mathbf{x}_i \in \mathbf{B}$ и стойност 0 в противния случай. Мярката за липса на равномерна разпределеност на безкрайна редица от точки в единичния хиперкуб се нарича *дискрепанс*. Най-широко изследваните мерки за дискрепанс се изразяват чрез L_2 -нормата или чрез L_∞ -нормата. В първия случай говорим за L_2 -дискрепанс, а във втория за *дискрепанс-звезда*.

Дефиниция 0.0.6. (*дискрепанс-звезда*) (англ. *star discrepancy*)

За всяка редица $S = \{\mathbf{x}_i\}_{i=1}^N \in (0, 1]^k$ и всеки осово ориентиран хиперпаралелепипед $\mathbf{Z} = \prod_{j=1}^k [0, z_j] \subset [0, 1]^k$, дискрепанс-звезда на редицата, $D_N^*(S)$, измерва максималното отклонение между емпиричното разпределение и равномерността въведена в (15) и се дефинира с равенството:

$$D_N^*(S) = \sup_{\mathbf{Z} \in [0, 1]^k} \left| \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\mathbf{Z}}(\mathbf{x}_i) - \prod_{j=1}^k z_j \right|. \quad (16)$$

Дефиниция 0.0.7. (*редица с малък дискрепанс*) (англ. *low discrepancy sequence, LDS*)

Казваме, че за всяко $N > 1$, редицата $S = \{\mathbf{x}_i\}_{i=1}^N$ в $(0, 1]^k$ е редица с малък дискрепанс, ако е изпълнено:

$$D_N^*(S) = \mathcal{O}(N^{-1}(\log N)^k) \text{ или } D_N^*(S) \leq C_k \frac{(\log N)^k}{N}. \quad (17)$$

Константата C_k в (17) зависи само от размерността k .

Оценка за грешката при Квази-Монте Карло методите

Грешката при КМК методите се дефинира като разлика между приближеното и точното решение $I(f)$ на многомерен определен интеграл, съгласно формула (14). Тя се оценява

чрез неравенството на Коксма-Хлавка [14, 92], но нека първо да бъдат въведени следните две дефиниции [92].

Дефиниция 0.0.8. (*вариация в смисъл на Витали*) (англ. *variation in sense of Vitali*)
Нека функцията $f(x)$ е дефинирана в единичния хиперкуб $[0, 1]^k$ и $\mathbf{B} \subseteq [0, 1]^k$ е област от вида $\mathbf{B} = \prod_{j=1}^k [a_j, b_j] \subset [0, 1]^k$. Дефинираме оператора $\Delta(f(x); \mathbf{B})$ като комутираща сума от стойностите на $f(x)$ във върховете на \mathbf{B} :

$$\Delta(f; \mathbf{B}) = \sum_{j_1=0}^1 \dots \sum_{j_k=0}^1 (-1)^{\sum_{i=1}^k j_i} f(j_1 a_1 + (1 - j_1) b_1, \dots, j_k a_k + (1 - j_k) b_k). \quad (18)$$

За множеството \mathcal{P} от подинтервали в $(0, 1)^k$, имащи вида \mathbf{B} , дефинираме вариация в смисъл на Витали чрез равенството:

$$V^{(k)}(f) = \sup_{\mathcal{P}} \sum_{\mathbf{B} \in \mathcal{P}} |\Delta(f; \mathbf{B})|. \quad (19)$$

Изразът (19) представлява мярка за променливостта на функцията $f(x)$. Ако $f(x)$ е k -кратна непрекъсната диференцируема функция дефинирана в областта $(0, 1)^k$, то вариацията в смисъл на Витали може да се запише и във вида:

$$V^{(k)}(f) = \int_0^1 \dots \int_0^1 \left| \frac{\partial^k f}{\partial x_1 \dots \partial x_k} \right| dx_1 \dots dx_k.$$

За да се намери горната граница на грешката на КМК метода, е необходимо да се разшири тази мярка.

Дефиниция 0.0.9. (*вариация в смисъл на Харди-Краус*) (англ. *variation in sense of Hardy-Krause*)

Нека функцията $f(x)$ е дефинирана в единичния хиперкуб $[0, 1]^k$. Нека $V^{(s)}(f; i_1, \dots, i_s)$ е вариация в смисъл на Витали за $f(x)$ ограничена до s -мерната област

$$\{(x_1, \dots, x_k) \in [0, 1]^k \mid x_j = 1 \text{ за } j \neq i_1, \dots, i_s\},$$

където $1 \leq s \leq k$ и $1 \leq i_1 < \dots < i_s \leq k$. Тогава

$$V(f) = \sum_{s=1}^k \sum_{1 \leq i_1 < \dots < i_s \leq k} V^{(s)}(f; i_1, \dots, i_s), \quad (20)$$

се нарича вариация на $f(x)$ в $[0, 1]$ в смисъл на Харди-Краус.

Ако $V(f)$ е крайна, тогава се казва, че функцията f е с ограничена вариация в смисъла на Харди-Краус. В сила е следната теорема [14, 92]:

Теорема 0.0.4. (*неравенство на Коксма-Хлавка*) (англ. *Koksma-Hlawka inequality*)

За всяка редица с малък дискрепанс $S = \{\mathbf{x}_i\}_{i=1}^N \in (0, 1]^k$, $N > 1$ и всяка функция $f(x)$ с ограничена в смисъл на Харди-Краус вариация $V(f)$, е изпълнено неравенството:

$$\left| I(f) - \frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_i) \right| \leq V(f) D_N^*(S), \quad (21)$$

където $D_N^*(S)$ е дискрепанс-звезда на редицата S .

Кафлиш в [14] е предложил просто и елегантно доказателство на неравенството (21) за случая на функции с ограничени k производни. Неравенството на Коксма-Хлавка показва, че грешката на КМК методите зависи от два фактора: (1) дискрепанса (несъответствието) на точките от редицата S и (2) вариацията на функцията $f(x)$. Ако подинтегралната функция е достатъчно гладка (например, тя може да бъде диференцируема поне веднъж по отношение на всяка променлива), тогава грешката при КМК методите е от порядък $\mathcal{O}(N^{-1}(\log N)^k)$ за редици с малък дискрепанс. Скоростта на сходимост може да бъде дори по-висока за периодични подинтегрални функции. Неравенството на Коксма-Хлавка дава теоретична оценка на грешката при КМК методите. При реални изчисления се срещат трудности при изчисляване на вариацията на функцията и на стойността за дискрепанс-звезда на съответната редица. Въпреки това порядъкът на грешката при КМК метода е асимптотически по-добър от този на съответния МК метод.

Първите конструкции на квазислучайни редици, т.е. редици с малък дискрепанс, съгласно дефиниция (0.0.7), са тясно свързани с редицата на Ван дер Корпут (*Van der Corput*) [111]. Тази квазислучайна редица е едномерна, дефинирана е в интервала $(0, 1)$ и е изградена чрез цифрова инверсия. Холтън (Halton) през 1960 г. [45] разширява идеята на Ван дер Корпут към произволна размерност чрез използване на различни прости основи по координати.

Нека се разгледа радикалната инверсия (англ. *radical inverse*) при основа b [111]. Известно е, че всяко цяло число $n \geq 0$ има единствено представяне при зададена основа $b \geq 2$ (цяло число):

$$n = \sum_{j=0}^{\infty} a_j b^j, \quad a_j \in \{0, 1, \dots, b-1\}. \quad (22)$$

Методът на радикалната инверсия „обръща“ числото n така:

$$\Phi_b(n) = \sum_{j=0}^{\infty} a_j b^{-(j+1)}. \quad (23)$$

По този начин числото $\Phi_b(n)$ се получава от n чрез симетрично завъртане на представянето (22) около “десетичната точка“. Нека се отбележи, че числото $\Phi_b(n) \in [0, 1)$ и при индекс j достатъчно голям е налице $a_j = 0$. Сега можем да дефинираме редица на Холтън.

Дефиниция 0.0.10. (редница на Холтън)

За дадена размерност $k \geq 1$ и за произволни взаимно прости цели числа b_1, b_2, \dots, b_k

по-големи от 1, редицата на Холтън се дефинира с равенството:

$$\mathbf{x}_n = (\Phi_{b_1}(n), \Phi_{b_2}(n), \dots, \Phi_{b_k}(n)) \in [0, 1)^k \text{ за всяко } n \geq 0. \quad (24)$$

Всеки елемент от редицата на Холтън представлява k -мерен вектор. В практиката при построяване на k -мерна редица на Холтън се използват първите k на брой прости числа: 2, 3, 5, 7, 11, ... Този избор дава възможност да се построи редица на Холтън с достатъчно голяма размерност и при това числата са взаимно прости.

Хамерсли (Hammersley) [47] е предложил да се прави избор за равноотдалечени точки при първата координата без да се използва просто число. Той дефинира и крайно множество на редицата на Холтън. Използвайки тази идея, Холтън е предложил следната формула за получаване на k -мерна редица чрез първите $k - 1$ на брой прости числа:

$$\mathbf{x}_n = \left(\frac{n}{N}, \Phi_{b_1}(n), \dots, \Phi_{b_{k-1}}(n) \right) \in [0, 1)^k, \quad n = 1, 2, \dots, N. \quad (25)$$

Ванг (Wang) [112] е направил друга модификация на първата координата на редицата на Холтън при $k \geq 2$:

$$\mathbf{x}_n = \left(\frac{2n-1}{2N}, \Phi_{b_1}(n), \dots, \Phi_{b_{k-1}}(n) \right) \in [0, 1)^k, \quad n = 1, 2, \dots, N. \quad (26)$$

Чрез формули (25), (26) се избягва първата координата да е 0, а за изчислителната сложност се получава порядъка $\mathcal{O}(kN \log_{b_{\min}}(N))$, където $b_{\min} = \min\{b_1, b_2, \dots, b_{k-1}\}$ при генериране на k -мерни редици на Холтън. Съществуват и други модификации на редицата на Холтън, като например предложената от Атанасов [8, 9]. За нея е доказано, че осигурява по-добра сходимост в сравнение с оригиналната редица по отношение на оценка на дискрепанса.

Друга фундаментална конструкция на редици с малък дискрепанс са редиците на Соболю (Sobol) [98], въведени през 1976. Редиците на Соболю [96, 97] притежават две основни разлики от редиците на Холтън: (1) при метода на радикалната инверсия се използва основа 2 и (2) този метод се комбинира с числа за посока (англ. *direction number*), получени от примитивни полиноми върху крайни полета. Редиците на Соболю се проектират така, че да осигуряват много добра равномерност върху единичния хиперкуб $[0, 1)^k$. Те принадлежат към класа цифрови (t, k) -редици с основа 2 (т.е. конструират се по битове в двоична система).

Дефиниция 0.0.11. (редица на Соболю (цифрова редица при основа 2))

Нека $n = 0, 1, 2, \dots$, и двоичното (битово) разлагане на n е:

$$n = \sum_{r=0}^{\infty} n_r 2^r, \quad n_r \in \{0, 1\}. \quad (27)$$

Всяка координата във вектора $\mathbf{n} = (n_0, n_1, n_2, \dots)^T$ се получава чрез побитово събиране над полето $\mathbb{F}_2 = \{0, 1\}$ (събиране по модул 2) по следния начин: За всяка координата

$j = 1, \dots, k$ на k -мерния вектор \mathbf{n} се задава генерираща двоична матрица:

$$C^{(j)} = \left(c_{i,r}^{(j)} \right)_{i \geq 1, r \geq 0}, \quad c_{i,r}^{(j)} \in \{0, 1\}. \quad (28)$$

Битовете на двоичната дроб на $x_n^{(j)}$ се пресмятат с формулата:

$$y_{n,i}^{(j)} = \sum_{r=0}^{\infty} c_{i,r}^{(j)} \pmod{2}, \quad i = 1, 2, \dots, \quad (29)$$

където събирането е по модул 2, а самата координата се получава от равенството:

$$x_n^{(j)} = \sum_{i=1}^{\infty} y_{n,i}^{(j)} 2^{-i} \in \{0, 1\}. \quad (30)$$

Тогава n -тата точка от редицата на Соболю като k -мерен вектор е:

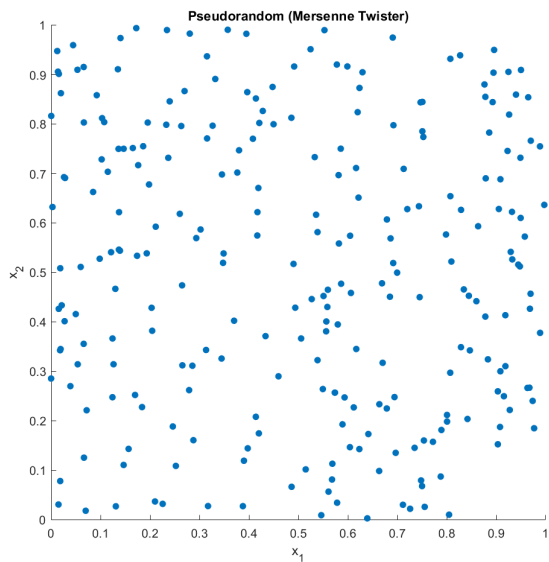
$$\mathbf{x}_n = (x_n^{(1)}, x_n^{(2)}, \dots, x_n^{(k)}) \in [0, 1]^k. \quad (31)$$

Ключов момент при описаната процедура за конструиране на цифрова редица с основа 2 е при избора на елементите на матрицата $C^{(j)}$, необходима за конструиране на j -тата координата на точка от редицата. При редиците на Соболю те се конструират по класическата схема чрез примитивен полином за всяка координата над полето \mathbb{F}_2 (т.е. полином, чиито коефициенти имат за общ делител само 1). Числото $y_{n,i}^{(j)}$ в (29) се нарича i -тото направляващо число на j -тата координата [96, 116]. Антонов и Салиев в [6] са използвали двоичен код на Грей за пресмятане на началните направляващи числа от примитивните полиноми. Те са представили схема на алгоритъм за генериране на редици на Соболю с N елемента, всеки от които е с размерност k . Изчислителната сложност на алгоритъма е от порядъка $\mathcal{O}(N(k + \log N))$.

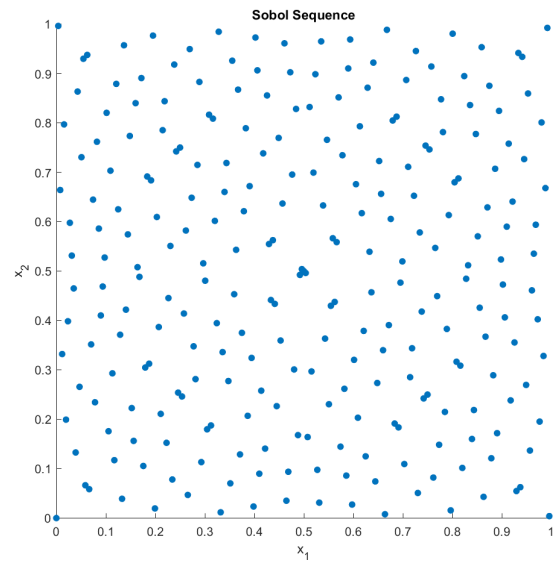
С развитието на теорията за (t, m, k) -мрежи и (t, k) -редици, тясно свързана с понятието *елементарен интервал* при основа b , се създават предпоставки за конструиране на широк клас квазислучайни редици [84, 85]. В този контекст редиците на Соболю представляват (t, k) -редици при основа 2. Към този клас принадлежат конструираните редици от Фор (Faure) през 1982 г. [35], който съществено обобщава метода на инверсията. Характерна особеност на редиците на Фор е, че тяхната конструкция се базира на $(0, m, k)$ -мрежа и $(0, k)$ -редица при основа просто число, по-голямо или равно на k . По-късно Нидерайтер (Niederreiter) обобщава редиците на Соболю и конструира редици с произволни основи, известни като редици на Нидерайтер [84]. Тезука (Tezuka) [105, 106, 107] доразвива подхода на Нидерайтер чрез използване на полиномиален аналог, близък по дух до конструкциите от типа на Холтгън.

За практически цели, редиците на Соболю, Холтгън и Фор са сред най-често използваните при решаване на задачи с помощта на КМК методи. В дисертационния труд разработените стохастични алгоритми за оценяване на екстремалните собствени числа на симетрични матрици са реализирани с редиците на Соболю и Холтгън. Те са сред най-широко реализираните квазиредици в софтуерни продукти за генериране на квазислучайни числа.

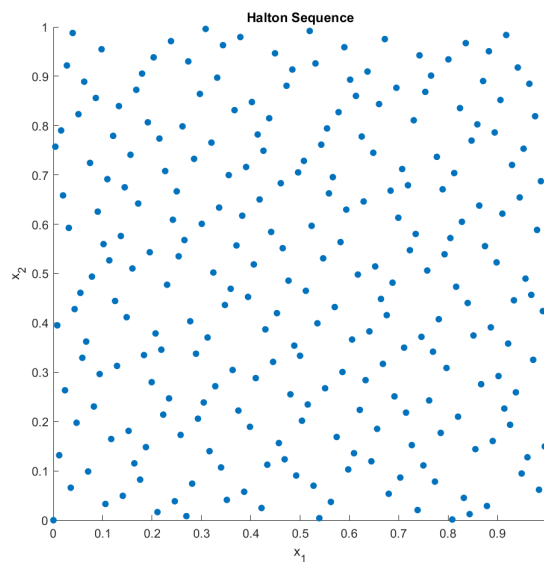
На Фигура (1) е представено сравнение между псевдослучайна редица и редици на Собол и Холтън. Случай (а) илюстрира разположението на 256 двумерни точки, генерирани с Mersenne Twister (MT) PRNG. В останалите два случая са използвани двумерните проекции на 16-мерни редици на Собол (случай (б)) и Холтън (случай (в)), като са разгледани първите две координати. Получените резултати ясно демонстрират по-равномерното разпределение на редиците с малък дискрепанс в единичния квадрат в сравнение с двумерната псевдослучайна редица.



(а) MT генератор на псевдослучайни числа



(б) редица на Собол



(в) редица на Холтън

Фигура 1: Сравнение на двумерна псевдослучайна редица, получена с MT PRNG, с двумерни проекции на 16-мерни редици на Собол и Холтън. Генерираните точки са 256, а проекцията е направена по първите две координати.

Въпреки че редиците с малък дискрепанс имат съществени предимства спрямо псевдослучайните, техният детерминистичен характер води до два практически значими недоста-

тъка: (1) липса на естествена вероятностна рамка за оценяване на грешката и построяване на доверителни интервали и (2) повишен риск от възникване на корелации при конструиране и използване на редиците с малък дискрепанс с висока размерност (виж Фигура (2) и Фигура (3)). Подобни корелации могат да се появят при неблагоприятно „съгласуване“ между геометрията на точковия набор при структурата на редиците с малък дискрепанс за големи размерности.

Обобщено, редиците с малък дискрепанс (класическите детерминистични редици на Соболев и Холтън) се отличават с висока равномерност при малки и средни размерности. При по-високи размерности обаче е възможна поява на корелации между координатите, което налага използването на допълнителни техники, като *разбъркване* (англ. *scrambling*) или *пропускане на начални членове* (операции от тип *skip* и *leap*) на редицата [62, 63, 88, 89].

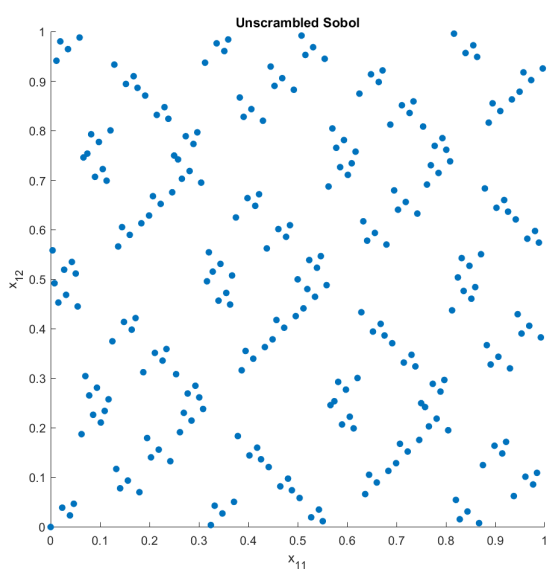
Рандомизирани редици с малък дискрепанс

Рандомизираните редици с малък дискрепанс преодоляват ограниченията на детерминистичните квазислучайни конструкции чрез въвеждане на контролирана стохастичност. Тя „разчупва“ фиксираната геометрия на точковия набор, но запазва основното предимство, а именно равномерното покритие на единичния хиперкуб $[0, 1]^k$ [18]. На практика рандомизацията се реализира чрез прилагане на специално конструирани трансформации (разбърквания) върху вече изградени редици с малък дискрепанс. Тези трансформации не разрушават равномерността на редиците, а я запазват в статистически смисъл [86].

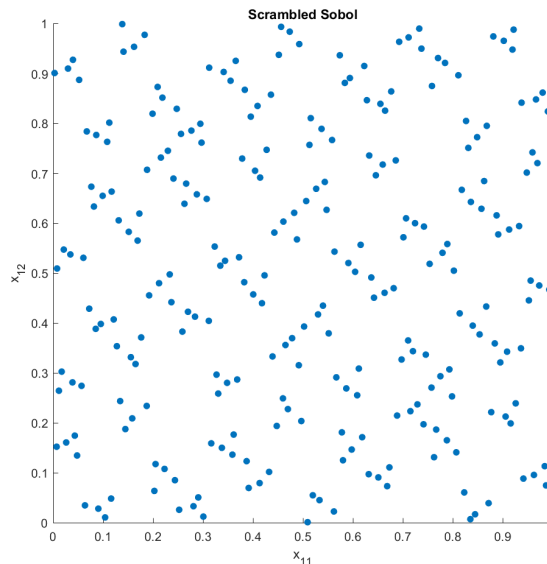
Съществуват два основни подхода за разбъркване (англ. *scrambling*). Първият се основава на *побитово (цифрово) случайно пермутиране*, при което цифрите на координатите в основа b (в частност при основа 2 при двоично представяне) се подлагат на случайни пермутации [86]. Този подход включва различни варианти на цифрови пермутации и побитово разбъркване, които запазват малкия дискрепанс, но осигуряват статистическа независимост между различни реализации на редицата. Вторият подход се основава на *събиране по модул 1* със случайно число или случаен вектор, генерирани чрез псевдослучаен генератор. Този метод е известен още като *случайно изместване* (англ. *shift*) по модул 1 и се характеризира с проста и ефективна реализация. При него всички точки от редицата се трансформират чрез добавяне на една и съща случайна стойност, след което резултатът се редуцира по модул 1, така че да остане в границите на единичния хиперкуб [53, 54].

На практика двата подхода често се комбинират, което води до множество техни варианти и комбинации. Например, могат да се използват различни стратегии за линеаризация на многомерните индекси, различни основи на цифровото представяне, както и различни видове псевдослучайни или квазислучайни генератори при конструиране на изместванията. Тези комбинации позволяват да се постигне по-добър баланс между равномерност и стохастична вариабилност [89]. Изборът на конкретна реализация (тип пермутации с основа b , използване на генератор за изместване или степен на разбъркването) зависи от свойствата на интегранда и размерността на задачата. По този начин рандомизираните редици с малък дискрепанс съчетават предимствата на детерминираното равномерно

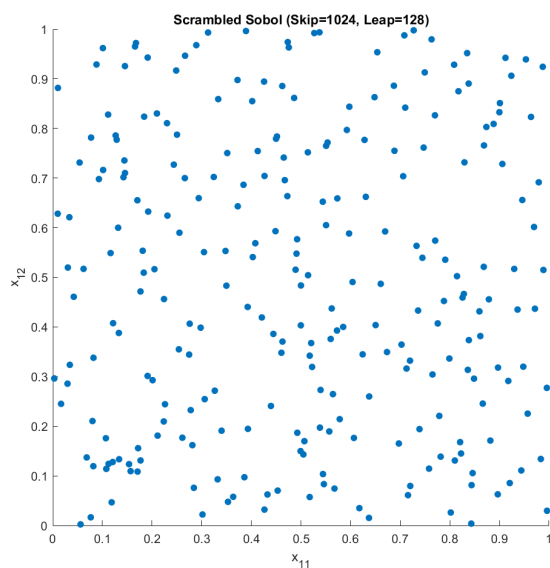
покрытие с гъвкавостта на вероятностните методи. Те осигуряват по-бърза сходимост в сравнение със стандартните Монте Карло техники и надеждна оценка на грешката чрез многократни независими рандомизации. Това ги прави особено подходящи за оценяване на многомерни интеграли и други числени задачи с висока изчислителна сложност [87].



(а) неразбъркана редица на Соболев



(б) разбъркана редица на Соболев

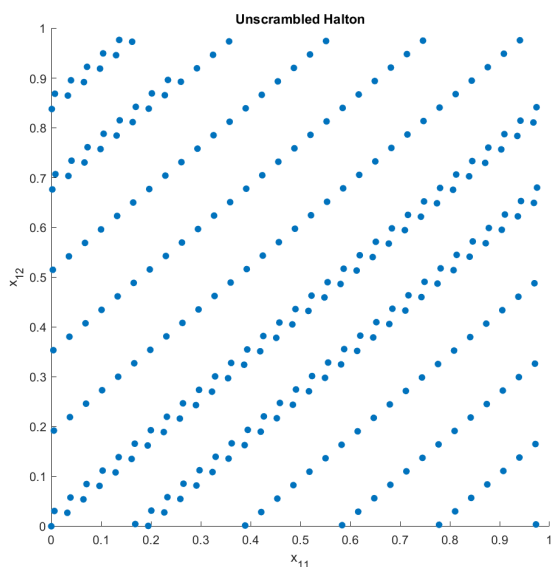


(в) разбъркана редица на Соболев с параметри *skip*, *leap*

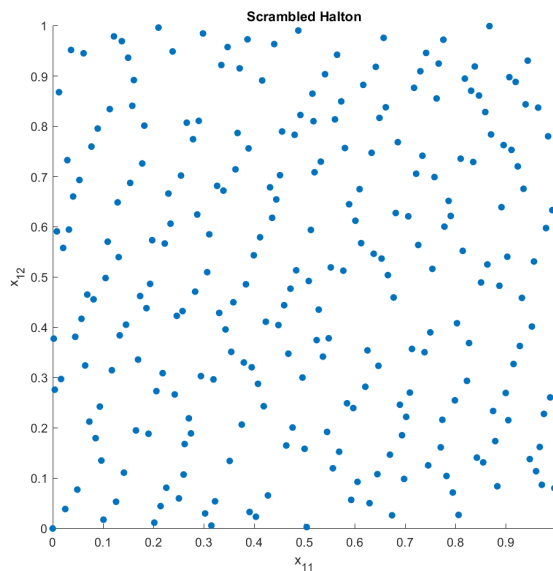
Фигура 2: Сравнение между двумерна проекция с наличие на корелация при 11-та и 12-та координата на редица на Соболев и съответните двумерни проекции на рандомизирани редици на Соболев без и с допълнително включени параметри $skip = 1024$ и $leap = 128$.

В нашите изследвания са използвани разбъркани редици на Соболев и Холтън, за да избегнем корелации, които могат да възникнат при по-високи измерения. В MATLAB, редиците на Холтън се генерират с помощта на типа разбъркване „RR2“, който прилага операция с обратна основа, за да пермутира радикално-инверсните коефициенти [58, 95],

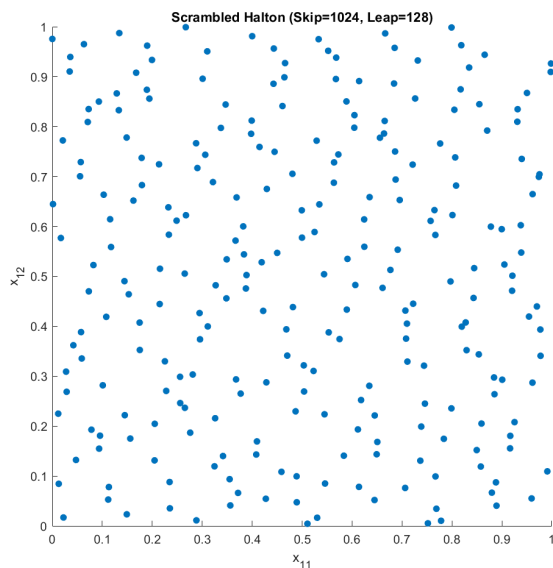
като по този начин запазва еднородността. Редиците на Соболев се генерират с помощта на разбъркване на „MatousekAffineOwen“, което комбинира произволно линейно разбъркване [77] с произволно цифрово отместване [88]. Двата метода на разбъркване позволяват да се добавя използването на операциите *skip* и *leap*. Така въведените рандомизации намаляват корелациите, подобряват оценката на дисперсията, а съответните редици се „държат“ по-скоро като истински случайни извадки.



(а) неразбъркана редица на Холтгън



(б) разбъркана редица на Холтгън



(в) разбъркана редица на Холтгън с параметри *skip*, *leap*

Фигура 3: Сравнение между двумерна проекция с наличие на корелация при 11-та и 12-та координата на редица на Холтгън и съответните двумерни проекции на рандомизирани редици на Холтгън без и с допълнително включени параметри $skip = 1024$ и $leap = 128$.

На Фигура (2) и Фигура (3) са представени сравнения между двумерни проекции на 16-

мерните класически детерминистични редици на Соболю и Холтъну и двумерните проекции на рандомизираните им варианти без и с добавяне на операции от тип *skip* и *leap*. Случаите (а) при Фигура (2) и Фигура (3) илюстрират наличие на корелации при двумерните проекции на детерминистични редици на Соболю и Холтъну по 11-та и 12-та координата. В останалите два случая е въведен стохастичен елемент чрез разбъркване (англ. *scrambling*) на редиците на Соболю и Холтъну: в случай (б) (виж Фигура (2) и Фигура (3)) е приложено стандартно разбъркване, а в случай (в) (виж отново Фигура (2) и Фигура (3)) е приложено разбъркване на редиците на Соболю и Холтъну с допълнително използване на параметри $skip = 1024$ и $leap = 128$. Резултатите за 256 точки в единичния квадрат. На Фигура (2) и Фигура (3) ясно показват, че разбъркването на редиците на Соболю по метода на „MatousekAffineOwen“ и при разбъркване на редиците на Холтъну с „RR2“ метода водят до значително намаляване на корелацията при редици с висока размерност. Двата метода на разбъркване заедно с включване на параметрите *skip* и *leap* са имплементирани в MATLAB.

Прескачането (англ. *skip/skipping*) има за цел най-вече да избегне начални неравномерности и „стартови“ (англ. *warm-up*) артефакти, характерни за много конструкции на редици с малък дискрепанс. Първите генерирани точки в цифровите (t, k) -редици могат да имат по-големи пропуски (англ. *gaps*), т.е. по-голяма дискретност поради не идеална инициализация на рекурсивни построения или да притежават по-ниска „разделителна способност“ на т.нар. числа по направление (англ. *direction numbers*). Ето защо отстраняването на тези точки чрез оператор за прескачане (англ. *skipping operator*) може да намали локалната дискретност [62].

Дефиниция 0.0.12. (*оператор за прескачане (англ. skipping operator)*)

Нека $\{\mathbf{x}_i\}_{i=0}^{\infty}$ е оригиналната безкрайна редица. Операторът за прескачане (*skipping operator*) \mathcal{S}_s с параметър $s \in \mathbb{N}$, който отстранява първите s точки в редицата, се въвежда с равенството:

$$\mathcal{S}_s(\{\mathbf{x}_i\}) = \{\mathbf{x}_{i+s}\}_{i=0}^{\infty}.$$

Дефиниция 0.0.13. (*оператор за прескачане през стъпка (англ. leaping operator)*)

Нека $\{\mathbf{x}_i\}_{i=0}^{\infty}$ е оригиналната безкрайна редица. Операторът за прескачане през стъпка (*leaping operator*) \mathcal{L}_l с параметър $l \in \mathbb{N}$, който избира всяка l -та точка от оригиналната редица, се въвежда с равенството:

$$\mathcal{L}_l(\{\mathbf{x}_i\}) = \{\mathbf{x}_{il}\}_{i=0}^{\infty} = (\mathbf{x}_0, \mathbf{x}_l, \mathbf{x}_{2l}, \mathbf{x}_{3l}, \dots).$$

Операторът за прескачане през фиксирана стъпка (англ. *leaping operator*) променя разстоянието между последователните точки и следователно оказва по-съществено влияние върху структурата на редицата. Този оператор може да намали локалната корелация между точките, поради което е полезен при паралелни изчисления. Параметърът l трябва да се подбира внимателно, за да не наруши общата равномерност на разпределението на точките в единичния хиперкуб.

Теоретичните резултати показват [18], че ако оригиналната редица удовлетворява нера-

венството:

$$D_N^* \leq C_k \frac{(\log(N))^k}{N},$$

то за всеки фиксиран параметър s новополучената редица след прилагане на оператора за прескачане удовлетворява:

$$D_N^*(\mathcal{S}_s(\{\mathbf{x}_i\})) \leq D_{N+s}^*(\{\mathbf{x}_i\}) + \frac{s}{N}.$$

Следователно, когато $N \rightarrow \infty$, получаваме $\frac{s}{N} \rightarrow 0$, така че асимптотичният порядък на дискретността остава непроменен. Оказва се, че често резултатите се подобряват при подходящо избран параметър s и крайна извадка N [18].

На практика операторите \mathcal{S}_s и \mathcal{L}_l могат да се комбинират, в резултат на което се получава нова безкрайна редица:

$$\{\mathbf{y}_i\}_{i=0}^{\infty} = \mathcal{L}_l(\mathcal{S}_s\{\mathbf{x}_i\}) = \{\mathbf{x}_{s+il}\}_{i=0}^{\infty}.$$

В дисертационния труд двата оператора са приложени при рандомизацията на редиците на Собол и Холтън с цел получаване на ефективен Степенен КМК алгоритъм за оценяване на максималната собствена стойност на две корелационни матрици: (1) корелационна матрица, конструирана на базата на годишни данни за портфейл, състоящ се от 32 глобални актива и (2) корелационна матрица, образувана от данните за доходността на 490 компании от *S&P500*.

Стохастични методи с вериги на Марков

Монте Карло (МК) и Квази-Монте Карло (КМК) методите условно се разделят на два класа: *директни* и *итерационни*.

За *директните* методи е характерно, че полученото приближено решение зависи единствено от стохастичната (статистическата) грешка. При класическия МК подход тази грешка се описва чрез r_N , определена от неравенство (10), и е от порядък $\mathcal{O}(N^{-1/2})$. При КМК подхода грешката се оценява чрез неравенството на Коксма-Хлавка (21) и е от порядък $\mathcal{O}(N^{-1}(\log N)^k)$ при използване на редици с малък дискрепанс, като редиците на Собол и Холтън. В дисертационния труд се използват рандомизирани редици с малък дискрепанс, които въвеждат контролирана стохастичност, без да се губи предимството на по-равномерното покритие на единичния куб. Поради това грешките, оценени чрез неравенствата (10) и (21), се разглеждат като стохастична (статистическа) грешка, зависеща от размера на извадката N .

Итерационните МК/КМК методи, освен стохастична грешка, се характеризират и със систематична грешка. Тя възниква, тъй като стохастичният метод не оценява директно точното решение, а k -тата итерация на сходящ итерационен процес, който приближава решението. В теорията на класическите числените методи за линейна алгебра *методът на простата итерация* и *методът на Гаус-Зайдел* [117] са примери за сходящи итерационни процеси за намиране на приближено решение на система линейни алгебрични уравнения (СЛАУ). Тези методи се характеризират със систематична грешка, т.е. разликата между

приближеното и точното решение, която зависи от свойствата на итерационната схема, включително от нормата на матрицата, участваща в построяването на итерационния процес.

В теорията на стохастичните итерационни методи за задачи на линейната алгебра се конструира случайна величина $\theta^{(k)}$ чрез верига на Марков с k стъпки (прехода) [21, 27, 119]. Броят стъпки може да бъде фиксиран или да се задава чрез случайна величина с математическо очакване, равно на k . Математическото очакване на случайната величина $\theta^{(k)}$ оценява търсеното решение като функционал, включващ матрично-векторни произведения. В този контекст систематичната грешка зависи от броя на преходите (итерации) k във веригата на Марков, а стохастичната грешка е резултат от размера на извадката N . Постигането на баланс между систематичната и стохастичната грешка, както и намаляването на дисперсията, са ключови фактори за изследване при разработването на ефективни стохастични итерационни методи и алгоритми за задачи от линейната алгебра.

Структура на дисертацията

Дисертационният труд се състои от увод, три глави, заключение, приложение и библиография. Дисертацията съдържа 118 страници, включва 15 фигури и 16 таблици, библиография от 121 заглавия, както и списък с математически означения и използвани абривиатури.

- В **Увода** е представена актуалността на тематиката, направен е обзор на основните научни резултати в областта, поставени са целите и задачите на дисертационния труд и подробно е описана методологията на изследването.
- **Глава 1** „Стохастичен Степенен метод: разработване и изследване на ефективни алгоритми за оценяване на максимална собствена стойност“ е посветена на разработването на ефективни стохастични степенни алгоритми за намиране на максималната собствена стойност на плътни симетрични матрици. Дефинирана е задачата и са разгледани методите, които се използват за нейното решаване, а именно, детерминистичния Степенен метод и стохастичния Степенен метод, в неговите два варианта - Степенен Монте Карло (СМК) и Степенен Квази-Монте Карло (СКМК) метод. Подробно е описана конструкцията на случайната величина с използване на верига на Марков, като се акцентира на използването на почти оптимални вероятности. Представен е псевдокод на разработените почти оптимални СМК и СКМК алгоритми и са определени условията за балансиране на стохастичната и систематичната грешка. Главата завършва с дискусия върху направените числени експерименти за матрици с различна размерност.
- **Глава 2** „Стохастичен Резолвентен метод: разработване и изследване на ефективни алгоритми за оценяване на минимална собствена стойност“ разглежда задачата за оценяване на минималната собствена стойност на симетрични квадратни матрици, чрез използване на нейната резолвентна матрица. Представен е Резолвентният Степенен метод и неговите стохастични реализации чрез Резолвентен Монте Карло

(РМК) и Резолвентен Квази-Монте Карло (РКМК) подход. В главата се изследва ролята на параметрите, влияещи върху сходимостта на безкрайния ред, чрез който се представя резолвентната матрица. Конструирани са почти оптимални РМК и РКМК алгоритми и се изследва балансът на стохастичната и систематична грешка. Този баланс зависи от броя реализации на случайните величини, дължината на веригата на Марков, степента на резолвентната матрица и ускоряващия параметър, включен в представянето на резолвентната матрица като степенен ред. В последната секция са представени проведените числени експерименти, получени при оценяването на минималната собствена стойност на тестови симетрични матрици.

- **Глава 3** „Стохастични методи за оценяване на пазарния риск на инвестиционни портфейли“ е свързана с прилагането на разработените алгоритми от Глава 1 към реална задача от областта на финансовата математика, свързана с оценяване на пазарния риск на инвестиционен портфейл. Оценяването му се извършва чрез максималната собствена стойност на корелационната матрица на портфейла. Намирането на най-голямата собствена стойност се извършва чрез почти оптималните СМК и СКМК алгоритми. При прилагането на почти оптималния СКМК алгоритъм са разгледани два варианта: (1) с разбъркани редици на Соболев и Холтърн с настройка по „подразбиране“; (2) с разбъркани редици на Соболев и Холтърн с използване на параметри *skip* и *leap*. Проведени са числени експерименти с реални финансови данни, които демонстрират приложимостта и ефективността на предложените алгоритми.

В заключителната част са включени: списък на публикациите, апробация, научните и научно-приложните приноси и благодарности.

Глава 1

Стохастичен Степенен метод: разработване и изследване на ефективни алгоритми за оценяване на максимална собствена стойност

1.1 Постановка на задачата

Нека $A = \{a_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}$ е дадена несингулярна симетрична матрица. *Собствените стойности (или спектър)* на симетричната матрица A са всички реални стойности на параметъра $\lambda(A)$, за които уравнението:

$$A\mathbf{x} = \lambda(A)\mathbf{x}, \text{ където } \mathbf{x} \in \mathbb{R}^n, \quad (1.1)$$

има ненулево решение. Ненулевите вектори, които са решение на уравнение (1.1) се наричат *собствени вектори* на матрицата A . За да има уравнение (1.1) ненулево решение, необходимо и достатъчно условие е стойностите на параметъра $\lambda(A)$ да са решения (корени) на уравнението $|A - \lambda(A)\mathbf{I}| = 0$, където \mathbf{I} е единичната n -мерна матрица [117]. Предполагаме, че собствените стойности на A удовлетворяват условието:

$$\lambda_{min} = |\lambda_n| \leq |\lambda_{n-1}| \leq |\lambda_{n-2}| \leq \dots \leq |\lambda_2| < |\lambda_1| = \lambda_{max}. \quad (1.2)$$

Разглеждаме задачата за разработване и изследване на ефективни алгоритми на базата на Стохастичния Степенен метод за оценяване на λ_{max} , която съгласно (1.2) е единствена.

1.2 Методи за изчисляване на максималната собствена стойност

1.2.1 Степенен метод

Степенният метод е итеративен числен метод за приближено намиране на максималната собствена стойност λ_{\max} на квадратна матрица [39, 51]. Итеративната процедура се реализира чрез следните три последователни стъпки, базирани на коефициента на Рейли [51]:

$$r(\mathbf{x}) = \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}},$$

както следва:

1. **Инициализация:** Избира се начален вектор \mathbf{x}_0 , такъв че $\|\mathbf{x}_0\| = 1$.
2. **Итерационен процес:** За $k = 0, 1, \dots$ се **изчислява:** $\mathbf{x}_{k+1} = r(\mathbf{x}_k)$.
3. **Оценяване на собствена стойност:** Съответната максимална собствена стойност за задачата (1.1) при (1.2) се оценява след k итерации като:

$$\lambda_{\max} \approx \lambda_{\max}^{(k)} = r(\mathbf{x}_k).$$

Систематична грешка: В случая на симетрична матрица грешката от прекъсване на Степенния метод при представяне на векторите \mathbf{x} и \mathbf{x}^T като линейни комбинации на ортонормиран базис е от порядъка [51]:

$$|\lambda_{\max}^{(k)} - \lambda_{\max}| = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right). \quad (1.3)$$

Сходимость: Скоростта на сходимость на итерационния процес при симетрични матрици се определя от коефициента на сходимость $\left|\frac{\lambda_2}{\lambda_1}\right|^2$ [51]. Затруднения в сходимостта възникват, когато първите две собствени стойности по модул са „близки“ една до друга. В този случай се използват техники за предварителна обработка (англ. *preconditioning*), целящи по-голямо разграничаване на двете доминиращи собствени стойности.

Изчислителна сложност: Броят на необходимите аритметични операции при прилагане на Степенния метод е $\mathcal{O}(kn^2)$, където k е броят на извършените итерации, а n е размерността на матрицата A .

1.2.2 Стохастичен Степенен метод

Стохастичният Степенен метод, който ще изследваме в дисертационния труд, е разработен от Караиванова и Димов в [26, 27].

Нека отново разгледаме задачата за собствени стойности (1.1) при условие (1.2). Най-голямата собствена стойност може да се представи като граница на коефициента на Рейли [27, 51]:

$$\lambda_{max} = \lim_{k \rightarrow \infty} \frac{(\mathbf{h}, A^k \mathbf{f})}{(\mathbf{h}, A^{k-1} \mathbf{f})}, \quad (1.4)$$

където $(\mathbf{h}, A^k \mathbf{f})$ е скалярно произведение, а векторите $\mathbf{f} = \{f_i\}_{i=1}^n$ и $\mathbf{h} = \{h_i\}_{i=1}^n$ са избрани произволно и нормирани. Ако k е произволно голямо естествено число, тогава:

$$\lambda_{max}^{(k)} = \frac{(\mathbf{h}, A^k \mathbf{f})}{(\mathbf{h}, A^{k-1} \mathbf{f})} \quad (1.5)$$

е приближение на λ_{max} с определена систематична грешка.

Стохастичният Степенен метод се основава на конструирането на случайни величини, чиито математически очаквания са равни на скалярните произведения в числителя и в знаменателя на израза (1.5). Нека разгледаме случайната величина $\theta^{(k)}$, предназначена за оценяване на скалярното произведение $(\mathbf{h}, A^k \mathbf{f})$ [21, 22], чрез дискретна верига на Марков [56, 60]. Конструирането на тази случайната величина може да бъде описано чрез следните стъпки:

1. **Конструиране** на дискретна крайна верига на Марков със състояния l_t :

$$l_0 \rightarrow \dots \rightarrow l_t \rightarrow \dots \rightarrow l_k \quad (1 \leq l_t \leq n, 0 \leq t \leq k), \quad (1.6)$$

с начален вектор на вероятностната плътност $\mathbf{p} = \{p_i\}_{i=1}^n \in \mathbb{R}^n$,

$\mathbb{P}\{l_0 = i\} = p_i \geq 0, \sum_{i=1}^n p_i = 1$ и матрица на преходната плътност

$P = \{p_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}, \sum_j p_{ij} = 1, \mathbb{P}\{l_t = j | l_{t-1} = i\} = p_{ij} \geq 0, (i, j = \overline{1, n}; t = \overline{1, k})$.

2. **Изчисляване** на теглото W_k итеративно:

$$W_0 = 1, W_t = W_{t-1} \frac{a_{l_{t-1} l_t}}{p_{l_{t-1} l_t}}, t = 1, \dots, k. \quad (1.7)$$

3. **Определяне** на случайната величина $\theta^{(k)}$ чрез формулата:

$$\theta^{(k)} = \frac{h_{l_0}}{p_{l_0}} W_k f_{l_k}. \quad (1.8)$$

В сила е следната теорема:

Теорема 1.2.1. *Нека случайната величина $\theta^{(k)}$ е дефинирана чрез формула (1.8). Тогава математическото очакване на случайната величина $\theta^{(k)}$ е равно на скалярното произведение $(\mathbf{h}, A^k \mathbf{f})$, т.е.:*

$$E[\theta^{(k)}] = (\mathbf{h}, A^k \mathbf{f}). \quad (1.9)$$

Доказателството на Теорема 1.2.1 може да бъде намерено в [22].

При конструирането на теглото W_k посредством веригата на Марков (1.6) се въвеждат множества от *допустими* плътности \mathfrak{P}_h и \mathfrak{P}_A , зависещи съответно от избора на вектора \mathbf{h} и матрицата A (виж [22]).

Дефиниция 1.2.1. Началният вектор на вероятностната плътност $\mathbf{p} = \{p_i\}_{i=1}^n \in \mathbb{R}^n$ се нарича допустим за вектор $\mathbf{h} = \{h_i\}_{i=1}^n \in \mathbb{R}^n$, т.е. $\mathbf{p} \in \mathfrak{P}_{\mathbf{h}}$, ако $p_i > 0$, когато $h_i \neq 0$ и $p_i = 0$, когато $h_i = 0$ за $i = 1, \dots, n$.

Дефиниция 1.2.2. Матрицата на преходната плътност $P = \{p_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}$ се нарича допустима за симетричната матрица $A = \{a_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}$, т.е. $P \in \mathfrak{P}_A$, ако $p_{ij} > 0$, когато $a_{ij} \neq 0$ и $p_{ij} = 0$, когато $a_{ij} = 0$ за $i, j = 1, \dots, n$.

Нека са дадени векторът на началната плътност \mathbf{p} и матрицата на преходната плътност P и са допустими съгласно дефиниции (1.2.1) и (1.2.2). Да разгледаме извадка от N на брой стойности на случайната величина $\theta^{(k)}$. Средната стойност на случайната величина $\theta^{(k)}$ се дефинира по следния начин:

$$\bar{\theta}_N^{(k)} = \frac{1}{N} \sum_{s=1}^N (\theta^{(k)})_s, \quad (1.10)$$

където s -тата реализация на случайната величина $\theta^{(k)}$ е означена с $(\theta^{(k)})_s$ за $1 \leq s \leq N$.

При използване на генератори на псевдослучайни числа (PRNGs) [56, 60], средната стойност (1.10) се разглежда като МК апроксимация на скаларното произведение $(\mathbf{h}, A^k \mathbf{f})$ с вероятностна грешка:

$$R_N^{(k)} = |(\mathbf{h}, A^k \mathbf{f}) - \bar{\theta}_N^{(k)}| < c_1 \sigma(\theta^{(k)}) N^{-1/2}, \quad (1.11)$$

където c_1 е константа, а $\sigma(\theta^{(k)})$ е стандартно отклонение на случайната величина [22].

Когато за генериране на N реализации на случайната величина $\theta^{(k)}$ чрез веригата на Марков (1.6) се използват редици с малък дискрепанс, като тези на Собол и Холтън, средната стойност (1.10) представлява КМК апроксимация на скаларното произведение $(\mathbf{h}, A^k \mathbf{f})$ [43], а грешката се определя по следния начин:

$$QR_N^{(k)} = |(\mathbf{h}, A^k \mathbf{f}) - \bar{\theta}_N^{(k)}| < c_2 \frac{(\log N)^{k+1}}{N}. \quad (1.12)$$

Тук $(k+1)$ е размерността на редицата с малък дискрепанс, като за получаване на оценката (1.12) е приложено неравенството на Коксма-Хлавка [60, 92].

В сила е следната теорема:

Теорема 1.2.2. Нека се разгледат случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$ при фиксираните стъпки k и $k-1$ на веригата на Марков (1.6). Тогава

$$\frac{E[\theta^{(k)}]}{E[\theta^{(k-1)}]} = \lambda_{max}^{(k)}. \quad (1.13)$$

Доказателство. Като се разгледа конструкцията на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$ при фиксираните стъпки k и $k-1$ във веригата на Марков, се получава:

$$\theta^{(k)} = \frac{h_{l_0}}{p_{l_0}} W_k f_{l_k} \quad \text{и} \quad \theta^{(k-1)} = \frac{h_{l_0}}{p_{l_0}} W_{k-1} f_{l_{k-1}}.$$

Като се вземе предвид равенството (1.5) и Теорема 1.2.1, се получава:

$$\frac{E[\theta^{(k)}]}{E[\theta^{(k-1)}]} = \frac{(\mathbf{h}, A^k \mathbf{f})}{(\mathbf{h}, A^{k-1} \mathbf{f})} = \lambda_{max}^{(k)}.$$

□

Стохастичният Степенен метод за оценяване на $\lambda_{max}^{(k)}$ се основава на:

$$\lambda_{max}^{(k)} \approx \frac{\sum_{s=1}^N (\theta^{(k)})_s}{\sum_{s=1}^N (\theta^{(k-1)})_s}. \quad (1.14)$$

Тук символът " \approx " показва, че изразът от дясната страна в (1.14) е приблизително решение за $\lambda_{max}^{(k)}$, което зависи асимптотично от k и N . Значението на този символ е едно и също във всички следствия по-долу.

Забележка 1. *Стохастичният Степенен метод се нарича Степенен Монте Карло (СМК) метод, когато извадките за двете случайни величини $\theta^{(k)}$ и $\theta^{(k-1)}$ се генерират чрез използването на генератори на случайни числа за конструиране на веригата на Марков (1.6).*

По този начин стохастичната грешка на СМК метода е функция на вероятностната грешка, дефинирана от (1.11).

Забележка 2. *Когато за конструирането на веригата на Марков при генериране на извадките на $\theta^{(k)}$ и $\theta^{(k-1)}$ се използват разбъркани редици с малък дискрепанс, като например разбърканите редици на Соболев [97], Стохастичният Степенен метод се нарича Степенен Квази-Монте Карло (СКМК) метод.*

Стохастичната грешка на СКМК метода се явява функция на грешката, дефинирана чрез (1.12).

Преформулирането на задачата в термините на интегралните уравнения с подходящи избрани ядра и области позволява използването на неравенството на Коксма-Хлавка при анализиране на сходимостта на разглежданите алгоритми. В общия случай, използването на редици с малък дискрепанс (т. нар. квазиредици) подобрява скоростта на сходимост на съответните КМК алгоритми, но не дава възможност за апостериорна оценка на грешката. Разбъркването (рандомизирането) на квазиредицата позволява избягването на този недостатък.

Теоретичното сравнение на грешката при СМК и СКМК метод показва, че и в двата случая грешката е произведение от два множителя. Единият, зависещ от редицата, а другият от функцията или случайната величина. При това, неравенството на Коксма-Хлавка е граница в най-лошия случай, докато грешката при СМК подход има вероятностен характер. Нещо повече, изчислителният опит показва, че вариацията $V(f)$ в неравенството на Коксма-Хлавка обикновено е свърхоценка, докато вторият множител дискрепансът на редицата показва действителния порядък на грешката.

1.3 Почти оптимални Степенен Монте Карло и Степенен Квази-Монте Карло методи и алгоритми

1.3.1 Прецизиране на конструкцията на $\theta^{(k)}$ и $\theta^{(k-1)}$ с почти оптимални вероятности за веригата на Марков: балансиране на грешките

Нека представим симетричната матрица A по следния начин:

$$A = \{a_{ij}\}_{i,j=1}^n = (\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_n)^T, \text{ където } \mathbf{a}_i = (a_{i1}, \dots, a_{in}), \quad i = 1, \dots, n.$$

В по-нататъшното изложение се разглеждат следните векторни и матрични норми:

$$\|\mathbf{h}\| = \|\mathbf{h}\|_1 = \sum_{i=1}^n |h_i|, \quad \|\mathbf{a}_i\| = \|\mathbf{a}_i\|_1 = \sum_{j=1}^n |a_{ij}|, \quad \|A\| = \|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|, \quad j = 1, \dots, n.$$

Съгласно неравенството (1.11), вероятностната грешка зависи от стандартното отклонение σ , съответно дисперсията σ^2 , както и от размера на извадката N . При фиксиран брой реализации N на веригата на Марков, вероятностната грешка се определя от дисперсията на случайната величина $\theta^{(k)}$, т.е. при по-малка дисперсия имаме по-малка вероятностна грешка. От своя страна, дисперсията зависи от избора на вектора на плътност $\mathbf{p} \in \mathfrak{P}_h$ и матрицата на преходната плътност $P \in \mathfrak{P}_A$. Всички допустими вектори на плътност $\mathbf{p} \in \mathfrak{P}_h$ и матрици на преходната плътност $P \in \mathfrak{P}_A$ могат да бъдат използвани за конструиране на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$. Тези случайни величини служат за оценяване на $\lambda_{max}^{(k)}$, съгласно (1.10). Достатъчно условие е дисперсиите им да са крайни.

Най-често използваният подход за конструиране на верига на Марков за задачи на линейната алгебра се основава на равномерното разпределение. В този случай координатите на вектора на плътност \mathbf{p} и елементите на матрицата на плътност на прехода P се задават равни на $\frac{1}{n}$, т.е.:

$$\mathbf{p} = \{p_i\}_{i=1}^n = \frac{1}{n} \text{ и } P = \{p_{ij}\}_{i,j=1}^n = \frac{1}{n}, \quad i, j = 1, \dots, n. \quad (1.15)$$

Когато изборът (1.15) се използва при конструирането на верига на Марков, Степенният МК и Степенният КМК методи се наричат *класически*. Алтернативни избори на \mathbf{p} и P водят до конструиране на нови оценки за $\lambda_{max}^{(k)}$. Получените решения обикновено се сравняват с тези от *класическия* подход. Казваме, че сме конструирали ефективен СМК или СКМК метод/алгоритъм, ако при това сравнение сме намалили/редуцирали дисперсията (изчислителната сложност) и сме подобрили точността.

Сред всички допустими вектори на плътност $\mathbf{p} \in \mathfrak{P}_h$ и матрици на плътност на прехода $P \in \mathfrak{P}_A$ съществува избор, който води до значително намаляване на дисперсията [21, 22]:

$$\mathbf{p} = \{p_i\}_{i=1}^n, \quad p_i = \frac{|h_i|}{\|\mathbf{h}\|} \text{ и } P = \{p_{ij}\}_{i,j=1}^n, \quad p_{ij} = \frac{|a_{ij}|}{\|\mathbf{a}_i\|}, \quad i, j = 1, \dots, n. \quad (1.16)$$

При избора чрез формули (1.16), \mathbf{p} се нарича *почти оптимален* начален вектор на плътност, а P се нарича *почти оптимална* матрица на плътност на прехода. Елементите на P се наричат *почти оптимални* вероятности. В този случай, конструкцията на случайната величина $\theta^{(k)}$ се различава съществено от тази във формулите (1.7) и (1.8) (виж по-долу конструкцията на *почти оптималния* Стохастичен Степенен метод).

В сила е следната теорема:

Теорема 1.3.1. *Нека \mathbf{p} и P са почти оптималният начален вектор на плътност и почти оптималната матрица на плътност. Тогава средната стойност (1.10) може да се изрази, както следва:*

$$\bar{\theta}_N^{(k)} = \frac{1}{N} \sum_{s=1}^N \text{sign}(h_{l_0^{(s)}}) \|\mathbf{h}\| \left\{ \prod_{t=1}^{t=k} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) \| \mathbf{a}_{l_{t-1}^{(s)}} \| \right\} f_{l_k^{(s)}}. \quad (1.17)$$

Доказателство. Като се използват формули (1.8) и (1.10) се получава:

$$\bar{\theta}_N^{(k)} = \frac{1}{N} \sum_{s=1}^N (\theta^{(k)})_s = \frac{1}{N} \sum_{s=1}^N \left\{ \frac{h_{l_0} W_k f_{l_k}}{p_{l_0}} \right\}_s = \frac{1}{N} \sum_{s=1}^N \frac{h_{l_0^{(s)}}}{p_{l_0^{(s)}}} \left\{ \prod_{t=1}^{t=k} \frac{a_{l_{t-1}^{(s)} l_t^{(s)}}}{p_{l_{t-1}^{(s)} l_t^{(s)}}} \right\} f_{l_k^{(s)}}.$$

Доказателството на теоремата следва след прилагане на формула (1.16), т.е.

$$\begin{aligned} \bar{\theta}_N^{(k)} &= \frac{1}{N} \sum_{s=1}^N \frac{h_{l_0^{(s)}}}{|h_{l_0^{(s)}}|} \|\mathbf{h}\| \left\{ \prod_{t=1}^{t=k} \frac{a_{l_{t-1}^{(s)} l_t^{(s)}}}{|a_{l_{t-1}^{(s)} l_t^{(s)}}|} \| \mathbf{a}_{l_{t-1}^{(s)}} \| \right\} f_{l_k^{(s)}} = \\ &= \frac{1}{N} \sum_{s=1}^N \text{sign}(h_{l_0^{(s)}}) \|\mathbf{h}\| \left\{ \prod_{t=1}^{t=k} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) \| \mathbf{a}_{l_{t-1}^{(s)}} \| \right\} f_{l_k^{(s)}}. \end{aligned}$$

□

От Теорема (1.2.2) и (1.3.1) се формулират следните следствия:

Следствие 1. *Нека \mathbf{p} и P са почти оптималният начален вектор на плътност и почти оптималната матрица на плътност. Тогава Стохастичният Степенен метод за оценяване на $\lambda_{\max}^{(k)}$ се дефинира като съотношение на средните стойности на случайните величини $\bar{\theta}_N^{(k)}$ и $\bar{\theta}_N^{(k-1)}$, използвайки формула (1.17), т.е.:*

$$\lambda_{\max}^{(k)} \approx \frac{\sum_{s=1}^N \text{sign}(h_{l_0^{(s)}}) \left\{ \prod_{t=1}^{t=k} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) \| \mathbf{a}_{l_{t-1}^{(s)}} \| \right\} f_{l_k^{(s)}}}{\sum_{s=1}^N \text{sign}(h_{l_0^{(s)}}) \left\{ \prod_{t=1}^{t=k-1} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) \| \mathbf{a}_{l_{t-1}^{(s)}} \| \right\} f_{l_{k-1}^{(s)}}}. \quad (1.18)$$

Следствие 2. *Да предположим, че L_1 -нормата на всички вектори-редове на симетричната матрица A е равна на константата a , т.е.:*

$$\|\mathbf{a}_i\|_1 = \|\mathbf{a}_i\| = \sum_{j=1}^n |a_{ij}| = a, \quad i = 1, \dots, n. \quad (1.19)$$

Тогава оценката за $\lambda_{max}^{(k)}$ по формула (1.18) може да се пренапише по по-прост начин:

$$\lambda_{max}^{(k)} \approx \frac{\sum_{s=1}^N \text{sign}(h_{l_0^{(s)}}) \left\{ \prod_{t=1}^{t=k} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) \right\} a f_{l_k^{(s)}}}{\sum_{s=1}^N \text{sign}(h_{l_0^{(s)}}) \left\{ \prod_{t=1}^{t=k-1} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) \right\} f_{l_{k-1}^{(s)}}}. \quad (1.20)$$

Следствие 3. Нека L_1 -нормата на всички вектори-редове на симетричната матрица A удовлетворява условието (1.19). Предполага се, че следното условие е валидно за всички елементи на A :

$$\text{sign}(h_{l_0^{(s)}}) \prod_{t=1}^{t=k-1} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) = 1, \quad s = 1, \dots, N. \quad (1.21)$$

Тогава $\lambda_{max}^{(k)}$ може да се оцени с по-проста формула, както следва:

$$\lambda_{max}^{(k)} \approx \frac{\sum_{s=1}^N \text{sign}(a_{l_{k-1}^{(s)} l_k^{(s)}}) a f_{l_k^{(s)}}}{\sum_{s=1}^N f_{l_{k-1}^{(s)}}}. \quad (1.22)$$

Следва да се отбележи, че чрез Следствие (2) и Следствие (3) се конструират опростени оценки за приближено пресмятане на $\lambda_{max}^{(k)}$ за определени класове симетрични матрици. Тези класове не са празно множество, съгласно Забележка 3.

Забележка 3. Множеството от симетрични матрици, които удовлетворяват условията на Следствие 3 е непразно. Например, нека се разгледа симетрична матрица A с положителни елементи, за която всички елементи по главния диагонал са равни, т.е. $a_{ii} = b > 0$, $i = 1, \dots, n$, и всички извън диагонални елементи също са равни, т.е. $a_{ij} = d > 0$, $i \neq j$, $i, j = 1, \dots, n$.

Тъй като систематичната грешка ϵ_1 зависи от броя k на преходите във веригата на Марков, а стохастичната грешка ϵ_2 , зависи от броя на реализациите N , се получава оценка за съответните грешки:

$$\left| \lambda_{max} - \frac{\bar{\theta}_N^{(k)}}{\bar{\theta}_N^{(k-1)}} \right| = \left| \lambda_{max} - \lambda_{max}^{(k)} + \lambda_{max}^{(k)} - \frac{\bar{\theta}_N^{(k)}}{\bar{\theta}_N^{(k-1)}} \right| \leq \left| \lambda_{max} - \lambda_{max}^{(k)} \right| + \left| \lambda_{max}^{(k)} - \frac{\bar{\theta}_N^{(k)}}{\bar{\theta}_N^{(k-1)}} \right| < \epsilon_1 + \epsilon_2 \quad (1.23)$$

Следва да се отбележи, че балансът между систематичната и стохастичната грешка се реализира при условие $\epsilon_1 = \epsilon_2 = \epsilon$.

Дефиниция 1.3.1. СМК и СКМК методи за оценка на максималната собствена стойност на симетрична матрица при избор на начален вектор на плътност \mathbf{p} и матрица на плътност на прехода P съгласно формула (1.16) се наричат почти оптимални СМК и СКМК методи.

1.3.2 Почти оптимален СМК и почти оптимален СКМК алгоритми: изчислителна сложност

Псевдокодът, представен в **Алгоритъм 1**, описва стъпките за оценяване на максималната собствена стойност на симетрична плътна матрица въз основа на формула (1.18), съгласно Следствие 1.

Algorithm 1 Pseudocode for computing the *Almost Optimal* PQMC algorithm

- 1: INPUT: matrix $A = \{a_{ij}\} \in \mathbb{R}^{n \times n}$; vectors $\mathbf{h}, \mathbf{f} \in \mathbb{R}^n$ and positive integers n, N, k
- 2: COMPUTE: $\|\mathbf{a}_i\| = \sum_{j=1}^n |a_{ij}|$, $1 \leq i \leq n$ and $\|\mathbf{h}\| = \sum_{i=1}^n |h_i|$
- 3: COMPUTE: The coordinates of a permissible density vector \mathbf{p} and the elements of a permissible density matrix P

$$p_i = \frac{|h_i|}{\|\mathbf{h}\|} \quad \text{and} \quad p_{ij} = \frac{|a_{ij}|}{\|\mathbf{a}_i\|}, \quad i, j = 1, \dots, n.$$

- 4: GENERATE: N elements of the $(k+1)$ -dimensional Sobol (or Halton) sequence
- 5: CONSTRUCT: N realizations of the Markov chain with integer elements

$$l_0^{(s)} \rightarrow \dots \rightarrow l_t^{(s)} \rightarrow \dots \rightarrow l_{k-1}^{(s)} \rightarrow l_k^{(s)}, \quad 1 \leq l_t^{(s)} \leq n, \quad 0 \leq t \leq k, \quad 1 \leq s \leq N$$

- 6: COMPUTE:

$$\bar{\theta}_N^{(k)} = \frac{1}{N} \sum_{s=1}^N \text{sign}(h_{l_0^{(s)}}) \|\mathbf{h}\| \left\{ \prod_{t=1}^{t=k} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) \|\mathbf{a}_{l_{t-1}^{(s)}}\| \right\} f_{l_k^{(s)}};$$

$$\bar{\theta}_N^{(k-1)} = \frac{1}{N} \sum_{s=1}^N \text{sign}(h_{l_0^{(s)}}) \|\mathbf{h}\| \left\{ \prod_{t=1}^{t=k-1} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) \|\mathbf{a}_{l_{t-1}^{(s)}}\| \right\} f_{l_{k-1}^{(s)}};$$

- 7: OUTPUT:

$$\lambda_{max}^{(k)} := \frac{\bar{\theta}_N^{(k)}}{\bar{\theta}_N^{(k-1)}}$$

Казваме, че **Алгоритъм 1** е *почти оптимален* СКМК алгоритъм (англ. *almost optimal PQMC algorithm*), когато в стъпка 4 се използват редици с малък дискрепанс, в частност редици на Собол или Холтън. Алгоритъмът се определя като *почти оптимален* СМК алгоритъм (англ. *almost optimal PMC algorithm*), който ще означаваме с **Алгоритъм 1a**, ако в стъпка 4 използваме генератори на псевдослучайни числа (PRNGs).

Забележка 4. Ако елементите на симетричната матрица A удовлетворяват условията в Следствие 2 или Следствие 3, тогава формулите, използвани в стъпка 6 в **Алгоритъм 1** и **Алгоритъм 1a**, могат да бъдат заменени с по-опростените изрази (1.20) или (1.22).

Оценяването на изчислителната сложност, респективно на броя на аритметичните операции на двата алгоритъма, е от съществено значение. Подобни оценки са важни при конструирането на ефективни алгоритми за задачи с големи размерности. В нашия случай е достатъчно да оценим изчислителната сложност на **Алгоритъм 1**. Съгласно представения псевдокод, изчислителната сложност на **Алгоритъм 1** зависи от броя на реализациите

на веригата на Марков, от изчислителното време за пресмятане на една стойност на случайната величина $\theta^{(k)}$, както и от времето, необходимо за определяне на всички елементи на матрицата на преходната плътност P .

Стъпки 2 и 3 в **Алгоритъм 1** изискват $\mathcal{O}(n^2)$ операции. Тази част от алгоритъма се разглежда като предварителна обработка. Основната изчислителна тежест се съдържа в стъпки 4-7. За стъпки 4-6, изчислителните разходи се оценяват като $N\tau_k$, където N е броят на реализациите на веригата на Марков, а τ_k е средното изчислително време за една реализация на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$ при фиксирано k .

Нека с τ_0 означим средното изчислително време за един преход ($l_{t-1} \rightarrow l_t$) във веригата на Марков при конструиране на $\bar{\theta}_N^{(k)}$ или $\bar{\theta}_N^{(k-1)}$ по тегло W_t (виж формула (1.7)). Реализацията на един преход изисква симулиране на дискретна случайна величина с n състояния и вероятности, определени от съответния ред на матрицата P . Изчислителната цена на такова симулиране е $\mathcal{O}(n)$. Следователно $\tau_0 = \mathcal{O}(n)$, а за фиксирано k се получава $\tau_k = k\tau_0 = \mathcal{O}(nk)$. Така общата изчислителна сложност за стъпки 4-6 е:

$$Cost(Alg1) = \mathcal{O}(Nnk) + \mathcal{O}(n^2) = \mathcal{O}(Nnk + n^2).$$

Стъпка 7 в **Алгоритъм 1** изисква $\mathcal{O}(1)$ операции. В обобщение, при условие $N > n$ изчислителната сложност на алгоритъма е $\mathcal{O}(Nnk)$. В този случай разходите за предварителна обработка могат да се пренебрегват, тъй като основната част на алгоритъма доминира по брой операции.

При оценяване на матрично-векторни скаларни произведения с плътни матрици с размери от няколко хиляди до десетки хиляди елементи, балансът между систематичната и стохастичната грешка се постига при $N = \mathcal{O}(n)$ [3, 20, 21]. За разредени многомерни матрици, за които броят на ненулевите елементи е от порядъка $\mathcal{O}(n)$, може да се очаква съществено намаляване на изчислителните разходи.

Числените експерименти, представени в следващия раздел, чрез **Алгоритъм 1** и **Алгоритъм 1a**, показват, че при плътни симетрични матрици с размери от няколкостотин до няколко хиляди елемента балансът между систематичните и стохастичните грешки също се постига при $N = \mathcal{O}(n)$.

1.4 Числени резултати и дискусия

В числените експерименти за *почти оптималния* СМК алгоритъм, обозначен като **Алгоритъм 1a**, е реализиран с помощта на два генератора на псевдослучайни числа (PRNGs): *Mersenne Twister* (MT) [78] и *Middle Square* (MS) [82, 90]. Генераторът MT има период $2^{19937} - 1$ и неговият изход не проявява дългосрочни корелации при разглеждания в 623 измерения. Стандартната реализация използва 32-битова дължина на думата, докато съществува и алтернативна версия с 64-битова дължина [94]. MT е разработен, за да преодолее повечето недостатъци на по-старите генератори на псевдослучайни числа, като например генератора MS, базиран на средноквадратичния метод [82].

В числените експерименти са разгледани две *почти оптимални* СМК реализации на **Алгоритъм 1a**, които при анализа на резултатите са означени като $PMC_{(MT)}$ и $PMC_{(MS)}$,

в зависимост от използвания генератор на псевдослучайни числа (MT и MS). Получените резултати са сравнени с тези на *класическия* СМК алгоритъм, реализиран съгласно *класическия* СМК подход. При тях началният вектор на вероятностната плътност и матрицата на преходната плътност се задават с формулата (1.15). *Класическите* СМК подходи са означени като $PMC_{(MT)}^{(n)}$ и $PMC_{(MS)}^{(n)}$.

В случая на **Алгоритъм 1** са разгледани две *почти оптимални* СКМК реализации, които при анализа на резултатите са означени като $PQMC_{(S)}$ и $PQMC_{(H)}$, в зависимост от използваните редици на Соболю (S) и Холтгън (H). Получените резултати са сравнени с тези на *класическия* СКМК алгоритъм. *Класическите* СМК подходи са означени като $PQMC_{(S)}^{(n)}$ и $PQMC_{(H)}^{(n)}$.

На практика редиците на Соболю и Холтгън обикновено се подлагат на разбъркване (англ. *scrambling*). По този начин, без да се губи предимството на равномерното покритие в единичния хиперкуб, се въвежда контролирана случайност и се редуцират корелациите, които възникват при високи размерности на редиците.

В проведените числени експерименти редиците на Холтгън са конструирани чрез тип разбъркване „RR2“, при който при операция *reverse-radix* се прилага случайно пермутиране на коефициентите на представянето в съответната бройна система [58, 95]. Този подход води до ефективно преразпределение на точките и подобрена равномерност. Редиците на Соболю са генерирани с използването на разбъркването „MatousekAffineOwen“, което комбинира случайно линейно разбъркване [77] със случайно цифрово изместване [88]. Генераторите на псевдослучайни числа MT и MS, както и разбъркваните редици на Соболю и Холтгън, са интегрирани в софтуерната среда MATLAB, версия 9.10 (R2021a), която използваме за нашите експерименти.

Числените резултати за оценяване на най-голямата собствена стойност на симетрични плътни матрици са получени въз основа на две тестови матрици със случайно генерирани елементи. Собствените стойности на тези матрици удовлетворяват условието (1.2). Разглеждат се следните два случая:

- (i) Матрица A е с размерност $n = 100$, $\lambda_{max} = \lambda_1 = 50.0408$, и $\lambda_2 = 4.0522$.
- (ii) Матрица A е с размерност $n = 500$, $\lambda_{max} = \lambda_1 = 250.2454$, и $\lambda_2 = 9.0721$.

Произволните вектори \mathbf{h} и \mathbf{f} са избрани с L_1 -норма, т.е.:

$$\|\mathbf{h}\| = \|\mathbf{f}\| = \sum_{i=1}^n \frac{1}{n} = 1.$$

Този избор е направен за различните размерности $n = 100$ и $n = 500$ на случайно генерираната симетрична матрица A . При стохастичните итерационни алгоритми условието за сходимост е нормата на матрицата да бъде строго по-малка от единица. Поради това при реализацията на СМК и СКМК алгоритмите матрицата A се мащабира с положителен параметър:

$$q = \frac{1}{|\lambda_1| + \delta_1}, \quad \delta_1 > 0,$$

така, че да бъде изпълнено условието $\|qA\| < 1$. В случая, когато предварителна оценка за най-голямата собствена стойност λ_1 не е налична, параметърът q може да бъде избран по алтернативен начин, а именно:

$$q = \frac{1}{\max_i \|\mathbf{a}_i\| + \delta}, \quad \delta > 0, \quad i = 1, \dots, n,$$

където $\|\mathbf{a}_i\|$ е L_1 -нормата на i -тия ред на A .

В числените експерименти собствените стойности λ_1 и λ_2 са избрани да бъдат достатъчно добре отделени, за да сме сигурни, че скоростта на сходимост на итерационния процес е добра. Това означава, че систематичната грешка да бъде пренебрежимо малка след приблизително четири до пет итерации. Тази постановка позволява да се анализира експериментално баланса между стохастичната и систематичната грешка. По-специално се изследва влиянието на броя на преходите k във веригата на Марков и броя на реализациите N на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$ върху този баланс. Освен това се анализира ефектът от намаляването на дисперсията в *почти оптималните* СМК/СКМК алгоритми върху точността на полученото приближение при фиксирана изчислителна сложност.

1.4.1 Случай (i): Размер на матрицата $n = 100$

В Таблица 1.1 са представени числените резултати за оценяване на най-голямата собствена стойност на симетричната матрица за случай (i), получени с помощта на *почти оптималния* $PMC_{(MT)}$ алгоритъм и *класическия* $PMC_{(MT)}^{(n)}$ алгоритъм при различни стойности на N и k . Аналогично, в Таблица 1.2 са показани съответните резултати за *почти оптималния* $PMC_{(MS)}$ алгоритъм и *класическия* $PMC_{(MS)}^{(n)}$ алгоритъм. В двете таблици са включени и абсолютните грешки, съпътстващи получените оценки. В тези експерименти максималната собствена стойност $\lambda_1 = 50.0408$ на симетричната матрица A с размерност $n = 100$ е оценявана при различен брой реализации N и преходи k във веригата на Марков.

Числените резултати показват, че при фиксиран брой реализации N увеличаването на броя преходи k води до намаляване на абсолютната грешка. Това намаляване се стабилизира до достигане на определена стойност на k , което свидетелства за наличие на оптимален баланс между систематичната и стохастичната грешка. При фиксирано $N = 128$ тази стойност е $k = 9$, докато при $N > 128$ оптималният брой на преходите е $k = 8$. По-нататъшното увеличаване на k не води до допълнително подобряване на точността. Минималните абсолютни грешки за всеки от алгоритмите $PMC_{(MT)}$, $PMC_{(MT)}^{(n)}$, $PMC_{(MS)}$ и $PMC_{(MS)}^{(n)}$ са отбелязани с удебелен шрифт в Таблица 1.1 и Таблица 1.2. Най-малката абсолютна грешка се наблюдава при $k = 8$, $N = 512$ за *почти оптималните* СМК алгоритми (виж Таблица 1.1 и Таблица 1.2). Това показва, че тези стойности на параметрите k и N осигуряват най-добър баланс между систематичната и стохастичната грешка. За *класическите* СМК алгоритми оптималният баланс се постига при $k = 8$ и $N = 2048$.

Сравнението между резултатите на *почти оптималните* СМК и *класическите* СМК алгоритми показва, че при фиксирани N и k абсолютната грешка при *почти оптималните* алгоритми е значително по-малка. Например, при $N = 512$ алгоритъмът $PMC_{(MT)}$ постига абсолютна грешка 5×10^{-5} , докато $PMC_{(MT)}^{(n)}$ достига грешка 1.5×10^{-3} (виж

Таблица 1.1). Аналогична тенденция се наблюдава и при сравнението между алгоритмите $PMC_{(MS)}$ и $PMC_{(MS)}^{(n)}$ в Таблица 1.2.

Таблица 1.1: Резултати от $PMC_{(MT)}$ и $PMC_{(MT)}^{(n)}$ при различни N и k за матрица A с размерност $n = 100$, за която $\lambda_2/\lambda_1 \approx 0.081$, където $\lambda_1 = 50.0408$ и $\lambda_2 = 4.0522$.

N	k	$\lambda_1^{(k)}$ using	Absolute Error	$\lambda_1^{(k)}$ using	Absolute Error
		$PMC_{(MT)}$	$PMC_{(MT)}$	$PMC_{(MT)}^{(n)}$	$PMC_{(MT)}^{(n)}$
128	6	50.012316	0.02848	49.893384	0.1474
	7	50.025579	0.01522	49.997142	0.0437
	8	5.039901	0.00090	50.019883	0.0209
	9	50.040709	0.00009	50.037795	0.0030
	10	50.045295	0.00450	50.066104	0.0253
	11	50.054378	0.01358	50.104543	0.0637
256	6	50.017864	0.02294	49.919982	0.1208
	7	50.023923	0.01688	50.012158	0.0286
	8	50.040628	0.00017	50.039528	0.0013
	9	50.044896	0.00410	50.050593	0.0098
	10	50.050896	0.01010	50.081757	0.0410
	11	50.070174	0.02937	50.093002	0.0522
512	6	50.032910	0.00789	49.676084	0.3647
	7	50.035045	0.00576	49.996040	0.0448
	8	50.040750	0.00005	50.039264	0.0015
	9	50.052027	0.01123	50.050158	0.0094
	10	50.058593	0.01779	50.072966	0.0322
	11	50.060141	0.01934	50.109447	0.0686
1024	6	50.015868	0.02493	49.770608	0.2702
	7	50.029828	0.01097	50.019435	0.0214
	8	50.040720	0.00008	50.039735	0.0011
	9	50.045403	0.00460	50.054351	0.0136
	10	50.053595	0.01280	50.069878	0.0291
	11	50.060302	0.01950	50.080112	0.0393
2048	6	50.035629	0.00517	49.851866	0.1889
	7	50.036291	0.00451	50.029563	0.0112
	8	50.041086	0.00029	50.041728	0.0009
	9	50.043850	0.00305	50.057179	0.0164
	10	50.066103	0.02530	50.077047	0.0362
	11	50.069460	0.02866	50.091154	0.0504

*Маркираните редове представляват оптималните k стойности за всяко N .

На Фигура 1.1 са представени оценките на собствените стойности при нарастващ брой реализации N (по оста X), като е използвана оптималната стойност на броя на преходите (k), определена от Таблица 1.1 и Таблица 1.2. Почти оптималните СМК алгоритми са сравнени със съответните им *класически* варианти. Резултатите показват, че при *класическите* варианти алгоритъмът $PMC_{(MT)}^{(n)}$ демонстрира по-висока точност и стабилност при увеличаване на N от $PMC_{(MS)}^{(n)}$. Според Таблица 1.1, при $N = 2048$ алгоритъмът $PMC_{(MT)}^{(n)}$ достига абсолютна грешка от 9×10^{-4} , като превъзхожда $PMC_{(MS)}^{(n)}$, за който

грешката е 1.8×10^{-3} (виж Таблица 1.2).

Визуализацията на Фигура 1.1 потвърждават, че с увеличаване на N получените оценки с алгоритъма $PMC_{(MT)}^{(n)}$ последователно се доближават до точната стойност λ_{max} . За разлика от него, алгоритъмът $PMC_{(MS)}^{(n)}$ показва по-големи колебания и в някои случаи надценява точната стойност при по-големи N , което показва по-малка стабилност. Това може да се дължи на псевдослучайни вариации и нарушаване на равномерността при използвания MS генератор в по-високи размерности.

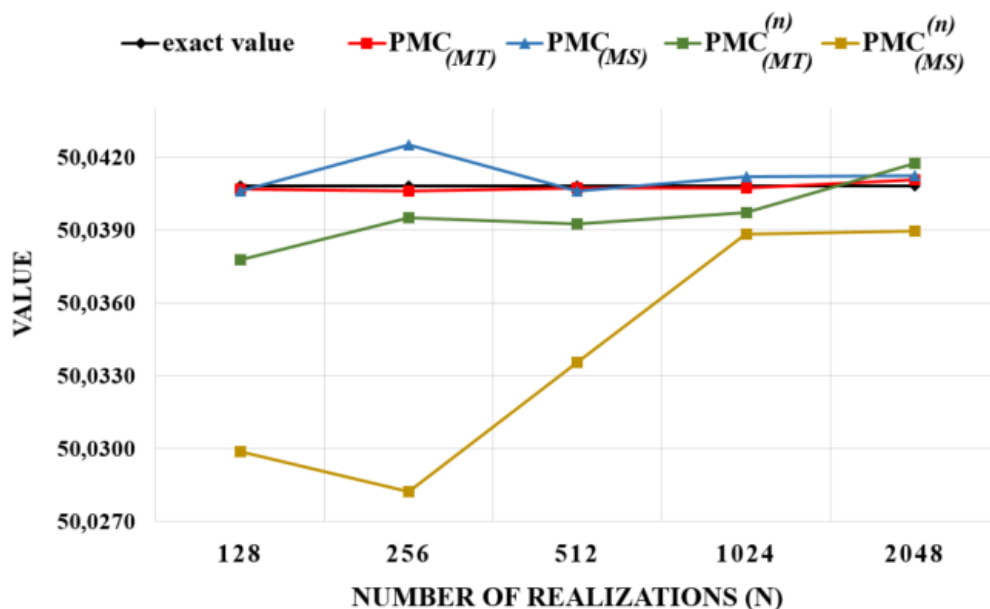
Таблица 1.2: Резултати от $PMC_{(MS)}^{(n)}$ и $PMC_{(MT)}^{(n)}$ при различни N и k за матрицата A с размерност $n = 100$, за която $\lambda_2/\lambda_1 \approx 0.081$, където $\lambda_1 = 50.0408$ и $\lambda_2 = 4.0522$.

N	k	$\lambda_1^{(k)}$ using	Absolute Error	$\lambda_1^{(k)}$ using	Absolute Error
		$PMC_{(MS)}$	$PMC_{(MS)}$	$PMC_{(MS)}^{(n)}$	$PMC_{(MS)}^{(n)}$
128	6	50.019125	0.02167	49.564508	0.4763
	7	50.023944	0.01686	49.988941	0.0519
	8	50.039440	0.00136	50.019388	0.0214
	9	50.040595	0.00020	50.029883	0.0109
	10	50.051262	0.01046	50.060923	0.0201
	11	50.084762	0.04396	50.109231	0.0684
256	6	50.011148	0.02965	49.574876	0.4659
	7	50.031044	0.00976	49.984883	0.0559
	8	50.042490	0.00169	50.028211	0.0126
	9	50.050490	0.00969	50.051251	0.0105
	10	50.076397	0.03560	50.084891	0.0441
	11	50.094549	0.05375	50.103125	0.0623
512	6	50.028121	0.01268	49.679757	0.3610
	7	50.030518	0.01028	49.997989	0.0428
	8	50.040617	0.00018	50.033570	0.0072
	9	50.058913	0.01811	50.053570	0.0128
	10	50.061048	0.02025	50.091035	0.0502
	11	50.064531	0.02373	50.109567	0.0688
1024	6	50.012038	0.02876	49.752343	0.2885
	7	50.031472	0.00933	50.009835	0.0310
	8	50.041209	0.00041	50.038850	0.0020
	9	50.044938	0.00414	50.056396	0.0156
	10	50.048139	0.00734	50.076755	0.0360
	11	50.053169	0.01237	50.093669	0.0529
2048	6	50.031902	0.00890	49.839506	0.02013
	7	50.038998	0.00180	50.022658	0.0181
	8	50.041243	0.00044	50.038951	0.0018
	9	50.043045	0.00225	50.063537	0.0227
	10	50.058875	0.01808	50.078798	0.0380
	11	50.084113	0.04331	50.098828	0.0580

*Маркираните редове представляват оптималните k стойности за всяко N .

Двете почти оптимални версии ($PMC_{(MT)}$ и $PMC_{(MS)}$) демонстрират сравнително равномерни траектории близки до точната собствена стойност и показват по-висока точ-

ност при по-малки стойности на N в сравнение с техните *класически* аналози (виж Фигура 1.1). Това се дължи на намалената дисперсия, произтичаща от избора на матрицата на преходната плътност на веригата на Марков (виж Таблица 1.9). Получените резултати ясно показват, че *почти оптималните* СМК алгоритми превъзхождат *класическите* СМК алгоритми както по отношение на точността, така и по изчислителната ефективност, особено когато се използва надежден генератор на псевдослучайни числа като МТ.



Фигура 1.1: Графично сравняване на приближените резултати за $\lambda_{max}^{(k)}$ с $\lambda_{max} = 50.0408$ при балансиран случай за грешките, т.е. когато $(k = 9, N = 128)$ и $(k = 8, N > 128)$. Приложени алгоритми: $PMC_{(MT)}$, $PMC_{(MS)}$, $PMC_{(MT)}^{(n)}$ и $PMC_{(MS)}^{(n)}$.

В Таблица 1.3 и Таблица 1.4 са представени резултатите, получени с *почти оптималните* и *класическите* СКМК алгоритми, при използване на редици на Холтърн и Соболев за изчисляване на максималната собствена стойност на симетричната матрица с размер $n = 100$. Максималната собствена стойност $\lambda_1 = 50.0408$ е оценена при различен брой реализации N и преходи k във веригата на Марков. За всеки от алгоритмите $PQMC_{(H)}$, $PMC_{(H)}^{(n)}$, $PQMC_{(S)}$ и $PQMC_{(S)}^{(n)}$ се наблюдава намаляване на абсолютната грешка с увеличаване на броя на преходите k , като оптимална сходимост обикновено се постига при $k = 11$ или $k = 12$ (виж Таблица 1.3 и Таблица 1.4). Алгоритъмът $PQMC_{(S)}$ систематично превъзхожда $PQMC_{(H)}$ за всички стойности на N , демонстрирайки по-малки грешки при оценяване на точното решение и по-гладка сходимост. Например, при $N = 2048$, $PQMC_{(S)}$ достига абсолютна грешка от 5×10^{-5} , докато при $PQMC_{(H)}$ тя е равна на 8.2×10^{-4} (виж Таблица 1.3 и Таблица 1.4). Това поведение отразява по-добрата равномерност на редиците на Соболев при по-високи размерности.

Класическите варианти ($PQMC_{(H)}^{(n)}$, $PQMC_{(S)}^{(n)}$) също показват устойчивост при големи стойности на N . Въпреки че при малки стойности на N грешките са значително големи, с увеличаване на броя на реализациите те постепенно намаляват. При $N = 2048$ алгоритъмът $PQMC_{(S)}^{(n)}$ постига абсолютна грешка 0.0336 в сравнение с 0.0707 за $PQMC_{(H)}^{(n)}$,

Таблица 1.3: Резултати от $PQMC_{(H)}$ и $PQMC_{(H)}^{(n)}$ при различни N и k за матрица A с размерност $n = 100$, за която $\lambda_2/\lambda_1 \approx 0.081$, където $\lambda_1 = 50.0408$ и $\lambda_2 = 4.0522$.

N	k	$\lambda_1^{(k)}$ using	Absolute Error	$\lambda_1^{(k)}$ using	Absolute Error
		$PQMC_{(H)}$	$PQMC_{(H)}$	$PQMC_{(H)}^{(n)}$	$PQMC_{(H)}^{(n)}$
128	9	49.667323	0.37348	49.051280	0.9895
	10	49.732875	0.30792	49.313447	0.7274
	11	49.801798	0.23900	49.574856	0.4659
	12	50.037989	0.00281	50.206208	0.1654
	13	50.152986	0.11219	50.314121	0.2733
	14	50.246958	0.20616	50.542813	0.5020
256	9	49.982000	0.05880	49.674872	0.3659
	10	50.001834	0.03897	49.739820	0.3010
	11	50.038996	0.00180	50.199525	0.1587
	12	50.067986	0.02719	50.229280	0.1885
	13	50.135149	0.09435	50.299014	0.2582
	14	50.379962	0.33916	50.387742	0.3469
512	9	49.987751	0.05305	49.676566	0.3642
	10	50.017989	0.02281	49.749688	0.2911
	11	50.039899	0.00090	50.176724	0.1359
	12	50.178512	0.13771	50.196566	0.1558
	13	50.228992	0.18819	50.299461	0.2587
	14	50.357840	0.31704	50.358776	0.3180
1024	9	49.452553	0.58825	49.453967	0.5868
	10	49.976987	0.06381	49.806135	0.2347
	11	50.039772	0.00103	50.172345	0.1315
	12	50.107985	0.06719	50.295866	0.2551
	13	50.282004	0.24120	50.298343	0.2575
	14	50.329662	0.28886	50.339248	0.2984
2048	9	49.681987	0.35881	49.670730	0.3701
	10	49.973344	0.06746	49.907160	0.1336
	11	50.039982	0.00082	50.111519	0.0707
	12	50.183245	0.14245	50.275169	0.2344
	13	50.238237	0.19744	50.291489	0.2507
	14	50.407066	0.36627	50.407572	0.3668

*Маркираните редове представляват оптималните k стойности за всяко N .

което потвърждава предимството на СКМК алгоритъма, базиран на редицата на Собол, дори в *класическата* конфигурация. Сравнението между *почти оптималните* и *класическите* СКМК алгоритми (виж Таблица 1.3 и Таблица 1.4) показва, че за фиксирани стойности на N и k абсолютната грешка при *почти оптималните* алгоритми е значително по-малка. Например, при $N = 256$ и $k = 11$, $PQMC_{(S)}$ постига грешка от 8.1×10^{-4} , докато при $PQMC_{(S)}^{(n)}$ тя е 0.1455 (виж Таблица 1.4). Аналогична тенденция се наблюдава и при сравнението между $PQMC_{(H)}$ и $PQMC_{(H)}^{(n)}$ в Таблица 1.3, където при същите параметри грешките са съответно 1.8×10^{-3} и 0.1587. Това значително подобрене се дължи на подходът при конструиране на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$ за *почти оптималния*

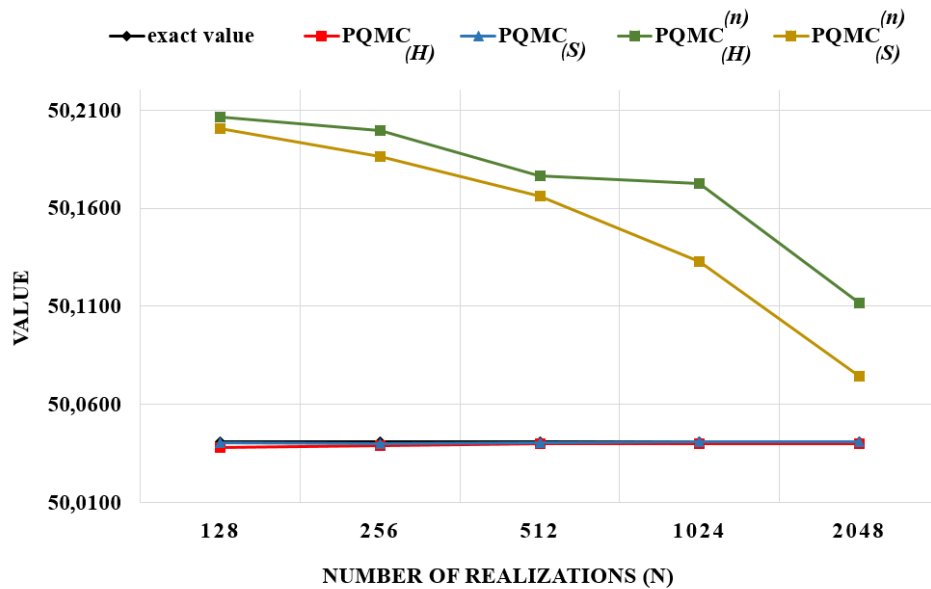
СКМК алгоритъм.

Фигура 1.2 допълва анализа чрез визуално сравнение между *класическите* и *почти оптималните* СКМК алгоритми. Алгоритмите $PQMC_{(S)}$ и $PQMC_{(H)}$ демонстрират висока точност и стабилност, при оценяване на максималната собствена стойност за всички размери N на извадката. За разлика от тях, $PQMC_{(S)}^{(n)}$ и $PQMC_{(H)}^{(n)}$ първоначално надценяват λ_{max} и се приближават към нея значително по-бавно с увеличаване на броя на реализациите N . Тези резултати ясно подчертават превъзходството на *почти оптималните* СКМК алгоритми спрямо *класическите* им аналози както по отношение на точността, така и на изчислителната сложност.

Таблица 1.4: Резултати от $PQMC_{(S)}$ и $PQMC_{(S)}^{(n)}$ при различни N и k за матрица A с размерност $n = 100$, за която $\lambda_2/\lambda_1 \approx 0.081$, където $\lambda_1 = 50.0408$ и $\lambda_2 = 4.0522$.

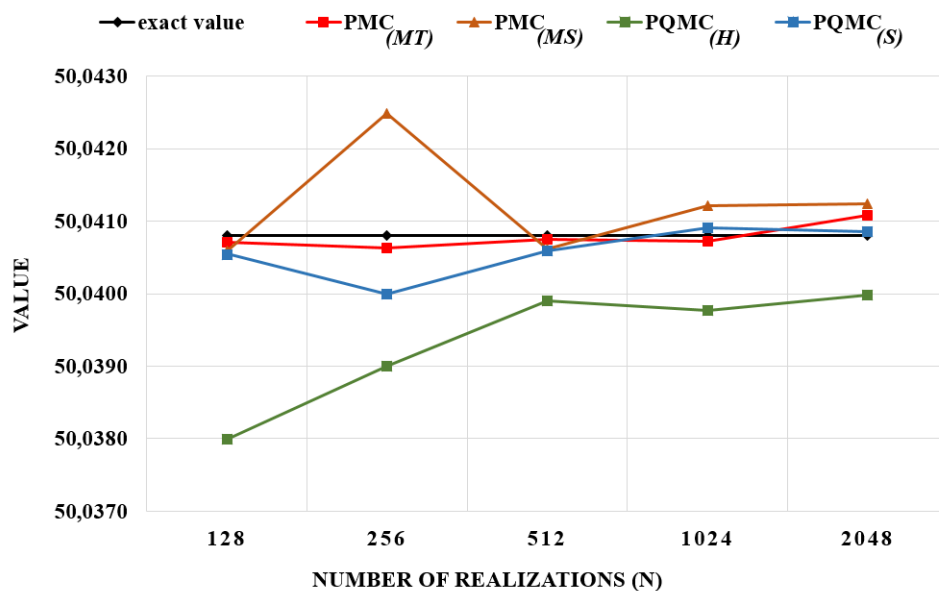
N	k	$\lambda_1^{(k)}$ using	Absolute Error	$\lambda_1^{(k)}$ using	Absolute Error
		$PQMC_{(S)}$	$PQMC_{(S)}$	$PQMC_{(S)}^{(n)}$	$PQMC_{(S)}^{(n)}$
128	9	49.393214	0.64759	49.215148	0.8257
	10	49.597734	0.44307	49.491180	0.5496
	11	49.679383	0.36142	49.624160	0.4166
	12	50.040544	0.00026	50.200481	0.1597
	13	50.091116	0.05032	50.299826	0.2590
	14	50.181771	0.14097	50.399651	0.3589
256	9	49.780806	0.25999	49.428552	0.6122
	10	49.994318	0.04648	49.742961	0.2978
	11	50.039989	0.00081	50.186264	0.1455
	12	50.049407	0.00861	50.225854	0.1851
	13	50.153503	0.11270	50.263529	0.2227
	14	50.327859	0.28706	50.386282	0.3455
512	9	49.966398	0.07440	49.543755	0.4970
	10	50.050000	0.00920	49.866152	0.1746
	11	50.040585	0.00021	50.166400	0.1256
	12	50.148019	0.10722	50.223877	0.1831
	13	50.209448	0.16865	50.271020	0.2302
	14	50.456838	0.41604	50.383010	0.3422
1024	9	49.765493	0.27531	49.617372	0.4234
	10	49.986946	0.05385	49.933245	0.1076
	11	50.040913	0.00011	50.132603	0.0918
	12	50.087435	0.04664	50.167252	0.1265
	13	50.218753	0.17795	50.280149	0.2393
	14	50.298935	0.25814	50.288215	0.2474
2048	9	49.707851	0.33295	49.681502	0.3593
	10	49.972883	0.06792	49.957605	0.0832
	11	50.040854	0.00005	50.074364	0.0336
	12	50.155992	0.11519	50.186986	0.1462
	13	50.222261	0.18146	50.239941	0.1991
	14	50.304974	0.26417	50.407318	0.3665

*Маркираните редове представляват оптималните k стойности за всяко N .



Фигура 1.2: Графично сравняване на приближените резултати за $\lambda_{max}^{(k)}$ с $\lambda_{max} = 50.0408$ при балансирания случай за грешките, т.е. когато $(k = 12, N = 128)$ и $(k = 11, N > 128)$. Приложени алгоритми: $PPMC_{(S)}$, $PPMC_{(H)}$, $PPMC_{(S)}^{(n)}$ и $PPMC_{(H)}^{(n)}$.

Фигура 1.3 сравнява *почти оптималните* СМК алгоритми, използващи генераторите МТ и MS, с *почти оптималните* СКМК алгоритми, базирани на разбърканите редици на Холтън и Собол. Скалата по оста Y е прецизирана така, че всички приближени стойности на $\lambda_{max}^{(k)}$ попадат в интервал с ширина $\delta < 0.005$ около точната стойност λ_{max} , което позволява ясно разграничаване между различните варианти на *почти оптималните* алгоритми.



Фигура 1.3: Графично сравнение на λ_{max} на симетричната матрица A ($n = 100$) с $\lambda_{max}^{(k)}$ в балансиращите случаи. Използвани са *почти оптималните* СМК и СКМК алгоритми.

Алгоритъмът $PMC_{(MT)}$ демонстрира най-добра сходимост сред СМК вариантите, като достига минимална грешка 5×10^{-5} при $N = 512$ и $k = 8$ (виж Таблица 1.1). За разлика от него, $PMC_{(MS)}$ (виж Таблица 1.2) показва малко по-големи грешки, макар и все още в допустими граници (например 1.8×10^{-4} при $N = 512$ и $k = 8$), което се потвърждава и визуално във Фигура 1.3.

От *почти оптималните* СКМК алгоритми, алгоритъмът $PQMC_{(S)}$ се откроява като по точен, демонстрирайки минимално отклонение от точната стойност, както е показано на Фигура 1.3, така и в табличните резултати. Например, при $N = 2048$ и $k = 11$, абсолютната грешка е 5×10^{-5} . В сравнение с него, при същите параметри $PQMC_{(H)}$ достига грешка 8.2×10^{-4} (виж Таблица 1.3).

Числените експерименти за плътната симетрична матрица с размер $n = 100$ показват, че оптималният баланс между стохастичната и систематичната грешка се постига при $k = 8$ за алгоритъма $PMC_{(MT)}$ и $k = 11$ за $PQMC_{(S)}$, независимо от броя реализации N . Резултатите демонстрират, че *почти оптималният* СМК алгоритъм, базиран на генератора МТ, както и *почти оптималният* СКМК алгоритъм, използващ разбъркани редици на Соболев, осигуряват значително по-висока точност. Това превъзходство се наблюдава спрямо вариантите с генератор MS, разбъркани редици на Холтън и техните *класически* аналози при еднакви изчислителни разходи.

1.4.2 Случай (ii): Размер на матрицата $n = 500$

Таблица 1.5 представя числените резултати за оценяването на λ_{max} на симетрична плътна матрица A с размерност $n = 500$ при различни стойности на броя на реализациите N и броя преходи k във веригата на Марков. Резултатите са получени с помощта на *почти оптималния* алгоритъм $PMC_{(MT)}$ и *класическия* алгоритъм $PMC_{(MT)}^{(n)}$. Таблица 1.6 представя съответните резултати за *почти оптималния* алгоритъм $PMC_{(MS)}$ и *класическия* алгоритъм $PMC_{(MS)}^{(n)}$. В двете таблици са включени и абсолютните грешки, съпътстващи оценките.

Числените експерименти показват, че при фиксиран брой реализации N увеличаването на броя на преходите k във веригата на Марков води до намаляване на абсолютната грешка. Това намаляване се стабилизира след достигане на определена стойност на k , което потвърждава наличието на оптимален баланс между систематичната и стохастична грешка. В разглеждания случай тази стойност е $k = 9$ (виж Таблица 1.5 и Таблица 1.6). При по-нататъшно увеличаване на k точността на оценките не се подобрява съществено. Минималните абсолютни грешки за всеки от алгоритмите ($PMC_{(MT)}$, $PMC_{(MT)}^{(n)}$, $PMC_{(MS)}$, и $PMC_{(MS)}^{(n)}$) при $k = 9$ са подчертани в Таблица 1.5 и Таблица 1.6.

Най-малката абсолютна грешка се наблюдава при ($k = 9, N = 512$) за алгоритъма $PMC_{(MT)}$ и при ($k = 9, N = 128$) за $PMC_{(MS)}$ (виж Таблица 1.5 и Таблица 1.6), което показва, че тези стойности на k и N осигуряват най-добрия баланс между систематичната и стохастичната грешка. При *класическите* СМК алгоритми оптималният баланс между двата вида грешки също се наблюдава при $k = 9$ (виж Таблица 1.5 и Таблица 1.6).

Сравнението между резултатите от *почти оптималните* и *класическите* СМК алгоритми (виж Таблица 1.5 и Таблица 1.6) показва, че при фиксирани стойности на N и

Таблица 1.5: Резултати от $PMC_{(MT)}$ и $PMC_{(MT)}^{(n)}$ при различни N и k за матрица A с размерност $n = 500$, за която $\lambda_2/\lambda_1 \approx 0.0363$, където $\lambda_1 = 250.2454$ и $\lambda_2 = 9.0721$.

N	k	$\lambda_1^{(k)}$ using	Absolute Error	$\lambda_1^{(k)}$ using	Absolute Error
		$PMC_{(MT)}$	$PMC_{(MT)}$	$PMC_{(MT)}^{(n)}$	$PMC_{(MT)}^{(n)}$
128	6	250.181100	0.06430	248.863095	1.3823
	7	250.196274	0.04913	249.073556	1.1718
	8	250.222396	0.02300	249.492834	0.7526
	9	250.232288	0.01311	250.032970	0.2124
	10	250.295027	0.04963	250.149293	0.0961
	11	250.301310	0.05591	250.391291	0.1459
256	6	250.071617	0.17378	249.499761	0.7456
	7	250.114415	0.13098	249.807572	0.4378
	8	250.228998	0.01640	250.124994	0.1204
	9	250.250130	0.00473	250.222275	0.0231
	10	250.283657	0.03826	250.311181	0.0658
	11	250.327734	0.08233	250.363882	0.1185
512	6	250.072767	0.17263	249.680711	0.5647
	7	250.161043	0.08436	250.072663	0.1727
	8	250.235302	0.01011	250.115109	0.1303
	9	250.245110	0.00029	250.192305	0.0531
	10	250.268302	0.02290	250.328365	0.0830
	11	250.306227	0.06083	250.458641	0.2132
1024	6	250.167399	0.07800	249.592590	0.6528
	7	250.211423	0.03398	250.075384	0.1700
	8	250.235306	0.01009	250.164200	0.0812
	9	250.245015	0.00038	250.205542	0.0399
	10	250.269913	0.02451	250.300328	0.0549
	11	250.287417	0.04202	250.495723	0.2503
2048	6	250.160017	0.08538	249.864998	0.3804
	7	250.193403	0.05200	250.039826	0.2056
	8	250.226172	0.01923	250.175939	0.0695
	9	250.246938	0.00154	250.224397	0.0210
	10	250.251779	0.00638	250.295497	0.0501
	11	250.258464	0.01306	250.480189	0.2348

*Маркираните редове представляват оптималните k стойности за всяко N .

k абсолютната грешка на *почти оптималните* алгоритми е значително по-малка. В балансирания случай при $k = 9$ грешката при *почти оптималните* СМК алгоритми е от порядъка на 10^{-4} , докато при *класическите* СМК алгоритми е от порядъка на 10^{-2} . Това подобрене се дължи на същественото намаляване на дисперсията при конструирането на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$ в *почти оптималните* СМК алгоритми (виж Таблица 1.9).

Фигура 1.4 илюстрира оценките за максималната собствена стойност при нарастващ брой реализации N като се използва оптималният брой преходи $k = 9$, обозначен с удебелен шрифт в Таблица 1.5 и Таблица 1.6. *Почти оптималните* СМК алгоритми са

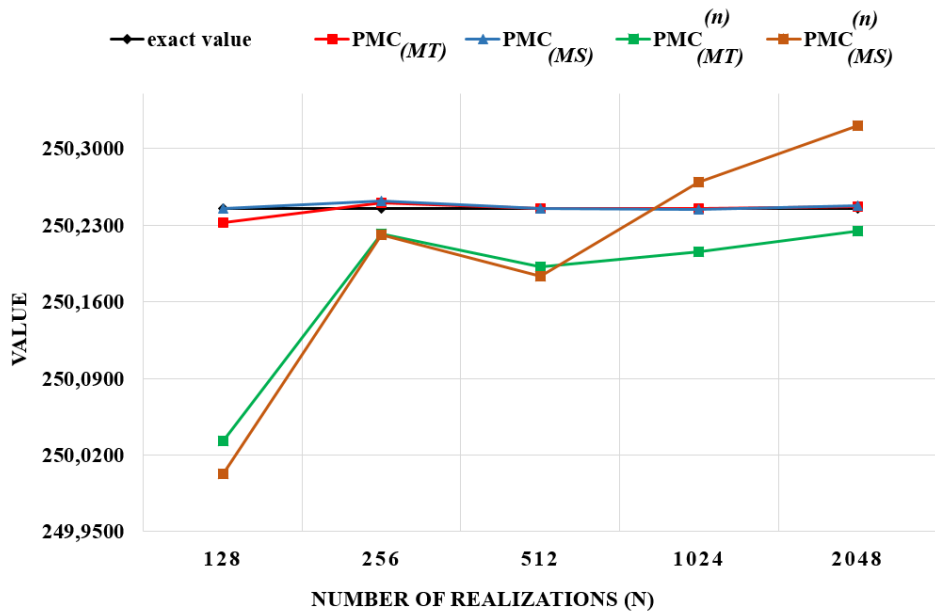
Таблица 1.6: Резултати от $PMC_{(MS)}$ и $PMC_{(MS)}^{(n)}$ при различни N и k за матрица A с размерност $n = 500$, за която $\lambda_2/\lambda_1 \approx 0.0363$, където $\lambda_1 = 250.2454$ и $\lambda_2 = 9.0721$.

N	k	$\lambda_1^{(k)}$ using	Absolute Error	$\lambda_1^{(k)}$ using	Absolute Error
		$PMC_{(MS)}$	$PMC_{(MS)}$	$PMC_{(MS)}^{(n)}$	$PMC_{(MS)}^{(n)}$
128	6	250.085322	0.16008	248.522547	1.7229
	7	250.185322	0.06008	248.931561	1.3138
	8	250.211445	0.03395	249.462722	0.7827
	9	250.245247	0.00015	250.003068	0.2423
	10	250.279476	0.03408	250.120230	0.1252
	11	250.356308	0.11109	250.408768	0.1634
256	6	250.104179	0.14122	249.129460	1.1159
	7	250.177974	0.06743	249.701820	0.5436
	8	250.203720	0.04168	250.015811	0.2296
	9	250.252044	0.00664	250.221130	0.0243
	10	250.290513	0.04511	250.363882	0.1185
	11	250.308344	0.06294	250.668091	0.4227
512	6	250.086724	0.15868	249.219447	1.0260
	7	250.177427	0.06797	249.501899	0.7435
	8	250.206884	0.03852	250.058060	0.1873
	9	250.244890	0.00051	250.183410	0.0620
	10	250.283279	0.03788	250.356760	0.1114
	11	250.297526	0.05213	250.501899	0.2565
1024	6	250.017168	0.22823	249.602843	0.6426
	7	250.189383	0.05602	250.051538	0.1939
	8	250.238977	0.00642	250.143770	0.1016
	9	250.244751	0.00065	250.268955	0.0236
	10	250.270855	0.02546	250.396782	0.1514
	11	250.306134	0.06073	250.493687	0.2483
2048	6	250.118546	0.12685	249.911061	0.3343
	7	250.187878	0.05752	250.034691	0.2107
	8	250.237612	0.00779	250.103221	0.1422
	9	250.247710	0.00231	250.320873	0.0755
	10	250.275206	0.02981	250.347819	0.1024
	11	250.290595	0.04520	250.494239	0.2488

*Маркираните редове представляват оптималните k стойности за всяко N .

сравнени със съответните им *класически* варианти. Алгоритъмът $PMC_{(MT)}^{(n)}$ демонстрира добра точност като при увеличаване на N оценките постепенно се приближават към λ_{max} . Визуалните резултати потвърждават тази тенденция: $PMC_{(MT)}^{(n)}$ се приближава устойчиво към точната стойност, докато $PMC_{(MS)}^{(n)}$ показва колебания и при по-големи N дори надценява λ_{max} . Това посочва възможна нестабилност и по-слаба равномерност в по-големи пространства, дължаща се на избрания генератор MS.

Фигура 1.4 демонстрира, че $PMC_{(MT)}$ и $PMC_{(MS)}$ поддържат сравнително гладка траектория много близка до собствената стойност λ_{max} и постигат по-висока точност при по-малки N в сравнение с *класическите* си аналози. Това поведение отново се обяснява



Фигура 1.4: Графично сравняване на λ_{max} за матрицата A ($n = 500$) с приближените резултати за $\lambda_{max}^{(k)}$, получени при балансирания случай на грешките, т.е. когато ($k = 9, N \geq 128$). Приложени алгоритми: $PMC_{(MT)}$, $PMC_{(MS)}$, $PMC_{(MT)}^{(n)}$ и $PMC_{(MS)}^{(n)}$.

с намалената дисперсия, произтичаща от избора на преходната плътност на веригата на Марков. Резултатите потвърждават, че *почти оптималните* СКМК алгоритми превъзхождат *класическите* по отношение както на точност, така и по изчислителна ефективност, особено при използване на надежден генератор като МТ.

Таблица 1.7 и Таблица 1.8 представят числените резултати за оценяване на λ_{max} , получени чрез *почти оптималните* и *класическите* СКМК алгоритми, използващи съответно рандомизирани редици на Холтън и Собол. В двете таблици максималната собствена стойност $\lambda_1 = 250.2454$ е оценена за различни стойности на N и k . За всички алгоритми ($PQMC_{(H)}$, $PQMC_{(H)}^{(n)}$, $PQMC_{(S)}$ и $PQMC_{(S)}^{(n)}$) абсолютната грешка намалява с увеличаване на k , като оптималното приближение се наблюдава при $k = 12$. Алгоритмите $PQMC_{(S)}$ и $PQMC_{(H)}$ постигат значително по-малки грешки и по-гладка сходимост в сравнение с *класическите* им варианти. Например, при $N = 2048$, $PQMC_{(S)}$ достига грешка 5.4×10^{-4} , докато при $PQMC_{(H)}$ тя е 4.59×10^{-3} (виж Таблица 1.7 и Таблица 1.8). Това отразява по-добрата равномерност на редиците на Собол при по-високи размерности. В балансирания случай при $k = 12$ абсолютната грешка е от порядъка на 10^{-4} за $PQMC_{(S)}$ и от порядъка 10^{-3} за $PQMC_{(H)}$.

Макар че *класическите* СКМК алгоритми първоначално показват по-големи грешки, те все пак демонстрират сравнително добра сходимост при по-големи N . Така например, при $N = 2048$, алгоритъмът $PQMC_{(S)}^{(n)}$ достига грешка 0.0455, докато при $PQMC_{(H)}^{(n)}$ грешката е 0.0726, което потвърждава предимството на алгоритмите, базирани на редиците на Собол, дори в класическа конфигурация.

Сравнението между *почти оптималните* и *класическите* СКМК алгоритми (виж Таблица 1.7 и Таблица 1.8), демонстрира, че при фиксирани N и k абсолютната грешка за *почти оптималните* варианти е значително по-малка. Например, при $N = 256$ и $k = 12$,

Таблица 1.7: Резултати от $PQMC_{(H)}$ и $PQMC_{(H)}^{(n)}$ при различни N и k за матрица A с размерност $n = 500$. Абсолютната грешка е за случая $\lambda_2/\lambda_1 \approx 0.0363$.

N	k	$\lambda_1^{(k)}$ using	Absolute Error	$\lambda_1^{(k)}$ using	Absolute Error
		$PQMC_{(H)}$	$PQMC_{(H)}$	$PQMC_{(H)}^{(n)}$	$PQMC_{(H)}^{(n)}$
128	9	249.471697	0.77370	245.712684	4.5327
	10	250.066438	0.17896	246.512652	3.7327
	11	250.167707	0.07769	251.486594	1.2412
	12	250.249391	0.00399	249.309133	0.9363
	13	250.357823	0.11242	251.999063	1.7537
	14	250.586571	0.34117	252.780009	2.5346
256	9	249.858433	0.38697	247.418360	2.8270
	10	249.991341	0.25406	248.248562	1.9968
	11	250.197841	0.04756	249.470953	0.7744
	12	250.239655	0.00574	250.770441	0.5250
	13	250.357573	0.11217	251.711421	1.4660
	14	250.472703	0.22730	251.797723	1.5523
512	9	249.592272	0.65313	248.042800	2.2026
	10	249.896185	0.34921	249.080394	1.1650
	11	250.089218	0.15618	249.591525	0.6539
	12	250.236033	0.00937	250.669970	0.4246
	13	250.369509	0.12411	250.978674	0.7333
	14	250.534697	0.28930	251.788278	1.5429
1024	9	249.959642	0.28576	248.998184	1.2472
	10	250.059089	0.18631	249.131528	1.1139
	11	250.127756	0.11764	249.672942	0.5725
	12	250.250718	0.00532	250.090116	0.1553
	13	250.387049	0.14165	250.829267	0.5839
	14	250.481235	0.23584	251.662872	1.4175
2048	9	249.972251	0.27315	248.999883	1.2455
	10	250.055927	0.18947	249.408467	0.8369
	11	250.166549	0.07885	249.987224	0.2582
	12	250.249986	0.00459	250.172833	0.0726
	13	250.321274	0.07587	250.581459	0.3361
	14	250.439084	0.19368	251.216429	0.9710

*Маркираните редове представляват оптималните k стойности за всяко N .

$PQMC_{(S)}$ постига грешка 2.8×10^{-4} , докато при $PQMC_{(S)}^{(n)}$ грешката е 0.3172 (виж Таблица 1.8). Аналогична тенденция се наблюдава и при сравнение на грешките от $PQMC_{(H)}$ и $PQMC_{(H)}^{(n)}$ (виж Таблица 1.7). Това подобрение отново се обяснява с конструирането на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$ при почти оптималния СКМК алгоритъм, чрез избора на матрицата на преходната плътност на веригата на Марков.

Фигура 1.5 сравнява числените резултати от класическите СКМК алгоритми и техните почти оптимални варианти при оптимална стойност $k = 12$. За всички реализации N , алгоритмите $PQMC_{(S)}$ и $PQMC_{(H)}$ демонстрират по-добра точност от класическите им аналози като следват почти точно λ_{max} . За разлика от тях, класическите им варианти,

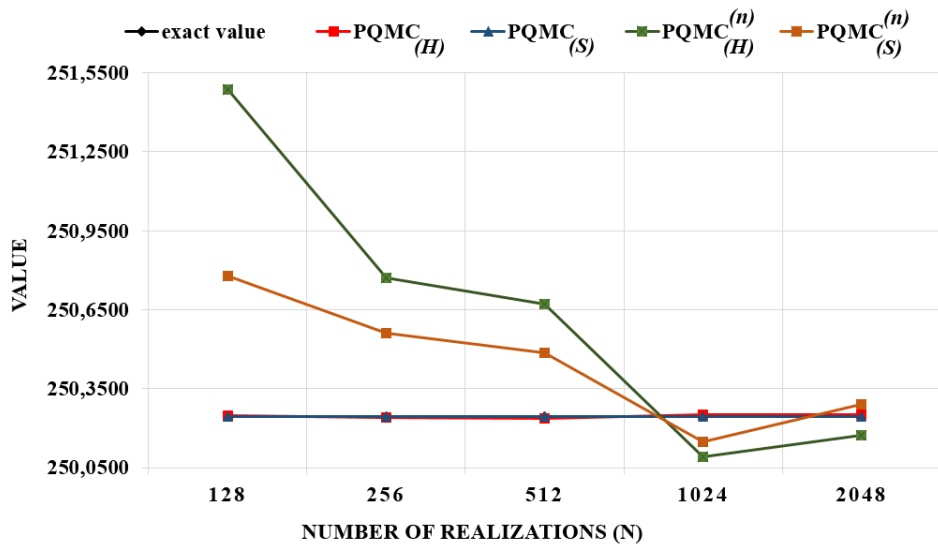
Таблица 1.8: Резултати от $PQMC_{(S)}$ и $PQMC_{(S)}^{(n)}$ при различни N и k за матрица A с размерност $n = 500$. Абсолютната грешка е за случая $\lambda_2/\lambda_1 \approx 0.0363$.

N	k	$\lambda_1^{(k)}$ using	Absolute Error	$\lambda_1^{(k)}$ using	Absolute Error
		$PQMC_{(S)}$	$PQMC_{(S)}$	$PQMC_{(S)}^{(n)}$	$PQMC_{(S)}^{(n)}$
128	9	249.987838	0.25756	248.764573	1.4808
	10	250.099240	0.14616	248.990612	1.2548
	11	250.236415	0.00898	249.415659	0.8297
	12	250.245037	0.00036	250.777310	0.5319
	13	250.330000	0.08460	250.989461	0.7441
	14	250.407154	0.16175	251.062045	0.8166
256	9	249.744662	0.50074	248.778926	1.4665
	10	249.982894	0.26251	249.094364	1.1510
	11	250.199189	0.04621	249.483392	0.7620
	12	250.245119	0.00028	250.562622	0.3172
	13	250.291367	0.04597	250.704555	0.4592
	14	250.369617	0.12422	251.039904	0.7945
512	9	249.638588	0.60681	248.786901	1.4585
	10	249.899316	0.34608	249.301034	0.9444
	11	250.099484	0.14592	249.632500	0.6129
	12	250.244872	0.00053	250.486175	0.2408
	13	250.378047	0.13265	250.679597	0.4342
	14	250.475610	0.23021	250.935292	0.6899
1024	9	249.879108	0.36629	249.156996	1.0884
	10	250.077200	0.16820	249.326758	0.9186
	11	250.140000	0.10540	249.811533	0.4339
	12	250.244996	0.00040	250.148126	0.0973
	13	250.351139	0.10574	250.651883	0.4065
	14	250.440000	0.19460	250.854128	0.6087
2048	9	250.062966	0.18243	249.216417	1.0290
	10	250.147824	0.09758	249.396625	0.8488
	11	250.193656	0.05174	249.937040	0.3084
	12	250.245937	0.00054	250.290864	0.0455
	13	250.272562	0.02716	250.555662	0.3103
	14	250.487070	0.24167	250.713019	0.4676

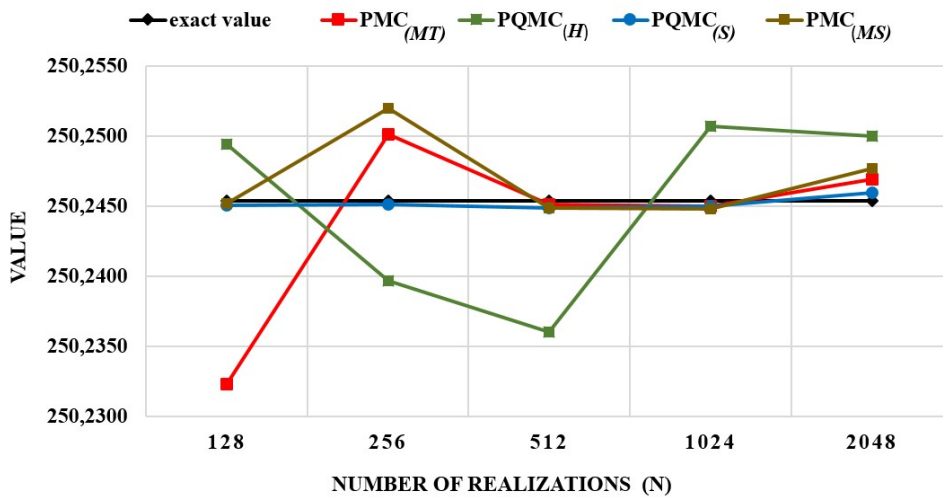
*Маркираните редове представляват оптималните k стойности за всяко N .

$PQMC_{(S)}^{(n)}$ и $PQMC_{(H)}^{(n)}$, се приближават към максималната собствена стойност по-бавно с нарастването на N .

Фигура 1.6 сравнява *почти оптималните* СМК алгоритми, използващи генераторите МТ и MS, с *почти оптималните* СКМК алгоритми, базирани на разбърканите редици на Холтън и Собол. След прецизиране на мащаба по оста Y се вижда, че всички приближени стойности попадат в интервал с ширина $\delta < 0.02$ около точната стойност λ_{max} , което позволява ясно разграничаване на решенията, получени от различните алгоритми. Алгоритъмът $PMC_{(MT)}$ демонстрира добро приближение, като минималната грешка достига 3.8×10^{-4} при $N = 1024$ и $k = 9$ (виж Таблица 1.5). Алгоритъмът $PMC_{(MS)}$ (виж Табли-



Фигура 1.5: Графично сравняване на λ_{max} за матрицата A ($n = 500$) с приближените резултати за $\lambda_{max}^{(k)}$, получени при балансирания случай на грешките, т.е. когато ($k = 12, N \geq 128$). Приложени алгоритми: $PQMC_{(S)}$, $PQMC_{(H)}$, $PQMC_{(S)}^{(n)}$ и $PQMC_{(H)}^{(n)}$.



Фигура 1.6: Графично сравнение на λ_{max} на симетричната матрица A ($n = 500$) с $\lambda_{max}^{(k)}$ в балансиращите случаи. Използвани са *почти оптималните* СКМ и СКМК алгоритми.

ца 1.6) допуска по-големи грешки за малки N , но остава в допустимите граници като с нарастването на N грешката достига 6.5×10^{-4} при $N = 1024$ и $k = 9$.

От двата *почти оптимални* СКМК алгоритми, $PQMC_{(S)}$ отново се откроява като по-прецизен в сравнение с $PQMC_{(H)}$. Той демонстрира стабилна сходимост и минимално отклонение от точната стойност както е представено визуално (виж Фигура 1.6), така и числено в Таблица 1.7 и Таблица 1.8. Алгоритъмът $PQMC_{(H)}$ също подобрява точността си с увеличаване на N , но изостава в сходимостта поради по-слабата равномерност на редицата на Холтън в по-големи размерности. По отношение на оптималния баланс между стохастичната и систематичната грешка, експериментите с плътна симетрична матрица с размерност $n = 500$ показват, че той се постига при $k = 9$ за $PMC_{(MT)}$ и при $k = 12$ за $PQMC_{(S)}$.

Резултатите за случай (ii) демонстрират, че *почти оптималните* СМК алгоритми осигуряват висока точност при умерени изчислителни разходи, а *почти оптималните* СКМК алгоритми, базирани на разбъркани редици на Собол, постигат значително по-добра точност от тези, използващи разбъркани редици на Холтън, при същата изчислителна сложност.

Таблица 1.9: Емпирична оценка на дисперсиите получена чрез *почти оптималния* и *класическия* СМК алгоритъм с генератори *MT* и *MS* при оценяване на максималната собствена стойност.

n	N	k	$Var(PMC_{(MT)})$	$Var(PMC_{(MT)}^{(n)})$	$Var(PMC_{(MS)})$	$Var(PMC_{(MS)}^{(n)})$
100	512	8	1.169203e-06	1.720495e-04	1.236100e-06	1.781005e-04
	2048	8	1.300635e-06	1.772653e-04	1.326996e-06	1.792424e-04
500	512	9	1.186151e-08	1.229778e-05	1.269925e-08	1.285468e-05
	2048	9	1.189763e-08	1.184918e-05	1.268721e-08	1.299595e-05

Като допълнение в нашите изследвания, в Таблица 1.9 са сравнени емпиричните оценки за дисперсиите на случайната величина $\theta^{(k)}$ при *класическите* и *почти оптималните* СМК алгоритми. Оценките са извършени за матрици с размерности $n = 100$ и $n = 500$. Резултатите показват, че при *почти оптималните* СМК алгоритми дисперсията е с около два порядъка по-малка в сравнение с *класическите* им варианти. Тази оценка не зависи съществено от използвания генератор или от броя на реализациите N , което съответства с теоретичните резултати. Фактът, че дисперсията при *почти оптималните* версии е с два порядъка по-малка от *класическите* версии, потвърждава отново направените изводи, че тези алгоритми осигуряват по-точно оценяване на максималната собствена стойност на симетрични матрици при намалена изчислителна сложност.

1.5 Обобщение

В настоящата глава е прецизирана конструкцията на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$ с използване на *почти оптимални* вероятности във веригата на Марков. Доказано е, че за определени класове симетрични матрици случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$ могат да се опишат по-опростено при оценяването на максималната собствена стойност. Разработени са *почти оптималните* СМК и СКМК алгоритми и са описани чрез псевдокод. Изследвана е ефективността на тези алгоритми като са сравнени с *класическите* им аналози. Числените експерименти са проведени за два случая на симетрична плътна матрица A с размерност $n = 100$ и $n = 500$. Получените резултати показват, че *почти оптималните* алгоритми съществено превъзхождат своите *класически* аналози както по точност, така и по изчислителна сложност. Балансът между систематичната и стохастичната грешка зависи от преходите k във веригата на Марков и от реализациите N на случайните величини.

При СМК алгоритмите се установява, че оптималният баланс между грешките се постига при $k = 8$ и $N = 512$, когато размерността на матрицата е $n = 100$. За случая на матрица с размерност $n = 500$, оптималният баланс е при $k = 9$. Използването на

генератора *Mersenne Twister* (MT) при СМК алгоритмите осигурява по-висока точност в сравнение с генератора *Middle Square* (MS). Абсолютните грешки при *почти оптималните* СМК алгоритми са от порядъка на 10^{-4} , докато при *класическите* варианти са от порядъка на 10^{-2} , което ясно подчертава ефективността на **Алгоритъм 1а**.

За СКМК алгоритмите оптималната стойност на броя преходи във веригата на Марков е $k = 11$ за $n = 100$ и $k = 12$ за $n = 500$. Вариантът на **Алгоритъм 1**, базиран на разбъркана редица на Соболев ($PQMC_{(S)}$), демонстрира по-голяма точност и по-гладка сходимост в сравнение с ($PQMC_{(H)}$), конструиран чрез разбъркана редица на Холтън. Абсолютните грешки при *почти оптималните* СКМК алгоритми са от порядъка на 10^{-4} , докато при *класическите алгоритми* те са от порядък 10^{-1} , при същите изчислителни разходи.

В заключение, *почти оптималните* СМК и СКМК алгоритми осигуряват по-висока точност при същата изчислителна сложност, което ги прави предпочитан избор за практически приложения. При тяхното конструиране трябва да се съобразяваме и с избора на генераторите на псевдослучайни числа, и от избора на редиците с малък дискрепанс. Описаните резултати са публикувани в [42, 43].

Глава 2

Стохастичен Резолвентен метод: разработване и изследване на ефективни алгоритми за оценяване на минимална собствена стойност

2.1 Постановка на задачата

Нека $A = \{a_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}$ е дадена несингулярна симетрична матрица. Собствените стойности на A са всички реални стойности на параметъра $\lambda(A)$, за които уравнението:

$$A\mathbf{x} = \lambda(A)\mathbf{x}, \text{ където } \mathbf{x} \in \mathbb{R}^n, \quad (2.1)$$

има ненулево решение. Нека предположим, че собствените стойности на A удовлетворяват условието:

$$0 < \lambda_{min} = \lambda_n < \lambda_{n-1} \leq \lambda_{n-2} \leq \dots \leq \lambda_2 < \lambda_1 = \lambda_{max}. \quad (2.2)$$

Разглеждаме задачата за разработване и изследване на числено ефективни стохастични алгоритми за оценяване на минималната собствена стойност λ_{min} , която съгласно (2.2) е единствена.

2.2 Методи за изчисляване на минималната собствена стойност

Да разгледаме несингулярната симетрична матрица и $R = \{r_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}$, за която собствените стойности изпълняват условието:

$$\mu_{min} = |\mu_n| < |\mu_{n-1}| \leq |\mu_{n-2}| \leq \dots \leq |\mu_2| < |\mu_1| = \mu_{max} \quad (2.3)$$

и за тях уравнението:

$$R\mathbf{y} = \mu(R)\mathbf{y} \quad (2.4)$$

има ненулево решение.

От теорията на линейната алгебра е известно, че при наличие на определена зависимост между матриците R и A възникват съответни връзки между техните собствени стойности и собствените вектори. По-долу са разгледани някои от тези случаи [20]:

1. **Обратна матрица:** Ако матрицата R е обратна на матрицата A , т.е. $R = A^{-1}$, то между собствените стойности е изпълнена връзката:

$$\mu = \frac{1}{\lambda}.$$

В този случай може да се предложи *Обратният Степенен метод* (англ. *Inverse Power method*) за намиране на максималната собствена стойност на R , съответно минималната собствена стойност на A . Недостиг на този подход е високата изчислителна цена, тъй като пресмятането на A^{-1} е трудоемко.

2. **Отместена матрица:** Нека матрицата R е зададена чрез параметър q по следния начин:

$$R_q = A - q\mathbf{I}.$$

Тогава собствените стойности удовлетворяват равенството $\mu = \lambda - q$, което води до *Отместения Степенен метод* (англ. *Shifted Power method*). Този метод е подходящ за изследване при какви условия на параметър q може да се оцени минималната собствена стойност на A .

3. **Резолвентна матрица;** При този случай R_q представлява резолвентна матрица, т.е.:

$$R_q = (\mathbf{I} - qA)^{-1}.$$

Между собствените стойности на R_q и A е валидна връзката:

$$\mu = \frac{1}{1 - q\lambda},$$

а съответният метод се нарича *Резолвентен Степенен метод* (англ. *Resolvent Power method*). Матрицата R_q може да се представи като граница от безкраен сходящ ред при условие $|q\lambda| < 1$ или еквивалентно $|q| \|A\| < 1$, а именно:

$$R_q = \sum_{i=0}^{\infty} q^i A^i = [\mathbf{I} - qA]^{-1}. \quad (2.5)$$

Следва да отбележим, че ако $q < 0$, тогава е в сила зависимостта:

$$\mu_{max} = \frac{1}{1 - q\lambda_{min}} \quad (2.6)$$

и ако $q > 0$, то

$$\mu_{max} = \frac{1}{1 - q\lambda_{max}}. \quad (2.7)$$

За целите на настоящето изследване от основен интерес е случаят, когато ускоряващият параметър удовлетворява условието $q < 0$.

2.2.1 Резолвентен степенен метод

При описанието на Степенния метод в Глава 1 използвахме, че най-голямата собствена стойност на A се представя като граница на коефициента на Рейли:

$$\lambda_{max} = \lambda_1 = \lim_{k \rightarrow \infty} \frac{(\mathbf{h}, A^k \mathbf{f})}{(\mathbf{h}, A^{k-1} \mathbf{f})}. \quad (2.8)$$

В този случай за приближено решение на λ_1 имаме:

$$\lambda_1 \approx \frac{(\mathbf{h}, A^k \mathbf{f})}{(\mathbf{h}, A^{k-1} \mathbf{f})}. \quad (2.9)$$

За да опишем Резолвентния Степенен метод, нека предположим, че $(\mathbf{I} - qA)$ е обратима, т.е. съществува резолвентна матрица $R_q = (\mathbf{I} - qA)^{-1}$, която се явява граница на Ноймановия ред:

$$R_q = (\mathbf{I} - qA)^{-1} = \mathbf{I} + qA + q^2 A^2 + \dots \quad (2.10)$$

Безкрайният матричен ред (2.10) е сходящ при условие $|q| \|A\| < 1$. Ще отбележим, че $R_q = (\mathbf{I} - qA)^{-1}$ е резолвентна матрица за уравнението:

$$\mathbf{x} = qA\mathbf{x} + \mathbf{f}, \quad (2.11)$$

където $\mathbf{x} \in \mathbb{R}^n$ е търсеното решение, $\mathbf{f} \in \mathbb{R}^n$ е произволен вектор и A е разглежданата симетрична матрица. Тогава решението е:

$$\mathbf{x} = (\mathbf{I} - qA)^{-1} \mathbf{f}, \quad (2.12)$$

при условие, че съществува $\det(\mathbf{I} - qA) \neq 0$, което е изпълнено за $q \neq 0$ и $\frac{1}{q}$ да не е собствена стойност на A .

Нека $m \in \mathbb{N}$ е цяло положително число и разглеждаме резолвентната матрица от степен m , т.е. $R_q^m = (\mathbf{I} - qA)^{-m}$. Тя има стандартно развитие по обобщен биномен ред (матричен аналог на $(1 - u)^{-m}$) [24, 25]:

$$R_q^m = (\mathbf{I} - qA)^{-m} = \sum_{i=0}^{\infty} q^i C_{i+m-1}^i A^i, \quad (2.13)$$

където биномните коефициенти се записват по следния начин:

$$C_{i+m-1}^i = \frac{(m+i-1)!}{(m-1)!i!}, \quad i = 0, 1, 2, \dots \quad (2.14)$$

Безкрайният ред за R_q^m (q се нарича параметър за ускорение) може да се запише като:

$$R_q^m = (I - qA)^{-m} = \mathbf{I} + mqA + \frac{m(m+1)}{2}q^2A^2 + \frac{m(m+1)(m+2)}{6}q^3A^3 + \dots \quad (2.15)$$

Прилагайки Степенния метод (2.8) за резолвентната матрица R_q , получаваме Резолвентния Степенен метод [24, 25], т.е.:

$$\lim_{m \rightarrow \infty} \frac{(\mathbf{h}, R_q^m \mathbf{f})}{(\mathbf{h}, R_q^{m-1} \mathbf{f})} = \lim_{m \rightarrow \infty} \frac{(\mathbf{h}, (\mathbf{I} - qA)^{-m} \mathbf{f})}{(\mathbf{h}, (\mathbf{I} - qA)^{-(m-1)} \mathbf{f})} = \mu_{max} = \frac{1}{1 - q\lambda} \quad (2.16)$$

Систематичната грешка при Резолвентния Степенен метод е от порядъка $\mathcal{O}(|\frac{\mu_2}{\mu_1}|^m)$ (виж [24, 39]). В зависимост от знака на ускоряващия параметър q , съгласно формули (2.6) и (2.7), Резолвентният Степенен метод може да бъде използван за приближено пресмятане както на минималната, така и на максималната собствена стойност на симетричната матрица A .

В [24, 25, 27] е доказано, че екстремалните собствени стойности λ_{min} и λ_{max} могат да бъдат пресметнати приближено чрез Резолвентния Степенен метод (2.16) и по следния начин:

$$\lambda = \frac{(\mathbf{h}, AR_q^m \mathbf{f})}{(\mathbf{h}, R_q^m \mathbf{f})} \approx \lambda_{max}, \quad \text{ако } q > 0, \quad (2.17)$$

и

$$\frac{(\mathbf{h}, AR_q^m \mathbf{f})}{(\mathbf{h}, R_q^m \mathbf{f})} \approx \frac{1}{q} \left(1 - \frac{1}{\mu^{(m)}}\right) \approx \lambda_{min}, \quad \text{ако } q < 0, \quad (2.18)$$

където $\mu^{(m)}$ е приближението на максималната собствена стойност μ_{max} на резолвентната матрица $R_q = [I - qA]^{-1} \in \mathbb{R}^{n \times n}$.

В практиката обикновено се използва приближена стойност за резолвентната матрица като безкрайният ред (2.13) се прекъсва при $i = k$, т.е.:

$$R_{q,k}^m = \sum_{i=0}^k C_{i+m-1}^i q^i A^i. \quad (2.19)$$

Тогав систематичната грешка е просто „опашката“ на реда:

$$\Delta_k^{(m)}(q) = (I - qA)^{-m} - R_{q,k}^m = \sum_{i=k+1}^{\infty} C_{i+m-1}^i q^i A^i. \quad (2.20)$$

Възниква задачата да се определят стойностите на параметрите k и m , за които е изпълнено условието:

$$|\Delta_k^{(m)}(q)| < \varepsilon, \quad (2.21)$$

където ε е фиксирано малко число в интервала $(0, 1)$, т.е. $0 < \varepsilon < 1$.

Да разгледаме случая, когато $|q||A| < 1$. При L_2 -норма имаме $\|A\|_2 = \lambda_1$ и получаваме:

$$|q| |\lambda_1| < 1.$$

В сила е следната оценка:

$$|\Delta_k^{(m)}(t)| \leq T_{m,k}(t) = \sum_{i=k+1}^{\infty} C_{i+m-1}^i t^i, \quad t = |q||\lambda_1| < 1. \quad (2.22)$$

(1) Нека $m = 1$. Тогава всички биномни коефициенти са равни на 1 и

$$T_{1,k}(t) = \sum_{i=k+1}^{\infty} t^i = \frac{t^{k+1}}{1-t} < \varepsilon. \quad (2.23)$$

От (2.23) получаваме:

$$t^{k+1} < \varepsilon(1-t) \iff k+1 > \frac{\ln(\varepsilon(1-t))}{\ln t}. \quad (2.24)$$

Да отбележим, че $\ln t < 0$. За минималната стойност на k получаваме:

$$k_{\min} + 1 = \left\lceil \frac{\ln(\varepsilon(1-t))}{\ln t} \right\rceil, \quad \text{където } t = |q||\lambda_1|. \quad (2.25)$$

(2) Нека $m \geq 1$. Означаваме

$$b_i = \binom{m+i-1}{i} t^i = C_{i+m-1}^i t^i.$$

Тогава отношението на два последователни члена от реда (2.13) е:

$$\frac{b_{i+1}}{b_i} = \frac{m+i}{i+1} t.$$

Ако за избраното $i = k+1$ е изпълнено условието:

$$\alpha := \frac{m+k+1}{k+2} t < 1,$$

то опашката (2.20) на реда (2.13) може да се оцени като геометрична прогресия:

$$T_{m,k}(t) \leq \frac{b_{k+1}}{1-\alpha} = \frac{\binom{m+k}{k+1} t^{k+1}}{1 - \frac{m+k+1}{k+2} t}. \quad (2.26)$$

Следователно, достатъчно е да изберем k така, че да е изпълнено неравенството:

$$\frac{\binom{m+k}{k+1} t^{k+1}}{1 - \frac{m+k+1}{k+2} t} < \varepsilon. \quad (2.27)$$

Условието $0 < \alpha < 1$ е еквивалентно на:

$$t(m+k+1) < k+2 \iff k+1 > \frac{tm-1}{1-t}. \quad (2.28)$$

От (2.28) се получава следното неравенство за m (като функция на k и t):

$$m < \frac{(k+1)(1-t)+1}{t} = \frac{k+2}{t} - (k+1) = \frac{1-t}{t}(k+1) + \frac{1}{t}. \quad (2.29)$$

Максималното цяло положително число m при фиксирано t се определя от формулата:

$$m_{max} = \left\lfloor \frac{(1-t)k+2}{t} - 1 \right\rfloor. \quad (2.30)$$

Също така от (2.28) се извежда неравенство за k (като функция на m и t):

$$k > \frac{t}{1-t}m - \frac{2-t}{1-t}, \quad (2.31)$$

като получаваме асимптотика (ако m расте, а t е фиксирано):

$$k \gtrsim \frac{t}{1-t}m, \quad \text{за големи стойности на } m. \quad (2.32)$$

Минималното цяло положително число k при фиксирано t се определя от формулата:

$$k_{min} = \left\lceil \frac{tm-1}{1-t} - 1 \right\rceil. \quad (2.33)$$

Ще отбележим, че връзката между параметрите m и k е получена при условие $t = |q| |\lambda_{max}| < 1$. Формулите (2.29) и (2.31) показват, че връзката между m и k е линейна.

На основата на Резолвентния Степенен метод в следващия параграф разглеждаме Стохастичния Резолвентен метод. Получената зависимост между параметрите m и k ни дава възможност в практически задачи да използваме $R_{q,k}^m$ за оценяване на минималната собствена стойност λ_{min} на симетричната матрица A чрез Стохастичния Резолвентен метод, както и да изследваме баланса между стохастичната и систематичната грешка при конструирането на съответните стохастични резолвентни алгоритми.

2.2.2 Стохастичен Резолвентен метод

Стохастичният Резолвентен метод е представен за първи път от Караиванова и Димов в [25, 30] за оценяване на екстремалните собствени стойности на матрицата A чрез формулите (2.17) и (2.18). Съществен елемент на метода е конструкцията на случайната величина $\theta^{(k)}$, въведена в Глава 1 чрез формула (1.8). В този контекст, Теорема 1.2.1 от Глава 1 се обобщава чрез разглеждане на безкрайна редица от случайни величини $\theta^{(i)}$, $i = 1, 2, 3, \dots$, дефинирани по формула (1.8), чиито математически очаквания са равни на скаларните произведения $(\mathbf{h}, A^i \mathbf{f})$, $i = 1, 2, 3, \dots$, т.е.:

$$E[\theta^{(i)}] = (\mathbf{h}, A^i \mathbf{f}), \quad i = 1, 2, 3, \dots \quad (2.34)$$

В сила е следната теорема:

Теорема 2.2.1. Нека случайните величини $\theta^{(i)}$, $i = 1, 2, 3, \dots$, са дефинирани чрез формула (1.8) от Глава 1. Нека параметърът q е избран така, че:

$$|q \lambda_{max}| < 1,$$

където λ_{max} е най-голямата собствена стойност на симетричната матрица A . Тогава за скаларното произведение $(\mathbf{h}, R_q^m \mathbf{f})$ е в сила равенството:

$$(\mathbf{h}, R_q^m \mathbf{f}) = (\mathbf{h}, (\mathbf{I} - qA)^{-m} \mathbf{f}) = \sum_{i=0}^{\infty} q^i C_{i+m-1}^i E[\theta^{(i)}]. \quad (2.35)$$

Доказателството на Теорема 2.2.1 може да бъде намерено в [22], като съществено се използва оценката (2.34).

Разглеждаме средните стойности на редицата от случайните величини $\theta^{(i)}$, $i = 1, 2, 3, \dots$, дефинирани чрез:

$$\bar{\theta}_N^{(i)} = \frac{1}{N} \sum_{s=1}^N (\theta^{(i)})_s, \quad i = 1, 2, 3, \dots \quad (2.36)$$

В случая на МК апроксимации на скаларните произведения $(\mathbf{h}, A^i \mathbf{f})$, стохастичните (вероятностни) грешки при фиксирана вероятност $\mathbb{P} < 1$ (например, при правилото на „трите сигми“ тя е равна на 0.997) се оценяват с неравенствата:

$$r_N^{(i)} = |(\mathbf{h}, A^i \mathbf{f}) - \bar{\theta}_N^{(i)}| < c_i \sigma(\theta^{(i)}) N^{-1/2}, \quad i = 1, 2, 3, \dots \quad (2.37)$$

При правилото на „трите сигми“ константите са $c_i = 3$.

Възниква задачата при фиксирано $\varepsilon \in (0, 1)$, да се определи минималния брой реализации N на случайните величини $\theta^{(i)}$, $i = 1, 2, 3, \dots$, така че да е изпълнено условието:

$$r_N^{(i)} < \varepsilon, \quad i = 1, 2, 3, \dots \quad (2.38)$$

Тази задача е решена в [24] при условие, че съществува константа σ , за която

$$\sigma \geq \sigma(\theta^{(i)}), \quad i = 1, 2, 3, \dots \quad (2.39)$$

Достатъчно условие за съществуването на константата σ е $\|A\| < 1$. Освен това, матрицата на преходната плътност $P = \{p_{ij}\}_{i,j=1}^n$ и началният вектор на плътността $\mathbf{p} = \{p_i\}_{i=1}^n$ трябва да бъдат допустими съответно за симетричната матрица A и за вектора \mathbf{h} .

За необходимия брой реализации N е получена оценката:

$$N \geq \left\{ \frac{c_k \sigma}{\varepsilon} \right\}^2. \quad (2.40)$$

На практиката скаларното произведение $(\mathbf{h}, R_q^m \mathbf{f})$ от равенството (2.35) се оценява чрез

използване на краен брой случайни величини $\theta^{(i)}$, $i = 1, \dots, k$, т.е.:

$$(\mathbf{h}, R_q^m \mathbf{f}) \approx \sum_{i=0}^k q^i C_{i+m-1}^i E[\theta^{(i)}]. \quad (2.41)$$

По този начин при оценяването на скаларното произведение $(\mathbf{h}, R_q^m \mathbf{f})$ възникват два вида грешки: стохастична и систематична. Балансът между двете грешки, така че и двете да са от един и същи порядък, оказва съществено влияние както върху изчислителната цена, така и върху точността. Това е причината за условието грешките да са по-малки или равни на фиксирано число ε в интервала $(0, 1)$. Равенствата (2.30) и (2.33) определят стойностите на параметрите m и k , за които систематичната грешка е по-малка от ε . Чрез формула (2.40) се контролира броят реализации N така, че стохастичната грешка да бъде от същия порядък. Параметърът c_k в (2.40) зависи от броя на използваните случайни величини $\theta^{(i)}$, $i = 1, \dots, k$, които се конструират чрез веригата на Марков при оценяването на скаларното произведение в (2.41).

Изразът (2.43) може да бъде представен и в еквивалентния разширен вид:

$$\lambda \approx \frac{E[\theta^{(1)}] + qC_m^1 E[\theta^{(2)}] + q^2 C_{m+1}^2 E[\theta^{(3)}] + \dots + q^k C_{k+m-1}^k E[\theta^{(k+1)}]}{1 + qC_m^1 E[\theta^{(1)}] + q^2 C_{m+1}^2 E[\theta^{(2)}] + \dots + q^k C_{k+m-1}^k E[\theta^{(k)}}]. \quad (2.42)$$

В зависимост от знака на ускоряващият параметър q се реализират две възможности: при $q > 0$ се получава приближена стойност на най-голямата собствена стойност $\lambda_1 = \lambda_{\max}$, а при $q < 0$ се оценява най-малката собствена стойност $\lambda_n = \lambda_{\min}$ на матрицата A .

Когато стойностите на случайните величини $\theta^{(i)}$, $i = 1, \dots, k$ в (2.44) се симулират с помощта на псевдослучайни генератори, Стохастичният Резолвентен метод се определя като Резолвентен Монте Карло (РМК) метод. При използване на редици с малък дискрепанс (например, редица на Собол), се получава Резолвентен Квази-Монте Карло (РКМК) метод.

Прилагайки Теорема 2.2.1 за оценяване на скаларните произведения в изразите (2.17) и (2.18), чрез използване на краен брой случайни величини $\theta^{(i)}$, $i = 1, \dots, k$, получаваме приближено пресмятане на екстремалните собствени стойности на A чрез Стохастичния Резолвентен метод, а именно:

$$\lambda \approx \frac{\sum_{i=0}^k q^i C_{i+m-1}^i E[\theta^{(i+1)}]}{\sum_{i=0}^k q^i C_{i+m-1}^i E[\theta^{(i)}]}, \quad (2.43)$$

където $\theta^{(0)} = 1$, а случайните величини $\theta^{(i)}$, $i = 1, \dots, k$, се дефинират чрез формула (1.8) от Глава 1.

Изразът (2.43) може да бъде представен и в еквивалентния разширен вид:

$$\lambda \approx \frac{E[\theta^{(1)}] + qC_m^1 E[\theta^{(2)}] + q^2 C_{m+1}^2 E[\theta^{(3)}] + \dots + q^k C_{k+m-1}^k E[\theta^{(k+1)}]}{1 + qC_m^1 E[\theta^{(1)}] + q^2 C_{m+1}^2 E[\theta^{(2)}] + \dots + q^k C_{k+m-1}^k E[\theta^{(k)}}]. \quad (2.44)$$

В зависимост от знака на ускоряващият параметър q се реализират две възможности. При $q > 0$ се получава приближена стойност на най-голямата собствена стойност $\lambda_1 = \lambda_{\max}$,

докато при $q < 0$ се оценява най-малката собствена стойност $\lambda_n = \lambda_{\min}$ на матрицата A . Когато стойностите на случайните величини $\theta^{(i)}$, $i = 1, \dots, k$ в (2.44) се генерират с помощта на псевдослучайни генератори, Стохастичният Резолвентен метод се определя като Резолвентен Монте Карло (РМК) метод. При използване на редици с малък дискрепанс (например, редици на Соболев), се получава Резолвентен Квази-Монте Карло (РКМК) метод.

Ще отбележим, че математическите очаквания $E[\theta^{(i)}]$ се апроксимират чрез извадкови средни:

$$E[\theta^{(i)}] \approx \bar{\theta}_N^{(i)} = \frac{1}{N} \sum_{s=1}^N (\theta^{(i)})_s, \quad i = 1, 2, \dots, k + 1. \quad (2.45)$$

За конструирането на изчислителен алгоритъм математическите очаквания $E[\theta^{(i)}]$, $i = 1, 2, \dots, k + 1$ в изразите (2.43) и (2.44) се заместват със съответните извадкови средни $\bar{\theta}_N^{(i)}$.

Конструкцията на случайните величини $\theta^{(i)}$, $i = 1, 2, \dots, k + 1$ е въведена в Глава 1 чрез формула (1.8) и се основава на използване на верига на Марков. При изграждане на тази верига от съществено значение е изборът на началния вектор на плътността \mathbf{p} и матрицата на преходните вероятности P . Числените резултати, представени в Глава 1, показваха, че изследваните *почти оптимални* Степенни стохастични алгоритми водят до по-добра точност в сравнение с *класическите* им аналози. В следващия параграф при конструирането на РМК и РКМК алгоритми ще бъде използван същият избор на началния вектор на плътността \mathbf{p} и матрицата на преходните вероятности P (виж формула (1.16) в Глава 1).

2.3 Почти оптимален Резолвентен Монте Карло и Резолвентен Квази-Монте Карло алгоритми

Нека отново представим симетричната матрица A по следния начин:

$$A = \{a_{ij}\}_{i,j=1}^n = (\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_n)^T, \quad \text{където } \mathbf{a}_i = (a_{i1}, \dots, a_{in}), \quad i = 1, \dots, n.$$

Да въведем следните векторни и матрични норми:

$$\|\mathbf{h}\| = \|\mathbf{h}\|_1 = \sum_{i=1}^n |h_i|, \quad \|\mathbf{a}_i\| = \|\mathbf{a}_i\|_1 = \sum_{j=1}^n |a_{ij}|, \quad \|A\| = \|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|, \quad j = 1, \dots, n.$$

За конструирането на резолвентен стохастичен алгоритъм отново използваме *почти оптимален* начален вектор на плътността и *почти оптимална* матрица на преходните вероятности, а именно:

$$\mathbf{p} = \{p_i\}_{i=1}^n, \quad p_i = \frac{|h_i|}{\|\mathbf{h}\|} \quad \text{и} \quad P = \{p_{ij}\}_{i,j=1}^n, \quad p_{ij} = \frac{|a_{ij}|}{\|\mathbf{a}_i\|}, \quad i, j = 1, \dots, n. \quad (2.46)$$

Алгоритъм 2, представен чрез псевдокод, описващ *почти оптималния* Резолвентен Квази-Монте Карло (РКМК) алгоритъм (англ. *almost optimal Resolvent Quasi-Monte Carlo*

(RQMC) algorithm) за оценяване на минималната собствена стойност на симетричната матрица A . В стъпка 4 на алгоритъма се прилага разбъркана редица на Соболев като редица в малък дискрепанс за изчисляване стойностите на случайните величини $\theta^{(i)}$, $i = 1, 2, \dots, k + 1$, чрез верига на Марков. В случай, че в стъпка 4 се използват генератори на псевдослучайни числа (PRNGs), алгоритъмът се определя като *почти оптимален* Резолвентен Монте Карло (РМК) алгоритъм (англ. *almost optimal Resolvent Monte Carlo (RMC) algorithm*). В по-нататъшното изложение *почти оптималният* РМК алгоритъм ще означаваме с **Алгоритъм 2а**.

Algorithm 2 Pseudocode for computing the almost optimal Resolvent QMC algorithm

- 1: INPUT: matrix $A = \{a_{ij}\} \in \mathbb{R}^{n \times n}$; vectors $\mathbf{h}, \mathbf{f} \in \mathbb{R}^n$, positive integers n, N, k, m and the real number q , where $(-1)/\|A\| < q < 0$
- 2: COMPUTE: $p_i = \frac{|h_i|}{\sum_{i=1}^n |h_i|}$; $p_{ij} = \frac{|a_{ij}|}{\sum_{j=1}^n |a_{ij}|}$, $1 \leq i \leq n$, $1 \leq j \leq n$
- 3: COMPUTE: $a_i = \sum_{j=1}^n |a_{ij}|$, $1 \leq i \leq n$, $\|\mathbf{h}\| = \sum_{i=1}^n |h_i|$
- 4: GENERATE: N elements of the $(k + 1)$ - dimensional Sobol sequence
- 5: CONSTRUCT: N realizations of the Markov chain with integer elements

$$l_1^{(s)} \rightarrow \dots \rightarrow l_i^{(s)} \rightarrow \dots \rightarrow l_k^{(s)} \rightarrow l_{k+1}^{(s)}, \quad 1 \leq l_i^{(s)} \leq n, \quad 1 \leq i \leq k + 1, \quad 1 \leq s \leq N$$

- 6: COMPUTE:

$$\bar{\theta}^{(i)} = \frac{1}{N} \|\mathbf{h}\| \sum_{s=1}^N \text{sign}\{h_{l_1^{(s)}} \prod_{i=1}^{k+1} (a_{l_i^{(s)} l_{i+1}^{(s)}})\} \prod_{i=1}^{k+1} a_{l_i^{(s)}} f_{l_{i+1}^{(s)}}; \quad 1 \leq i \leq k + 1$$

- 7: COMPUTE:

$$C_{i+m-1}^i = \frac{(i + m - 1)!}{i!(m - 1)!}, \text{ where } 1 \leq i \leq k$$

- 8: OUTPUT:

$$\lambda_{min} \approx \frac{\bar{\theta}^{(1)} + qC_m^1 \bar{\theta}^{(2)} + q^2 C_{m+1}^2 \bar{\theta}^{(3)} + \dots + q^k C_{k+m-1}^k \bar{\theta}^{(k+1)}}{1 + qC_m^1 \bar{\theta}^{(1)} + q^2 C_{m+1}^2 \bar{\theta}^{(2)} + \dots + q^k C_{k+m-1}^k \bar{\theta}^{(k)}}$$

Алгоритъм 2 и **Алгоритъм 2а** принадлежат към класа на итерационните Монте Карло и Квази-Монте Карло алгоритми. Този клас алгоритми се характеризира с наличието на два вида грешки - стохастична и систематична. Стохастичната грешка произтича от вероятностната природа на МК алгоритмите и е от порядъка $\mathcal{O}(N^{-1/2})$. При КМК алгоритмите, които използват детерминистични редици с малък дискрепанс, за оценка на грешката се използва неравенството на Коксма-Хлавка [14, 92] и грешката е от порядъка $\mathcal{O}(N^{-1}(\log N)^k)$. При използване на редици с малък дискрепанс в пространства с висока размерност могат да възникнат корелации между отделни координати, което води до влошаване на сходимостта. Подобен ефект беше илюстриран с редиците на Соболев в уводната част на дисертационния труд. Въвеждането на рандомизирани квазислучайни редици, които добавят контролирана случайност, запазвайки равномерността, позволява грешката при КМК алгоритмите да бъде разглеждана като стохастична по своя характер.

Систематичната грешка възниква поради факта, че в практическите реализации на итерационните стохастични алгоритми случайните величини се конструират с помощта на крайна (прекъсната) верига на Марков. Използването на крайна верига на Марков е аналогично на прилагането на прекъснат ред на Нойман при детерминистичните итерационни алгоритми и води до наличието на априорна систематична грешка в получените оценки.

Почти оптималните РКМК и РМК алгоритми оценяват скаларните произведения, които участват в числителя и знаменателя на изразите (2.17) и (2.18), като за резолвентна матрица се използва $R_{q,k}^m$, съгласно формула (2.19). По този начин систематичната грешка зависи от параметъра на прекъсване k в представянето на резолвентната матрица като граница на безкраен ред с условие $|q|||A|| < 1$. В безкрайния ред (2.15), q е параметър на ускорението, а m е степента на резолвентната матрица. Ще отбележим, че параметърът m влияе върху сходимостта на Резолвентния Степенен метод, като сходимостта е от порядъка:

$$\mathcal{O}\left(\left|\frac{\mu_2}{\mu_1}\right|^m\right), \quad (2.47)$$

където μ_2 и μ_1 са двете най-големи собствени стойности на резолвентната матрица R_q .

Като се вземат предвид формулите (2.6) и (2.7), порядъкът за сходимостта (2.47) се изразява чрез собствените стойности на A и ускоряващия параметър q :

$$\mathcal{O}\left(\left|\frac{1 - q\lambda_1}{1 - q\lambda_2}\right|^m\right), \text{ при } q > 1 \text{ и } \mathcal{O}\left(\left|\frac{1 - q\lambda_n}{1 - q\lambda_{n-1}}\right|^m\right), \text{ при } q < 1, \quad (2.48)$$

където λ_1 и λ_2 са най-големите собствени стойности, а λ_n и λ_{n-1} са най-малките собствени стойности на A . За да бъдат приложени описаните *почти оптимални* алгоритми, е необходимо да съществува разделение между двете собствени стойности, които определят порядъка на метода (2.48). В числените експерименти, представени в следващия параграф, са разгледани примери на симетрични матрици, при които е налице разделение между двете най-големи (съответно най-малки) за матрицата собствени стойности.

Псевдокодът на **Алгоритъм 2** и **Алгоритъм 2а** оценява минималната собствена стойност на матрицата A , тъй като в стъпка 1 ускоряващият параметър се избира $q \in (-1/||A||, 0)$. В случай, че изберем $q \in (0, 1/||A||)$, псевдокодът описва *почти оптимален* стохастичен алгоритъм за оценяване на максималната собствена стойност на A . Биномните коефициенти в стъпка 7 могат да бъдат предварително изчислени. В стъпка 6 извадковите средни $\bar{\theta}_N^{(i)}$, $i = 1, 2, \dots, k + 1$ зависят от получените стойности на случайните величини $\theta^{(i)}$, $i = 1, 2, \dots, k + 1$. Стойностите на $\theta^{(i)}$, $i = 1, 2, \dots, k + 1$ се пресмятат чрез N реализации на веригата на Марков в стъпка 5.

Изследването на баланса между стохастичната и систематичната грешка, т.е. двете грешки да бъдат по-малки от предварително зададено малко число $0 < \varepsilon < 1$, е от съществено значение. Този баланс определя както точността при оценяването на екстремални собствени стойности чрез **Алгоритъм 2** и **Алгоритъм 2а**, така и тяхната изчислителна цена.

Балансът между стохастичната и систематична грешка може да се контролира чрез след-

ните параметри:

1. броят на реализациите N на веригата на Марков (големината на извадката на случайните величини $\theta^{(i)}$, $i = 1, 2, \dots, k + 1$);
2. броят на преходите k във веригата на Марков, параметър, който съответства на прекъснатия итерационен процес (2.19) с фиксирани k итерации;
3. степента m на резолвентната матрица;
4. параметърът за ускорение q .

Числените тестове, представени в следващия параграф, показват, че изборът на тези параметри трябва да е в съзвучие с получените теоретични оценки за връзките между тях (виж изрази (2.27), (2.29), (2.31), (2.40) и (2.48)).

2.4 Числени резултати и дискусия

В този параграф са представени числените резултати чрез прилагането на **Алгоритъм 2** и **Алгоритъм 2а** за оценяване на минималната собствена стойност на случайно генерирани симетрични матрици с размерности $n = 500$ и $n = 1000$. *Почти оптималният* РКМК алгоритъм, който сме обозначили като на **Алгоритъм 2**, е реализиран с генератор BRODA's Sobol RSG [12]. Този генератор ни дава възможност да генерираме разбъркани редици на Собол с необходимата размерност при конструирането на веригата на Марков. **Алгоритъм 2а** е реализиран с МТ генератора на псевдослучайни числа [78]. Двата алгоритъма са написани на езика за програмиране C++, като предварително се използва подпрограма, която генерира произволна симетричната матрица A със зададена размерност n , така че всичките ѝ собствени стойности да са в единичната окръжност и да удовлетворяват неравенството:

$$0 < c < \lambda_n < \lambda_{n-1} \leq \lambda_{n-2} \leq \dots \leq \lambda_2 < \lambda_1 < 1.$$

Константата c е избрана да бъде $c = 0.15$.

Тестовите случаи са както следва:

- (i) Матрица A с размерност $n = 500$, като двете най-големи и двете най-малки собствени стойности са: $\lambda_1 = 0.620713$, $\lambda_2 = 0.335683$, $\lambda_{499} = 0.15452$ и $\lambda_{500} = 0.152992$.
- (ii) Матрица A с размерност $n = 1000$, като двете най-големи и двете най-малки собствени стойности са: $\lambda_1 = 0.688181$, $\lambda_2 = 0.443078$, $\lambda_{999} = 0.154138$ и $\lambda_{1000} = 0.151911$.

Двете най-големи и двете най-малки собствени стойности на случайно генерираните матрици A са различни, с което гарантираме сходимост на Резолвентния Степенен метод.

Числените тестове са проведени върху високопроизводителен изчислителен клъстер разположен в изчислителния център на ИИКТ-БАН, състоящ се от 12 броя сървъри Fujitsu Primergy RX 2540 M4, като всеки от тях е със следната конфигурация: 128 GB оперативна памет; два броя процесори - CPU 2x Intel Xeon Gold 5118 2.30 GHz с 24 ядра; 2 броя по 800 GB SSD; 36 TB HDD, Графична карта - NVIDIA Tesla V100 с 32 GB

оперативна памет. Сървърите са свързани с InfiniBand за бърз обмен на данни с малка латентност. Операционната система е Red Hat Enterprise Linux 7.6 и NVIDIA GRID. Ще отбележим, че реализацията на **Алгоритъм 2** е хибридна, т.е. разбърканата редица на Соболев се генерира чрез генератор BRODA's Sobol RSG като се използва графична карта NVIDIA Tesla V100 с 5120 CUDA ядра, а останалите изчисления се извършват с помощта на двата броя процесори - CPU 2x Intel Xeon Gold 5118 2.30 GHz с 24 ядра.

Проведените тестовите показаха, че при големи стойности на N времето за изпълнение на **Алгоритъм 2** е съществено по-малко от времето за изпълнение на **Алгоритъм 2а**. При числените експерименти произволните вектори \mathbf{h} и \mathbf{f} са избрани с L_1 -норма по следния начин:

$$\|\mathbf{h}\| = \|\mathbf{f}\| = \sum_{i=1}^n \frac{1}{n} = 1.$$

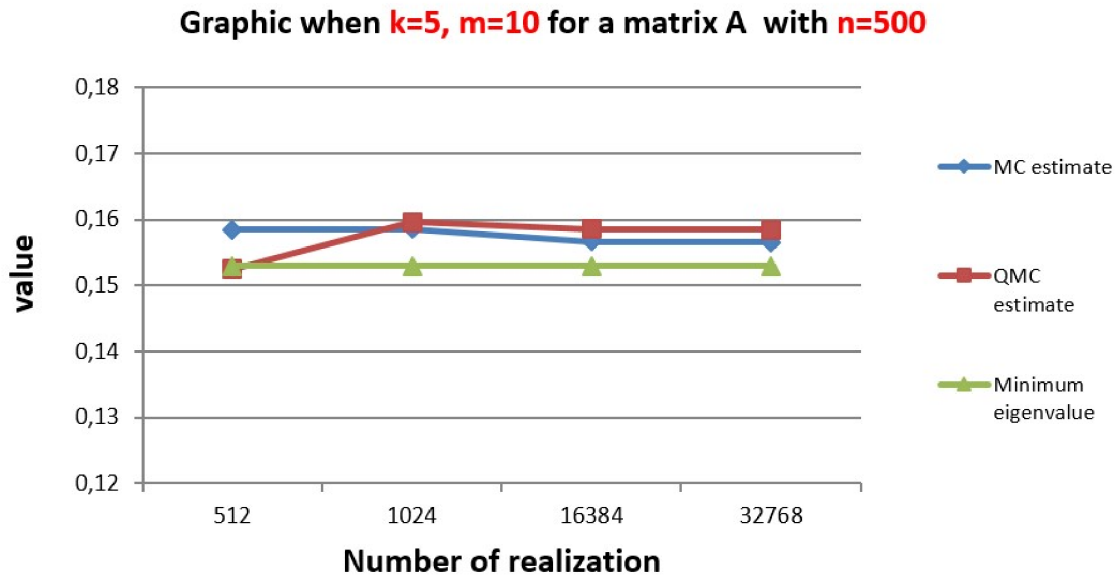
Този избор е направен и за двата случая на случайно генерираната симетрична матрица A с размерности $n = 500$ и $n = 1000$. В Таблица 2.1 и Таблица 2.2, **Алгоритъм 2** е означен като $RQMC_{(S)}$ (англ. *Resolvent Quasi-Monte Carlo using Sobol sequences*), а **Алгоритъм 2а** е записан като $RMC_{(MT)}$ (англ. *Resolvent Monte Carlo using Mersenne Twister generator*).

2.4.1 Случай (i): Размер на матрицата $n=500$

Таблица 2.1: Числени резултати за матрицата $A \in \mathbb{R}^{500 \times 500}$ при различни стойности на N , k и m с параметъра на ускорение $q = -0.157428$. Точната минимална собствена стойност $\lambda_{\min} = \lambda_{500} = 0.152992$.

N	k	m	λ_{\min}	Absolute error	λ_{\min}	Absolute error
			$RMC_{(MT)}$	$RMC_{(MT)}$	$RQMC_{(S)}$	$RQMC_{(S)}$
$512 * 2^0$	5	5	0.1585637	0.0055717	0.1561046	0.0031126
	5	10	0.1585163	0.0055243	0.1524884	0.0005036
	10	5	0.1556132	0.0026212	0.1560595	0.0030675
$512 * 2^1$	5	5	0.158395	0.005403	0.1596503	0.0066583
	5	10	0.1585518	0.0055598	0.1595946	0.0066026
	10	5	0.1580794	0.0050874	0.1596616	0.0066696
$512 * 2^5$	5	5	0.1572588	0.0042668	0.1591139	0.0061219
	5	10	0.1566524	0.0036604	0.1585388	0.0055468
	10	5	0.156875	0.003883	0.1591292	0.0061372
$512 * 2^6$	5	5	0.157129	0.004137	0.1589384	0.0059464
	5	10	0.1565796	0.0035876	0.1584563	0.0054643
	10	5	0.1570216	0.0040296	0.158961	0.005969

В Таблица 2.1 са представени резултати от оценяването на най-малката собствена стойност на симетричната матрица с размерност $n = 500$, получени чрез *почти оптималните* алгоритми $RMC_{(MT)}$ и $RQMC_{(S)}$ при различни стойности на параметрите N , k и m . В таблицата са включени съответните абсолютни грешки на получените оценки. В проведените числени експерименти параметърът на ускорение е фиксиран на стойност $q = -0.157428$, а систематичната грешка е предварително зададена на ниво $\varepsilon_{sys} = 0.001$.



Фигура 2.1: Сходимост към точната минимална собствена стойност на матрицата A при използване на **Алгоритъм 2** и **Алгоритъм 2а** с параметри $k = 5, m = 10$ за различни на брой N

Анализът на резултатите се основава на балансирането на двата основни източника на неточност: систематичната и стохастичната грешка. Систематичната грешка произтича от прекъсването на биномния (Ноймановия) ред при $i = k + 1$, което е еквивалентно на разглеждането на краен брой преходи във веригата на Марков, както и от избора на степента m на резолвентната матрица. От своя страна, стохастичната грешка възниква вследствие на извадковото оценяване на математическите очаквания чрез краен брой реализации N . В числените експерименти параметрите m и k не се третират като свободни оптимизационни променливи, а се определят въз основа на теоретичните оценки за „опашката“ на реда $T_{m,k}(t)$ при $t = |q||\lambda_1| < 1$. По-конкретно, те се избират така, че да удовлетворяват условието $T_{m,k}(t) \leq \varepsilon_{sys} = 0.001$ (виж теоретичните оценки (2.27), (2.29)). Следователно, стойностите на m и k , използвани в Таблица 2.1, представляват обоснован избор, който гарантира, че систематичната грешка остава под предварително зададения праг 10^{-3} , при $t = |q||\lambda_1| < 1$ за разглежданата матрица. При така фиксираната систематична грешка от порядък $\mathcal{O}(10^{-3})$ чрез подходящ избор на двойките (m, k) , основният параметър, определящ общата точност на оценката, остава броят реализации N , който контролира стохастичната компонента на грешката. При Монте Карло подхода, реализиран в **Алгоритъм 2а**, стохастичната грешка е от порядък $\mathcal{O}(N^{-1/2})$ при фиксирана доверителна вероятност (например, съгласно правилото на трите сигми). При Квази-Монте Карло подхода, описан в **Алгоритъм 2**, може да се очаква по-бърза сходимост. С цел получаване на статистическа грешка се използват рандомизирани редици на Соболев, с което се добавя контролирана случайност без да се губят предимствата на детерминистичните редици с малък дискрепанс.

Следва да отбележим, че стохастичната грешка зависи не само от N , но и от параме-

търа k , тъй като с увеличаване на k нараства броят на оценяваните случайни величини $\theta^{(i)}, i = 1, 2, \dots, k + 1$. Това води до натрупване на дисперсии както в числителя, така и в знаменателя на формулите (2.43) и (2.44), което от своя страна влияе върху необходимия брой реализации N . Следователно, постигането на оптимален баланс между двата вида грешки изисква избор на такъв брой реализации N , при който стохастичната грешка е от същия порядък като предварително зададената систематична грешка $\varepsilon_{sys} = 0.001$. По този начин се избягва доминирането на единия източник на грешка и се предотвратява неефективното използване на изчислителни ресурси без съществено подобрение на крайната абсолютна точност.

Резултатите за матрицата A с размерност $n = 500$ (Таблица 2.1) и фиксирано $q = -0.157428$, потвърждават направените по-горе теоретични изводи. Наблюдава се, че при подходящ избор на параметрите (m, k) , гарантиращ ограничаването на систематичната грешка под предварително зададен праг 10^{-3} , може да се постигне съществено по-малка абсолютна грешка при сравнително малък брой реализации N , съпоставим с размерността на матрицата. Най-добрата точност е получена при комбинацията $k = 5, m = 10$ и $N = 512$. При тези стойности на параметрите абсолютната грешка достига минимални стойности, които са от същия порядък като предварително зададената систематична грешка.

От друга страна, с увеличаване на N (например, при $N = 16384$ и $N = 32768$, виж Фигура 2.1 и Таблица 2.1) не се наблюдава монотонно намаляване на абсолютната грешка. Вместо това тя често достига ниво на насищане или проявява колебливо поведение. Това е характерно за ситуации, при които систематичната грешка контролирана от параметрите m и k , вече определя долна граница за постижимата точност. Допълнително влияние оказват и числени ефекти, свързани с нарастването на биномните коефициенти при по-големи стойности на m и/или k , които увеличават чувствителността на дробно-рационалната формула (2.44) към малки флуктуации в оценките на случайните величини $\theta^{(i)}, i = 1, \dots, k$. Следователно, в разглежданите числени експерименти увеличаването на N не гарантира автоматично намаляване на абсолютната грешка, когато систематичната компонента е вече определяща или когато усилването породено от биномните коефициенти, прави метода по-чувствителен към остатъчния стохастичен шум.

2.4.2 Случай (ii): Размер на матрицата $n=1000$

В Таблица 2.2 са представени резултати от оценяването на най-малката собствена стойност на симетричната матрица с размерност $n = 1000$, получени чрез *почти оптималните* алгоритми $RMC_{(MT)}$ и $RQMC_{(S)}$. В таблицата са включени и съответните абсолютни грешки на получените оценки. В проведените числени експерименти параметърът на ускорение е фиксиран на стойност $q = -0.2302458$, като абсолютната му стойност е по-голяма в сравнение със случая на матрица с размерност $n = 500$. Увеличаването на $|q|$ води до нарастване на параметъра $t = |q||\lambda_1|$. В резултат на това, за да бъдат изпълнени условията (2.27) и (2.29) при същите параметри k и m (виж Таблица 2.1) е необходимо систематичната грешка да бъде фиксирана на по-високо ниво $\varepsilon_{sys} = 0.01$.

Резултатите за матрицата A с размерност $n = 1000$ (виж Таблица 2.2 и Фигура 2.2) показват, че най-добрият оптимален баланс между систематичната и стохастичната ком-

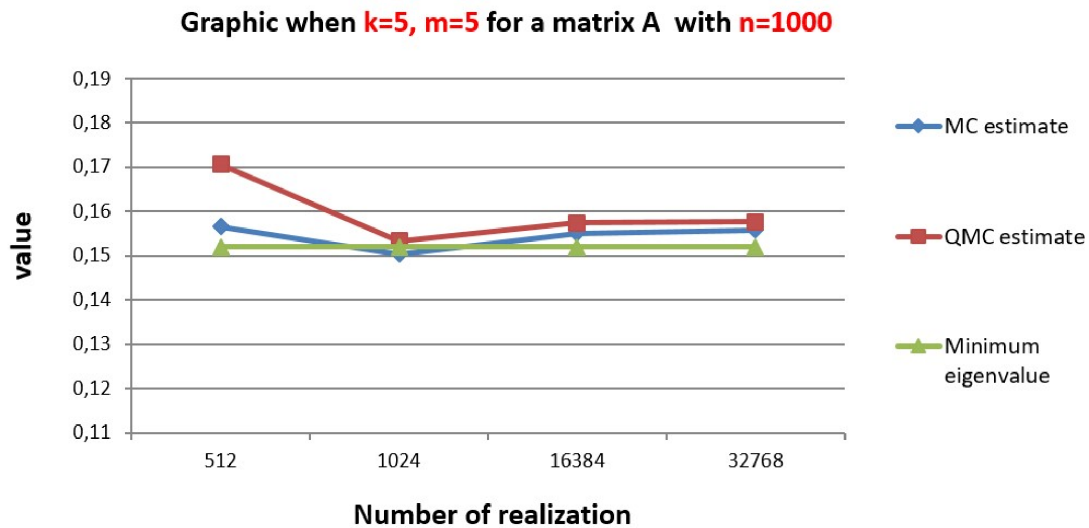
Таблица 2.2: Числени резултати за матрицата $A \in \mathbb{R}^{1000 \times 1000}$ при различни стойности на N, k и m с параметър на ускорение $q = -0.2302458$. Точната минимална стойност $\lambda_{\min} = \lambda_{1000} = 0.151911$.

N	k	m	λ_{\min}	Absolute error	λ_{\min}	Absolute error
			$RMC_{(MT)}$	$RMC_{(MT)}$	$RQMC_{(S)}$	$RQMC_{(S)}$
$512 * 2^0$	5	5	0.1556059	0.0045949	0.1705751	0.0186641
	5	10	0.168251	0.01634	0.1886959	0.0367849
	10	5	0.1593428	0.0074318	0.1701307	0.0182197
$512 * 2^1$	5	5	0.1504153	0.0014957	0.1533797	0.0014687
	5	10	0.1417805	0.0101305	0.1499974	0.0019136
	10	5	0.156739	0.004828	0.1534383	0.0015273
$512 * 2^5$	5	5	0.1550015	0.0030905	0.1573842	0.0054732
	5	10	0.1495483	0.0023627	0.1539357	0.0020247
	10	5	0.1559548	0.0040438	0.1575454	0.0056344
$512 * 2^6$	5	5	0.1557509	0.0038399	0.1576339	0.0057229
	5	10	0.1537008	0.0017898	0.1562911	0.0043801
	10	5	0.1562905	0.0043795	0.15764225	0.0057312

понента на грешката се постига при стойности $k = 5, m = 5$ и $N = 1024$. При тази комбинация абсолютната грешка е минимална и е от същия порядък като предварително зададената систематична грешка. Това потвърждава теоретичните оценки, според които параметрите m и k , определени въз основа на фиксирано ниво на систематична грешка, осигуряват необходимата точност на апроксимацията на резолвентната матрица. Параметърът N следва да бъде избран така, че стохастичната грешка да бъде съпоставима по порядък със систематичната.

Резултатите, получени чрез алгоритми $RMC_{(MT)}$ и $RQMC_{(S)}$ при по-големи стойности на m, k и N , показват известно подобрене на абсолютната грешка по отношение на константния множител, но без промяна в нейния порядък. Същевременно, увеличаването на тези параметри води и до нарастване на изчислителната сложност на алгоритмите, което налага търсене на компромис между постижима точност и необходимите изчислителни ресурси. Освен това, с увеличаване на N (например, при $N = 16384$ и $N = 32768$, виж Фигура 2.2), не се наблюдава монотонно намаляване на абсолютната грешка. Това поведение може да бъде обяснено с числени ефекти, свързани с повишената чувствителност на дробно-рационалната формула (2.44) към малки флуктуации в оценките на случайните величини $\theta^{(i)}, i = 1, \dots, k + 1$. В тези случаи систематичната грешка и усилването на остатъчния стохастичен шум ограничават възможностите за допълнително подобряване на точността чрез увеличаване на N .

От практическа гледна точка, резултатите и за двата разгледани случая на симетрични плътни матрици с размерност от порядъка на няколко хиляди елементи показват, че оптималният баланс между систематичната и стохастичната грешка се постига, когато параметрите m и k са избрани съгласно теоретичните оценки при предварително фиксирано ниво на систематична грешка, а броят реализации N е от същия порядък като размерността на матрицата. Следва да отбележим, че изборът на нивото на систематичната грешка зависи от стойността на параметъра на ускорение q , както и от наличната



Фигура 2.2: Сходимост към точната минимална собствена стойност на матрицата A при използване на **Алгоритъм 2** и **Алгоритъм 2а** с параметри $k = 5, m = 5$ за различни на брой N

априорна информация за нормата на матрицата. При твърде малки стойности на N стохастичната грешка доминира и води до недостатъчна точност на оценките. Обратно, при прекомерно големи стойности на N не се наблюдава съществено подобрене на точността, докато изчислителната цена на **Алгоритъм 2** и **Алгоритъм 2а** нараства значително. Това подчертава необходимостта от внимателен избор на параметъра N , съобразен както с теоретичните оценки, така и с наличните изчислителни ресурси.

2.5 Обобщение

Оценяването на минималната собствена стойност λ_{min} на дадена симетрична матрица A чрез степенните стохастични алгоритми, изследвани в Глава 1 изисква намирането на обратната матрица A^{-1} , което представлява изчислително скъпа операция при задачи с голяма размерност. В настоящата глава е изследван Стохастичен Резолвентен метод за оценяване на минималната собствена стойност на симетрични матрици, който представлява ефективна алтернатива на степенния стохастичен метод. Основното предимство на Резолвентния метод е, че елиминира необходимостта от явно изчисляване на обратната матрица, като вместо това използва оценяване на резолвентната матрица $R_q = (\mathbf{I} - qA)^{-1}$ и нейните степени. Знакът на параметъра на ускорението q определя дали се оценява минималната собствена стойност λ_{min} или максималната собствена стойност λ_{max} . Установено е, че сходимостта на Резолвентния Степенен метод зависи съществено както от параметъра на ускорение q , така и от степента m на резолвентната матрица.

Получените теоретични оценки [24, 25] показват, че систематичната грешка намалява с увеличаване на параметъра m , докато скоростта на сходимост се определя от отношението между съответните крайни собствени стойности на матрицата в зависимост от знака

на ускоряващия параметър. Допълнително, въвеждането на прекъсване на биномния ред чрез параметъра k води до нов компонент на систематичната грешка, за която бяха получени аналитичните оценки (2.27), (2.29), (2.31), (2.40). На тази основа беше установена линейна зависимост между параметрите m и k , която позволява определянето на минималните им стойности при предварително фиксиране ниво на систематична грешка. Този резултат има съществено практическо значение, тъй като позволява оптимален избор на параметрите и съответно намаляване на изчислителната сложност на алгоритъма.

В стохастичната постановка на метода беше показано, че освен систематичната грешка възниква и стохастична грешка. За практическа реализация бяха предложени Резолвентен Монте Карло (РМК) и Резолвентен Квази-Монте Карло (РКМК) алгоритми, при които математическите очаквания $E[\theta^{(i)}]$ се апроксимират чрез извадкови средни, получени от N реализации на прекъснатата верига на Марков с k прехода. Теоретичният анализ показва, че стохастичната грешка зависи основно от броя реализации N , като при МК подхода тя е от порядък $\mathcal{O}(N^{-1/2})$, докато при КМК подхода се оценява чрез неравенството на Коксма–Хлавка при фиксирана размерност на редиците с малък дискрепанс. Систематичната грешка, от своя страна, се дължи на прекъсването на безкрайния ред, чрез който се представя резолвентната матрица $R_q^m = (\mathbf{I} - qA)^{-m}$, при $i = k$ (виж формула (2.19)).

Едновременното наличие на стохастична и систематична грешка наложи да се търсят зависимости между параметрите m , k и N , при които и двата вида грешки са достатъчно малки и от съпоставим порядък спрямо предварително зададената точност. Подходящият избор на тези параметри е съществен при конструирането на ефективни резолвентни стохастични алгоритми, тъй като позволява постигането на необходимата точност без неоправдано увеличаване на изчислителните разходи и гарантира надеждността на получените числени резултати.

Получените теоретични резултати (2.27), (2.29), (2.31), (2.40) предоставят практически насоки за избора на параметрите m и k , при фиксирано ниво на систематичната грешка, при условие че $t = |q||\lambda_{max}| < 1$. Паралелно с това, при зададена обща точност ε броят реализации N може да бъде определен чрез оценката (2.40), така че стохастичната грешка да бъде съпоставима по порядък със систематичната. Предложените *почти оптимални* РМК и РКМК алгоритми, представени чрез псевдокодове, осигуряват допълнително намаляване на дисперсията на стохастичните оценки и повишават ефективността на Резолвентния Стохастичен метод. Използването на редици с малък дискрепанс, като редиците на Соболев, позволява допълнително подобряване на точността и стабилността на получените резултати. Проведените числени експерименти за симетрични матрици с размерности $n = 500$ и $n = 1000$ потвърждават теоретичните оценки и демонстрират ефективността на предложените алгоритми. Установено беше, че най-добра точност се постига при стойности на параметрите m и k , съответстващи на теоретично предсказаните зависимости при фиксирана систематична грешка, както и при стойности на броя реализации N , при които стохастичната грешка е съизмерима със систематичната. Наблюдавано беше също, че увеличаването на N над определена граница не води до съществено подобряване на точността, което потвърждава, че общата грешка се определя от баланса между двата вида грешки. Освен това, сравнението между **Алгоритъм 2** и **Алгоритъм 2а** показва, че използване-

то на рандомизирана редица на Собол може да осигури по-бърза сходимост и по-висока точност, независимо че размерността на използваните редици, зависи от по-високи стойности за k . Част от получените резултати, описани в тази глава, са публикувани в статиите [41, 42].

В обобщение, разработените **Алгоритъм 2** и **Алгоритъм 2а** могат да бъдат използвани при оценяване на минимални собствени стойности на симетрични матрици с голяма размерност. Ще отбележим, че те могат да се приложат и за оценяване на максимални собствени стойности на симетрични матрици при положителна стойност на ускоряващия параметър. Получените теоретични оценки и числени резултати показват, че чрез подходящ избор на параметрите m , k и N може да се постигне оптимален баланс между точност и изчислителна сложност. Формулираният принцип за балансиране на грешките предоставя практически механизъм за избор на параметрите. Освен това, реализацията на РКМК алгоритъм в хибриден GPU/CPU режим позволява ефективно използване на наличните ресурси и бързо получаване на надеждни оценки за λ_{min} за матрици с голяма размерност. Това прави метода особено подходящ за приложения, при които *класическите* детерминистични методи са трудно приложими поради високата размерност или високата изчислителна цена, и създава основа за бъдещи изследвания и приложения на Стохастичните Резолвентни методи в различни научни и приложни области.

Глава 3

Стохастични методи за оценяване на пазарния риск на инвестиционни портфейли

В тази глава от дисертационния труд се представя приложение на разработения **Алгоритъм 1**, описан в Глава 1, към реален проблем в областта на финансовата математика, а именно оценяване на пазарния риск на инвестиционен портфейл. По този начин се осъществява преход от теоретичен и тестови анализ на стохастичните степенни методи и алгоритми към тяхното практическо използване в контекста на финансовото моделиране. Целта е да се демонстрира приложимостта и ефективността на разработените стохастични числени алгоритми при количественото измерване на пазарния риск, базирано на спектралните характеристики на корелационната матрица на доходностите на активите. В частност, оценяването на екстремална собствена стойност се използва като инструмент за идентифициране на доминиращите рискови фактори в портфейла и за анализ на неговата чувствителност към пазарни колебания. По този начин се илюстрира практическата значимост на разработените алгоритми и тяхната приложимост при задачи, свързани с оценяване на пазарния риск и вземане на инвестиционни решения.

3.1 Постановка на задачата

Основната задача, която се поставя в тази глава, е да се оцени максималната собствена стойност на корелационни матрици и да се анализира точността на получените оценки чрез прилагането на разработените почти оптимални Степенен Монте Карло (СМК) и Степенен Квази-Монте Карло (СКМК) алгоритми, представени в Глава 1. Тези алгоритми се използват за численото оценяване на най-голямата собствена стойност на матриците, като същевременно позволяват числено да се изследва грешката на приближението. В рамките на СКМК алгоритъм се въвеждат допълнително два оператора: (1) оператор за пропускане (англ. *skipping operator*) и (2) оператор за прескачане през стъпка (англ. *leaping operator*). Тяхното използване е мотивирано от необходимостта да се намали влиянието на началните неравномерности, характерни за рандомизираните редици с малък дискрепанс.

3.2 Методи и данни

Финансовите данни, използвани за конструирането на корелационните матрици (виж Приложение), са извлечени от два основни източника: (1) уеб платформата *Portfolio Visualizer* и (2) базата данни на *Yahoo Finance*. Достъпът до втория източник е реализиран чрез програмната библиотека *yFinance*, която е достъпна чрез хранилището на платформата *PyPI*.

При оценяването на най-голямата собствена стойност на конструираните корелационни матрици, получени от финансовите данни, се извършва анализ на баланса между систематичната и стохастичната грешка. От друга страна, с цел интерпретация на резултатите в контекста на финансовия анализ се оценява пазарният риск на разглежданите портфейли чрез изчисляване на дяла на обяснената вариация (англ. *Fraction of Variance Explained*, FVE), съответстващ на максималната собствена стойност на корелационната матрица. Този показател позволява да се определи степента, в която съвместната динамика на доходностите на активите се обяснява от доминиращия общ фактор на пазара.

3.2.1 Обща формулировка на ковариационна и корелационна матрица

При анализа на зависимостите между две или повече случайни величини (променливи), основен интерес представляват характеристиките ковариация и коефициент на корелация, които описват степента и посоката на линейната зависимост между тях. Ковариацията между две случайни величини X и Y се дефинира чрез следната зависимост:

$$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)], \quad (3.1)$$

където $\mu_X = E[X]$ и $\mu_Y = E[Y]$ са математическите очаквания на случайните величини X и Y [99, 118]. Ковариацията представлява мярка за линейна зависимост между две случайни променливи. Нейната стойност зависи от използваните мерни единици на разглежданите величини (например, ръст, измерван в сантиметри и тегло, измервано в килограми). Стойностите на ковариацията не са ограничени и могат да приемат както положителни, така и отрицателни стойности. Положителната ковариация показва, че променливите имат тенденция да се изменят в една и съща посока, докато отрицателната ковариация показва, че те се изменят в противоположни посоки.

Коефициентът на корелация (известен като коефициент на корелация на Пиърсън) се дефинира чрез нормализиране на ковариацията със стандартните отклонения на съответните променливи:

$$Corr(X, Y) = \rho_{XY} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}, \quad (3.2)$$

където σ_X и σ_Y са стандартните отклонения на случайните величини X и Y [99, 118]. Корелационният коефициент представлява нормализирана мярка на ковариацията, която позволява сравнение на линейните зависимости между различни променливи независимо от техните мащаби и мерни единици. Стойностите на този коефициент са ограничени в

интервала $[-1, 1]$. При стойност -1 се наблюдава отрицателна линейна зависимост, при стойност 0 липсва линейна зависимост, а при стойност 1 е налице положителна линейна зависимост между разглежданите случайни величини.

Нека е даден случаен вектор $\mathbf{X} = (X_1, \dots, X_n)^T$, чиито компоненти се разглеждат като случайни величини. Тогава ковариационната матрица на този вектор се дефинира чрез:

$$\Sigma = Cov(\mathbf{X}) = E [(\mathbf{X} - \mu)(\mathbf{X} - \mu)^T] = \begin{bmatrix} Cov(X_1, X_1) & \dots & Cov(X_1, X_n) \\ \vdots & \ddots & \vdots \\ Cov(X_n, X_1) & \dots & Cov(X_n, X_n) \end{bmatrix}, \quad (3.3)$$

където елементите $\Sigma_{ij} = Cov(X_i, X_j)$ се изчисляват чрез формула (3.1), а диагоналните елементи са $\Sigma_{ii} = \sigma_i^2$, т.е. представляват дисперсиите (вариациите) на съответните случайни величини [68, 99]. Съответно, корелационната матрица \hat{R} на случайния вектор \mathbf{X} се задава чрез:

$$\hat{R} = [\rho_{ij}]_{i,j=1}^n, \quad \text{където } \rho_{ij} = Corr(X_i, X_j), \quad i, j = 1, \dots, n. \quad (3.4)$$

Елементите ρ_{ij} на корелационната матрица \hat{R} се определят съгласно формула (3.2), като диагоналните елементи винаги са равни на единица, т.е. $\rho_{ii} = 1$ [49, 68].

В контекста на анализа на портфейли от финансови активи, корелационната матрица предоставя съществена информация за силата на линейните зависимости между доходностите на различните активи [49]. По-специално, най-голямата собствена стойност на корелационната матрица често се интерпретира като индикатор за наличието на доминиращ общ фактор в динамиката на пазара. В условията на силна пазарна синхронизация тази собствена стойност нараства, което означава, че систематичният риск играе водеща роля, а възможностите за ефективна диверсификация на портфейла намаляват. Например, ако два актива са перфектно положително корелирани ($\rho = 1$), включването им едновременно в портфейла не води до диверсификационен ефект. Обратно, при по-малки или отрицателни стойности на корелацията активите могат да допринесат за намаляване на общия риск на портфейла чрез диверсификация.

От друга страна, ковариационната матрица обобщава зависимостите между всички активи в портфейла чрез включване на всички двойки ковариации. Тази матрица играе фундаментална роля в оптимизационните модели за управление на портфейли, тъй като позволява количествено измерване на риска и взаимозависимостите между активите в портфейли, съдържащи множество финансови инструменти.

3.2.2 Конструиране на ковариационна и корелационна матрица от емпирични данни от финансови пазари

В емпиричните изследвания на финансовите пазари входните данни най-често се представят под формата на матрица от наблюдения върху доходностите на активи. Нека $B \in \mathbb{R}^{r \times n}$ е матрицата с данни, в която броят на редовете r съответства на броя на наблюденията на доходностите (например, дневни/месечни или годишни наблюдения), а броят на колоните

n съответства на броя на разглежданите величини. При анализ на финансови пазари тези променливи обикновено представляват доходностите на различни активи или компании. Така всяка колона на матрицата B съдържа наблюденията за доходността на даден актив, а всеки ред представя едновременно наблюдение на доходностите на всички активи в даден момент от време.

Матрицата B може да се представи по следния начин:

$$B = [b_{ij}]_{i,j=1}^{r,n} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{r1} & b_{r2} & \cdots & b_{rn} \end{bmatrix} = (\mathbf{b}_1, \dots, \mathbf{b}_i, \dots, \mathbf{b}_r)^T, \quad (3.5)$$

където $\mathbf{b}_i = (b_{i1}, \dots, b_{in})$ представлява n -мерен вектор от наблюдавани доходности за всички активи в момента i , ($i = 1, \dots, r$).

Центриране на данните и извадкова ковариационна и корелационна матрица

Нека $\bar{\mathbf{b}}$ е n -мерният вектор на извадковите средни стойности на колоните на матрицата B , дефиниран по следния начин:

$$\bar{\mathbf{b}} = \left(\frac{1}{r} \sum_{i=1}^r b_{i1}, \dots, \frac{1}{r} \sum_{i=1}^r b_{in} \right).$$

Центрираната матрица на наблюденията се дефинира чрез изваждане на съответните средни стойности от всяко наблюдение:

$$B_c = B - \mathbf{e} \bar{\mathbf{b}}^T, \quad (3.6)$$

където $\mathbf{e} \in \mathbb{R}^r$ е вектор от единици. Извадковата ковариационна матрица на наблюдаваните доходности се оценява чрез:

$$\Sigma = \frac{1}{r} B_c^T B_c. \quad (3.7)$$

При малки стойности на r във формула (3.7) делението се извършва с $r - 1$. Елементът Σ_{ij} на тази матрица измерва съвместната изменчивост между доходностите на активите i и j . Диагоналните елементи Σ_{ii} представляват извадковите вариации на съответните активи. Ако \mathbf{w} е векторът на портфейлните тегла, дисперсията на доходността на портфейла Π се задава чрез квадратичната форма:

$$Var(\Pi) = \mathbf{w}^T \Sigma \mathbf{w}, \quad (3.8)$$

което показва, че рискът на портфейла зависи както от индивидуалните вариации на отделните активи (диагоналните елементи на ковариационната матрица), така и от взаимните зависимости между тях (извъндиагоналните елементи). Нека стандартните отклонения на доходностите са дадени чрез:

$$\sigma_i = \sqrt{\Sigma_{ii}}, \quad i = 1, \dots, n.$$

Диагоналната матрица на стандартните отклонения се дефинира по следния начин:

$$D_\sigma = \text{diag}(\sigma_1, \dots, \sigma_n).$$

Тогава извадковата корелационна матрица се получава чрез нормализиране на ковариационната матрица:

$$\hat{R} = D_\sigma^{-1/2} \Sigma D_\sigma^{-1/2}. \quad (3.9)$$

На ниво елементи тази зависимост може да бъде записана като:

$$\hat{R} = [\rho_{ij}]_{i,j=1}^n, \quad \rho_{ij} = \frac{\Sigma_{ij}}{\sigma_i \sigma_j}, \quad i, j = 1, \dots, n.$$

Диагоналните елементи на корелационната матрица са равни на единица, тъй като всяка променлива е перфектно корелирана със самата себе си [49].

3.2.3 Дял на обяснената вариация

Дял на обяснената вариация (англ. *Fraction Variance Explained*, FVE) представлява количествена характеристика, свързана със собствените стойности на ковариационната или корелационната матрица. Тази мярка намира широко приложение в методите за многомерния статистически анализ, като главния компонентен анализ (англ. *Principal Component Analysis*, PCA), както и във факторните модели [55]. Тя служи за оценка на това каква част от общата вариация в данните може да бъде обяснена от един или няколко водещи компонента или фактори.

Нека $A \in \mathbb{R}^{n \times n}$ е симетрична положително полуопределена матрица със собствени стойности, подредени в низходящ ред: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$. В този случай дялът на обяснената вариация, FVE, съответстващ на най-голямата собствена стойност λ_1 , се дефинира като частно между тази стойност и сумата на всички собствени стойности на матрицата:

$$FVE_1 = \frac{\lambda_1}{\sum_{j=1}^n \lambda_j} = \frac{\lambda_1}{\text{trace}(A)}. \quad (3.10)$$

Тъй като сумата на собствените стойности е равна на следата (англ. *trace*) на матрицата, тази величина измерва каква част от общата вариация се обяснява от първата главна компонента. Аналогично, ако се разглеждат първите k най-големи собствени стойности на матрицата A , тогава дялът на обяснената вариация се определя като частно между сумата на тези k стойности и сумата на всички собствени стойности на матрицата. По този начин се оценява каква част от общата вариация може да бъде описана от първите k главни компоненти, т.е.:

$$FVE_k = \frac{\sum_{i=1}^k \lambda_i}{\text{trace}(A)}. \quad (3.11)$$

Важен случай е, когато матрицата A съвпада с извадковата корелационна матрица \hat{R} . В

този случай следата на матрицата е равна на размерността на матрицата, т.е. на броя на променливите n , т.е. $trace(\hat{R}) = n$. Това означава, че сумата от всички собствени стойности на корелационната матрица също е равна на n . Следователно, дялът на обяснената вариация, съответстващ на най-голямата собствена стойност, се получава като отношение между λ_1 и n :

$$FVE_1 = \frac{\lambda_1}{n}. \quad (3.12)$$

Корелационната матрица \hat{R} обикновено е симетрична и положително полуопределена. При големи размерности на корелационната матрица за оценяването на максималната собствена стойност можем да приложим разработените почти оптимални стохастични алгоритми от Глава 1.

В емпиричните изследвания на финансовите пазари, по-специално при анализа на портфейли от активи, често се наблюдава, че най-голямата собствена стойност на корелационната матрица обяснява значителна част от общата вариация. Това явление се интерпретира като наличие на доминиращ общ фактор, който обикновено се асоциира със системния пазарен риск. Когато стойността на първия дял на обяснената вариация е висока, това означава, че съществена част от съвместното движение на доходностите на активите може да бъде описана чрез един доминиращ фактор.

При корелационни матрици с по-малки размерности, големите стойности на максималното собствено число на тази матрица често водят до висок дял на обяснената вариация FVE_1 . Това от своя страна е индикатор за силно колективно движение на активите и за по-малка ефективност на диверсификацията. В такава ситуация значителна част от риска на портфейла произтича от общи пазарни шокове, които трудно могат да бъдат неутрализирани чрез разпределение на капитала между множество активи. От друга страна, по-малките стойности на FVE_1 предполагат децентрализирана структура на зависимостите между активите, при която рискът е разпределен между по-голям брой относително независими фактори. Това създава по-благоприятни условия за прилагане на оптимизационни модели за управление на портфейли, насочени към подобряване на съотношението между риск и доходност.

Поради тези причини оценяването на най-голямата собствена стойност на корелационната матрица и съответния дял на обяснената вариация представлява важен етап в анализа на пазарния риск на портфейл от финансови активи.

3.3 Стохастични алгоритми за оценяване на максималната собствена стойност на корелационна матрица

При оценяването на най-голямата собствена стойност на двете разглеждани корелационни матрици в секция 3.4 се прилагат разработените в Глава 1 на дисертационния труд *почти оптимални* Степенен Монте Карло (СМК) и Степенен Квази-Монте Карло (СКМК) алгоритми. За тези алгоритми бе показано, че превъзхождат *класическите* им аналози, тъй като при съпоставими изчислителни разходи осигуряват по-висока числена точност и по-стабилна (по-гладка) сходимост на оценките.

Почти оптималният СКМК алгоритъм е представен чрез псевдокод и е означен като **Алгоритъм 1** в Глава 1, секция 1.3.2. В стъпка 4 на **Алгоритъм 1** се генерират разбъркани редици на Холтън и Собол, които се използват при построяването на веригата на Марков при краен брой стъпки.

В числените експерименти, представени в следващата секция, вариантите на **Алгоритъм 1** на базата на разбъркани редици на Холтън (H) и Собол (S), се реализират чрез подпрограми в MATLAB. Съответните процедури за генериране на тези редици са описани чрез псевдокода на **Алгоритъм 3** и **Алгоритъм 4**.

Algorithm 3 Pseudocode for generating a scrambled Halton Sequence with *skip* and *leap* parameters

- 1: CREATE FUNCTION: `generateHaltonSequence` (N, k+1)
 - 2: GENERATE: $p \leftarrow \text{haltonset}(k, \text{'skip'} = s_1, \text{'leap'} = l_1)$
 - 3: GENERATE: $p \leftarrow \text{scramble}(p, \text{'RR2'})$
 - 4: GENERATE: $\text{points} \leftarrow \text{net}(p, N)$
 - 5: OUTPUT: Return *points*
-

Algorithm 4 Pseudocode for generating a scrambled Sobol Sequence with *skip* and *leap* parameters

- 1: CREATE FUNCTION: `generateSobolSequence` (N, k+1)
 - 2: GENERATE: $p \leftarrow \text{sobolset}(k, \text{'skip'} = s_1, \text{'leap'} = l_1)$
 - 3: GENERATE: $p \leftarrow \text{scramble}(p, \text{'MatousekAffineOwen'})$
 - 4: GENERATE: $\text{points} \leftarrow \text{net}(p, N)$
 - 5: OUTPUT: Return *points*
-

При избор на $s_1 = 0$ и $l_1 = 0$ в **Алгоритъм 3** и **Алгоритъм 4** се генерират разбъркани редици на Холтън и Собол без допълнителни операции по пропускане на начални точки (англ. *skip*) и прескачане през стъпка (англ. *leap*). В този случай, съответните версии на почти оптималния СКМК алгоритъм се означават с $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$, където индексът (d) обозначава настройка по подразбиране (англ. *default*). Когато параметрите s_1 и l_1 приемат ненулеви стойности, редиците се конструират с прилагане на операторите *skip* и *leap*, които позволяват контрол върху началната точка и подизбор на елементи от редицата. В този случай съответните версии на почти оптималния СКМК алгоритъм се означават с $PQMC_{(H)}^{(s,l)}$ и $PQMC_{(S)}^{(s,l)}$.

За реализацията на почти оптималния Степенен Монте Карло (СМК) алгоритъм, т.е. **Алгоритъм 1a** от Глава 1, се използват генератори на псевдослучайни числа *Mersenne Twister* (MT) и *Middle Square* (MS). Съответните реализации се означават с $PMC_{(MT)}$ и $PMC_{(MS)}$.

Всички алгоритми и подпрограми са реализирани в MATLAB версия 9.10 (R2021a). Използваният синтаксис на функциите `haltonset`, `sobolset` и `scramble` в **Алгоритъм 3** и **Алгоритъм 4** е съгласуван с официалната документация на MATLAB [120, 121].

3.4 Числени резултати и дискусия

В настоящата секция са представени числени резултати за оценяване на максималната собствена стойност на две корелационни матрици чрез разработените *почти оптимални* СКМК и СМК алгоритми. Разгледаните корелационни матрици са конструирани на базата на доходностите на два портфейла от активи, което позволява да се сравни поведението на алгоритмите при различни размерности и различна информационна наситеност на данните.

Първата корелационна матрица се получава от месечните доходности на портфейл от $n = 32$ глобални актива, извлечени от платформата *Portfolio Visualizer*, за периода от януари 2007 г. до март 2024 г. (виж Приложение). Втората корелационна матрица е изградена на базата на доходностите на $n = 490$ компании, включени в индекса *S&P500*, оценяващ състоянието на американския фондов пазар за петгодишен период (2021-2025 г.) (виж Приложение). За данните от *Portfolio Visualizer* размерностите са $r = 207$ и $n = 32$, което отговаря на 207 месечни наблюдения за период от 17 години и 3 месеца. За данните, извлечени от *Yahoo Finance*, размерностите са $r = 5$ и $n = 490$, т.е. петгодишни наблюдения за 490 компании. Прилагайки последователно центрирането (т.е. формула (3.6)), оценката на ковариационната матрица чрез (3.7) и нормализацията до корелационна матрица чрез (3.9), се получават следните две корелационни матрици:

- (i) Корелационна матрица $\hat{R} = [\rho_{ij}]_{i,j=1}^n$ с размерност $n = 32$, $\lambda_{max} = \lambda_1 = 18.14714$ и дял на обяснена вариация $FVE_1 = 56.71\%$;
- (ii) Корелационна матрица $\hat{R} = [\rho_{ij}]_{i,j=1}^n$ с размерност $n = 490$, $\lambda_{max} = \lambda_1 = 201.090512$ и дял на обяснена вариация $FVE_1 = 41.04\%$;

Точната стойност на максималното собствено число за двете матрици са изчислени посредством вградените функции на MATLAB и се използват като референтни величини при оценката на числената точност на разглежданите алгоритми.

В проведените числени експерименти произволните вектори \mathbf{h} и \mathbf{f} , използвани при конструирването на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$, са избрани така, че да имат L_1 -норма:

$$\|\mathbf{h}\| = \|\mathbf{f}\| = \sum_{i=1}^n \frac{1}{n} = 1.$$

Числените експерименти са проведени за матрица $A = q\hat{R}$ с **Алгоритъм 1** и **Алгоритъм 1a** от Глава 1, като целта е да се гарантира условието за сходимост $\|A\| < 1$. Това условие е изпълнено при избор на положителен параметър q от вида:

$$q = \frac{1}{|\lambda_1| + \delta}, \quad \delta > 0.$$

Когато предварителната оценка за λ_1 не е налична, параметърът q може да се избере и

по алтернативен начин алтернативен начин чрез L_1 -нормата на \hat{R} :

$$q = \frac{1}{\|\hat{R}\|_1 + \delta}, \quad \|\hat{R}\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |\rho_{ij}| \quad \text{и} \quad \delta > 0.$$

3.4.1 Случай (i): корелационна матрица с размерност $n = 32$

В Таблица 3.1 са представени числените резултати за оценяване на най-голямата собствена стойност на корелационната матрица \hat{R} при размерност $n = 32$. Резултатите са получени чрез *почти оптималния* СМК алгоритъм като са използвани два генератора на псевдослучайни числа: *Mersenne Twister* (MT) и *Middle Square* (MS). Съответните реализации на **Алгоритъм 1a** са означени с $PMC_{(MT)}$ и $PMC_{(MS)}$. Максималната собствена стойност $\lambda_1 = \lambda_{\max}$ е оценявана при различен брой преходи k във веригата на Марков и различен брой реализации N на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$. В таблицата са включени и абсолютните грешки на съответните оценки.

Получените резултати показват, че при фиксиран брой реализации N и увеличаването на броя преходи k води до намаляване на абсолютната грешка, като това намаляване се стабилизира около $k = 8$. Това е индикатор за достигане на практически оптимален баланс между систематичната и стохастичната компонента на грешката. По-нататъшното увеличаване на k не води до съществено подобрене на точността. От друга страна, при увеличаване на броя симулации N се наблюдава монотонно намаляване на абсолютната грешка, което е в съответствие с очакваното поведение на Монте Карло оценките. Минималните абсолютни грешки за всеки от алгоритмите $PMC_{(MT)}$ и $PMC_{(MS)}$ са маркирани с удебелен шрифт в Таблица 3.1. Резултатите допълнително показват, че при разглежданата матрица порядъкът на грешката при двата генератора е съпоставим.

На Фигура 3.1 са представени приближените оценки на $\lambda_1^{(k)}$ при нарастващ брой реализации N , при фиксирана балансираща стойност $k = 8$. Графичното сравнение илюстрира сходимостта на приближените решения към точната стойност λ_1 при увеличаване на N . Изводът, който може да бъде направен е, че и двата алгоритъма се стремят към референтната стойност на λ_{\max} , но $PMC_{(MT)}$ демонстрира по-висока точност и по-устойчива сходимост спрямо $PMC_{(MS)}$, особено при по-малки стойности на N .

Получените резултати показват, че за постигане на оперативна точност до два или три знака след десетичната запетая са достатъчни $N = 512$ или $N = 1024$ симулации. Подобна точност е адекватна за количествена оценка на систематичния пазарен риск в контекста на портфейлни приложения. Прилагайки формула (3.12) за корелационната матрица с размерност $n = 32$ и използвайки оценките при $N = 1024$, получаваме $FVE_1 \approx 56.70\%$ за $PMC_{(MT)}$ и $FVE_1 \approx 56.72\%$ за $PMC_{(MS)}$.

Тези резултати потвърждават, че *почти оптималният* СМК алгоритъм представлява надежден подход за оценяване на максималната собствена стойност на корелационна матрица с неголяма размерност при сравнително малък брой симулации. Изборът на подходящ генератор и на балансиращ брой итерации k способства за ефективно управление на компромиса между систематична и стохастична грешка.

Получената приближена оценка за λ_1 с точност от порядъка $\mathcal{O}(10^{-3})$ позволява да се

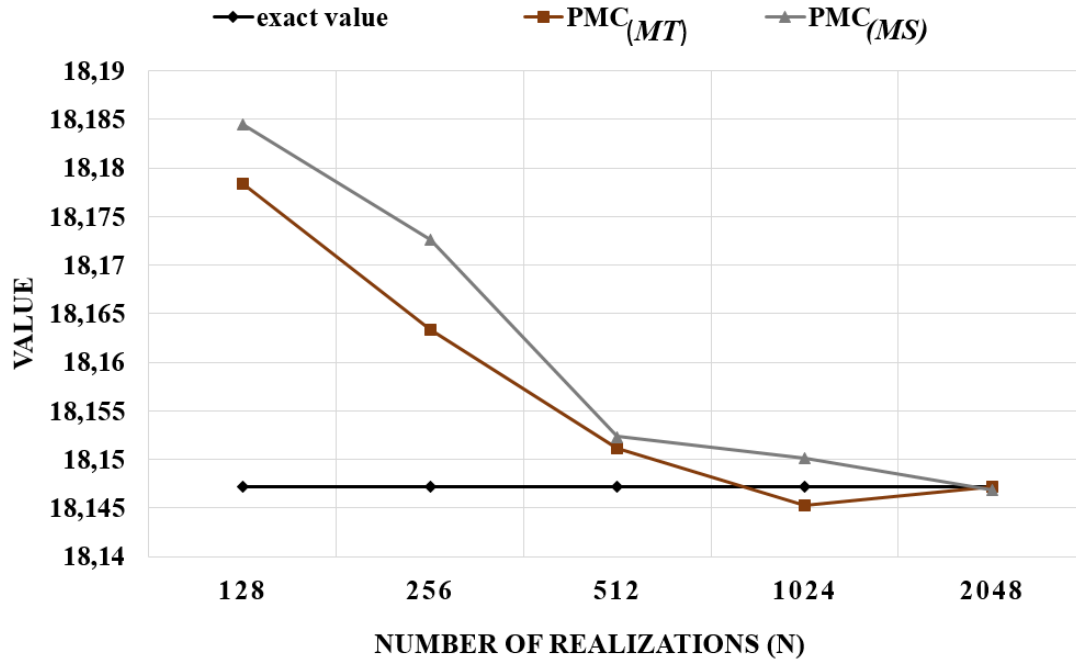
Таблица 3.1: Резултати за максималните стойности на $\lambda_{max}^{(k)}$ и абсолютната грешка за $PMC_{(MT)}$ и $PMC_{(MS)}$ при различни N и k , когато $\lambda_{max} = 18.14714$

N	k	$\lambda_1^{(k)}$ using $PMC_{(MT)}$	Absolute Error $PMC_{(MT)}$	$\lambda_1^{(k)}$ using $PMC_{(MS)}$	Absolute Error $PMC_{(MS)}$
128	5	17.809541	0.3376	17.766635	0.3805
	6	17.955782	0.1914	17.931153	0.2160
	7	18.031908	0.1152	18.016774	0.1304
	8	18.178382	0.0312	18.184488	0.0373
	9	18.217756	0.0706	18.234182	0.0870
	10	18.259830	0.1127	18.278015	0.1309
	11	18.293246	0.1461	18.336586	0.1894
256	5	17.886432	0.2607	17.880110	0.2670
	6	18.013889	0.1333	18.025125	0.1220
	7	18.098671	0.0485	18.090287	0.0569
	8	18.163375	0.0162	18.172589	0.0254
	9	18.214595	0.0675	18.220052	0.0729
	10	18.265757	0.1186	18.270934	0.1238
	11	18.307841	0.1607	18.317397	0.1703
512	5	17.894326	0.2528	17.893265	0.2539
	6	18.019977	0.1272	18.014737	0.1324
	7	18.099165	0.0480	18.094732	0.0524
	8	18.151116	0.0040	18.152321	0.0052
	9	18.194868	0.0477	18.218433	0.0713
	10	18.231143	0.0840	18.244924	0.0978
	11	18.275571	0.1284	18.284115	0.1370
1024	5	17.977295	0.1698	17.947696	0.1994
	6	18.102537	0.0446	18.038992	0.1081
	7	18.118865	0.0283	18.098850	0.0483
	8	18.145221	0.0019	18.150097	0.0030
	9	18.176356	0.0292	18.214874	0.0677
	10	18.202606	0.0555	18.276008	0.1289
	11	18.250397	0.1033	18.294888	0.1477
2048	5	18.007899	0.1392	17.981049	0.1661
	6	18.106753	0.0404	18.056203	0.0909
	7	18.126731	0.0204	18.104529	0.0426
	8	18.147226	0.0001	18.146851	0.0003
	9	18.164272	0.0171	18.172111	0.0250
	10	18.193708	0.0466	18.215959	0.0688
	11	18.218186	0.0710	18.260127	0.1130

*Маркираните редове представляват оптималните k стойности за всяко N .

изчисли с висока прецизност дялът на обяснената вариация FVE_1 , величина от съществено значение при анализа на систематичния риск, колективната динамика на активите и устойчивостта на диверсификацията в инвестиционни портфейли.

Представените числени резултати в Таблица 3.2 се отнасят за оценките на най-голямата



Фигура 3.1: Сравнение между точната максимална собствена стойност λ_1 и приближена собствена стойност $\lambda_1^{(k)}$ за $PMC_{(MT)}$, $PMC_{(MS)}$, получени в балансиращите случаи при различни стойности на N и $k = 8$

собствена стойност на корелационната матрица \hat{R} при размерност $n = 32$, получени чрез четирите версии на *почти оптималния* СКМК алгоритъм, а именно: $PQMC_{(H)}^{(d)}$, $PQMC_{(S)}^{(d)}$, $PQMC_{(H)}^{(s,l)}$ и $PQMC_{(S)}^{(s,l)}$. Получените резултати с вариантите $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ са при избор $s_1 = 0$ и $l_1 = 0$ в **Алгоритъм 3** и **Алгоритъм 4**, т.е. разбъркани редици на Холтън и Соболю са с настройки *по подразбиране* (англ. *default*).

Числените резултати с алгоритмите $PQMC_{(H)}^{(s,l)}$ и $PQMC_{(S)}^{(s,l)}$ са при избор $s_1 = 1024$ и $l_1 = 128$ в **Алгоритъм 3** и **Алгоритъм 4**, т.е. при генериране на разбърканите редици на Холтън и Соболю са включени операции по пропускане (англ. *skip*) и прескачане през стъпка (анг. *leap*). Максималната собствена стойност $\lambda_1 = \lambda_{\max}$ на корелационната матрица е оценена при различен брой преходи k във веригата на Марков и различен брой реализации N на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$. В Таблица 3.2 са включени и абсолютните грешки на съответните оценки.

Резултатите показват, че броят на преходите k във веригата на Марков и броят на реализациите N на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$ влияят както върху баланса между статистическа и систематична грешка, така върху изчислителната сложност на алгоритмите. При фиксиран брой реализации N , увеличаването на броя преходи k води до намаляване на абсолютната грешка. Това намаляване се стабилизира около $k = 11$ при $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ и около $k = 13$ при $PQMC_{(H)}^{(s,l)}$ и $PQMC_{(S)}^{(s,l)}$.

Този резултат е индикатор за достигане на практически оптимален баланс между систематичната и стохастичната компонента на грешката при съответните стойности на k . По-нататъшното увеличаване на k не води до съществено подобрение на точността. При фиксиран брой преходи k във веригата на Марков и с увеличаването на броя на реали-

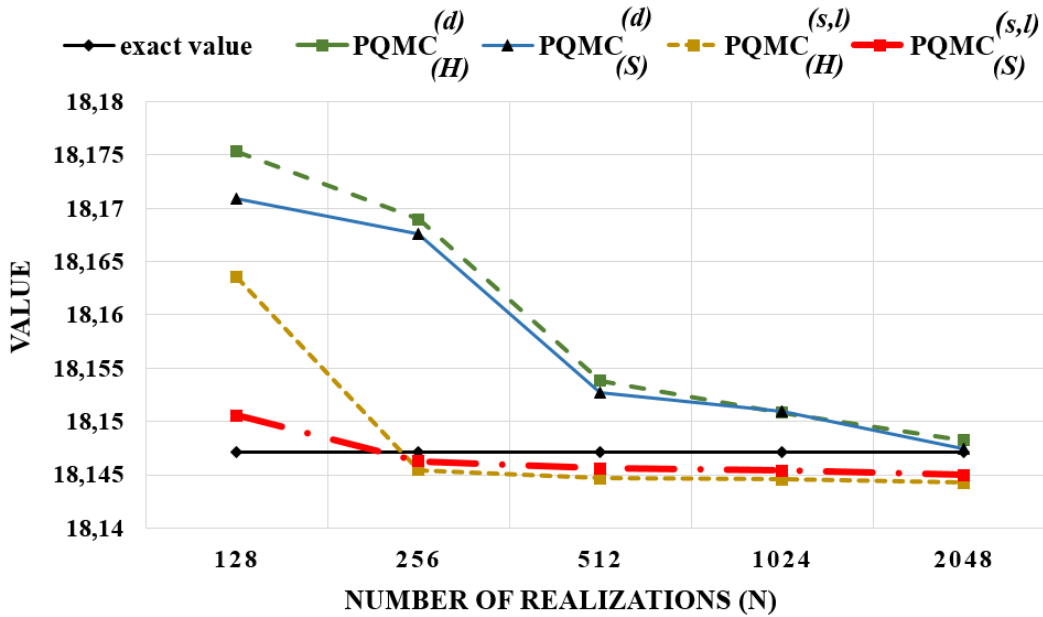
Таблица 3.2: Резултати за максимални собствени стойности за $\lambda_{max}^{(k)}$ и абсолютна грешка чрез сценарий по *подразбиране* и сценарий, използващ *skip* и *leap* при различни N и k , когато $\lambda_{max} = 18.14714$

N	k	$\lambda_1^{(k)}$ using $PQMC_{(H)}^{(d)}$	Absolute Error $PQMC_{(H)}^{(d)}$	$\lambda_1^{(k)}$ using $PQMC_{(S)}^{(d)}$	Absolute Error $PQMC_{(S)}^{(d)}$	$\lambda_1^{(k)}$ using $PQMC_{(H)}^{(s,l)}$	Absolute Error $PQMC_{(H)}^{(s,l)}$	$\lambda_1^{(k)}$ using $PQMC_{(S)}^{(s,l)}$	Absolute Error $PQMC_{(S)}^{(s,l)}$
128	8	17.757547	0.3896	17.761379	0.3858	17.672733	0.4744	17.706637	0.4405
	9	17.888100	0.2590	17.888185	0.2590	17.787523	0.3596	17.792723	0.3544
	10	18.016669	0.1305	18.038214	0.1089	17.856741	0.2904	17.859089	0.2881
	11	18.175307	0.0282	18.170859	0.0237	17.903166	0.2440	17.914975	0.2322
	12	18.298811	0.1517	18.264832	0.1177	18.036970	0.1102	18.097564	0.0496
	13	18.479146	0.3320	18.474626	0.3275	18.145462	0.0017	18.146253	0.0009
	14	18.643362	0.4962	18.639031	0.4919	18.194240	0.0471	18.191902	0.0448
	15	18.712889	0.5657	18.706247	0.5591	18.290732	0.1436	18.290924	0.1438
256	8	17.882593	0.2645	17.883812	0.2633	17.761315	0.3858	17.763949	0.3832
	9	17.994408	0.1527	18.016913	0.1302	17.854804	0.2923	17.867712	0.2794
	10	18.061271	0.0859	18.062487	0.0847	17.936785	0.2104	17.921160	0.2260
	11	18.168944	0.0218	18.167556	0.0204	18.036010	0.1111	18.043959	0.1032
	12	18.284126	0.1370	18.251619	0.1045	18.097978	0.0492	18.098070	0.0491
	13	18.482373	0.3352	18.469322	0.3222	18.145462	0.0017	18.146253	0.0009
	14	18.545036	0.3979	18.516617	0.3695	18.209384	0.0622	18.185109	0.0380
	15	18.608838	0.4617	18.607262	0.4601	18.289967	0.1428	18.270105	0.1230
512	8	17.927397	0.2197	17.962654	0.1845	17.774098	0.3730	17.778972	0.3682
	9	18.020330	0.1268	18.021557	0.1256	17.846070	0.3011	17.867473	0.2797
	10	18.081105	0.0660	18.081198	0.0659	17.938077	0.2091	17.949995	0.1971
	11	18.153831	0.0067	18.152703	0.0056	18.009669	0.1375	18.014807	0.1323
	12	18.263301	0.1162	18.248667	0.1015	18.038078	0.1091	18.038550	0.1086
	13	18.425061	0.2779	18.411903	0.2648	18.144697	0.0024	18.145602	0.0015
	14	18.506259	0.3591	18.504204	0.3571	18.204814	0.0577	18.191490	0.0444
	15	18.551586	0.4044	18.546687	0.3995	18.249184	0.1020	18.242947	0.0958
1024	8	17.962214	0.1849	17.978111	0.1690	17.776881	0.3703	17.792702	0.3544
	9	18.035437	0.1117	18.042592	0.1045	17.867701	0.2794	17.887652	0.2595
	10	18.103422	0.0437	18.109582	0.0376	17.980078	0.1671	17.980502	0.1666
	11	18.150822	0.0037	18.150986	0.0038	18.010761	0.1364	18.037651	0.1095
	12	18.294487	0.1473	18.284447	0.1373	18.098692	0.0484	18.098718	0.0484
	13	18.385875	0.2387	18.384413	0.2373	18.144277	0.0029	18.144964	0.0022
	14	18.446470	0.2993	18.443825	0.2967	18.202841	0.0557	18.203207	0.0561
	15	18.530500	0.3834	18.521017	0.3739	18.236167	0.0890	18.235788	0.0886
2048	8	18.000431	0.1467	18.000619	0.1465	17.787254	0.3599	17.788274	0.3589
	9	18.061446	0.0857	18.064399	0.0827	17.879783	0.2674	17.887915	0.2592
	10	18.108451	0.0387	18.110469	0.0367	17.980078	0.1671	17.980502	0.1666
	11	18.148212	0.0011	18.147480	0.0003	18.010923	0.1362	18.040106	0.1070
	12	18.226453	0.0793	18.221684	0.0745	18.118819	0.0283	18.120939	0.0262
	13	18.385443	0.2383	18.379872	0.2327	18.144277	0.0029	18.144964	0.0022
	14	18.443228	0.2961	18.435573	0.2884	18.202381	0.0552	18.203555	0.0564
	15	18.509033	0.3619	18.492101	0.3450	18.247648	0.1005	18.243565	0.0964

*Маркираните редове представляват оптималните k стойности за всяко N .

защите N се наблюдава монотонно намаляване на абсолютната грешка при $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$.

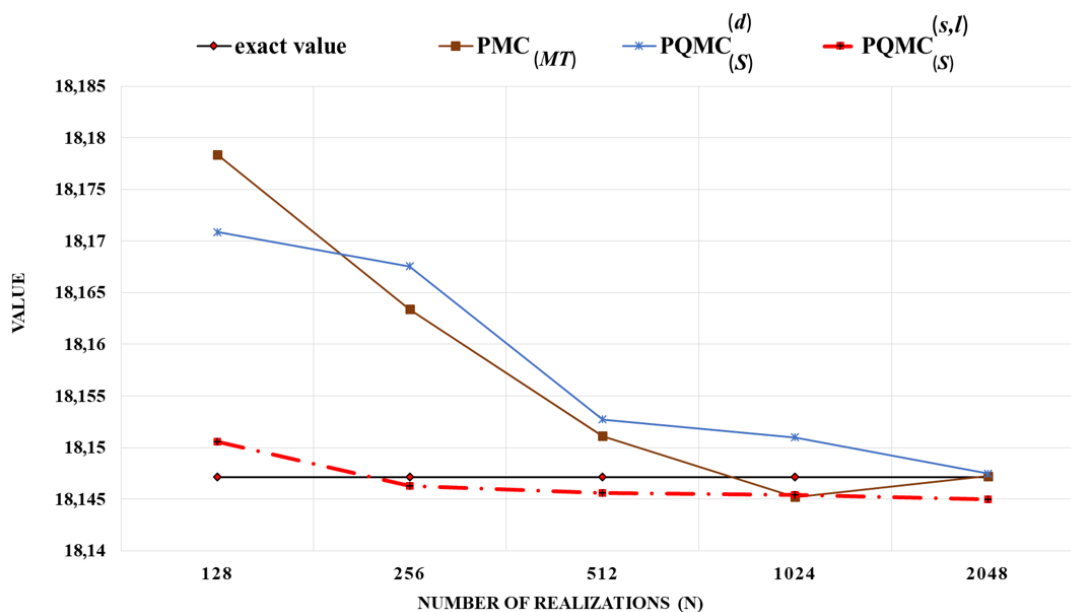
От друга страна, при алгоритмите $PQMC_{(H)}^{(s,l)}$ и $PQMC_{(S)}^{(s,l)}$ се наблюдава силно намаляване на абсолютната грешка при малки стойности на N , а при по-нататъшно увеличаване на N , грешката остава от същия порядък или дори нараства. Този проблем е коментирани в [62], като е препоръчано изборът на параметрите *skip* и *leap* да се прави внимателно за разглежданите задачи. В Таблица 3.2 минималните абсолютни грешки и съответните решения за всеки от алгоритмите $PQMC_{(H)}^{(d)}$, $PQMC_{(S)}^{(d)}$, $PQMC_{(H)}^{(s,l)}$ и $PQMC_{(S)}^{(s,l)}$ са маркирани с удебелен шрифт.



Фигура 3.2: $(PQMC_{(H)}^{(d)}, PQMC_{(S)}^{(d)})$ и $(PQMC_{(H)}^{(s,l)}, PQMC_{(S)}^{(s,l)})$. Сравнение между максималната собствена стойност λ_{\max} на корелационната матрица и оценените стойности $\lambda_1^{(k)}$, получени при $k = 11$ в случая по *подразбиране* и $k = 13$ в случая с пропускане и прескачане и различни стойности на N .

Направените обобщения на четирите варианта на *почти оптималните* СКМК алгоритми ясно се демонстрират на Фигура 3.2, където са сравнени приближените оценки на $\lambda_1^{(k)}$ с точното решение λ_1 при нарастващ брой реализации N . Графично са илюстрирани резултатите за $PQMC_{(S)}^{(d)}$, $PQMC_{(H)}^{(d)}$, $PQMC_{(S)}^{(s,l)}$ и $PQMC_{(H)}^{(s,l)}$ при балансиращите стойности за $k = 11$ във вариантите по *подразбиране* и $k = 13$ при използване на параметрите за пропускане и прескачане. Графиките при алгоритмите $PQMC_{(S)}^{(d)}$ и $PQMC_{(H)}^{(d)}$ показват, че при увеличаване на N се наблюдава сходимост на приближеното решение към точната стойност λ_1 .

При алгоритмите $PQMC_{(H)}^{(s,l)}$ и $PQMC_{(S)}^{(s,l)}$ се наблюдава намаляване на абсолютната грешка при $N = 128$ и $N = 256$, а при по-нататъшно увеличаване на N , грешката остава почти непроменена, дори леко нараства независимо, че при $N = 512$ и $N = 1024$ продължава да е по-малка от вариантите по *подразбиране*. На Фигура 3.3 се демонстрира, че алгоритмите $PMC_{(MT)}$ и $PQMC_{(S)}^{(d)}$ показват устойчива сходимост към λ_{\max} и по-висока



Фигура 3.3: Сравнение между максимална собствена стойност λ_{\max} и приближените собствени стойности $\lambda_1^{(k)}$, получени при $k = 8$ за $PMC_{(MT)}$, $k = 11$ за $PQMC_{(S)}^{(d)}$ и $k = 13$ за $PQMC_{(S)}^{(s,l)}$ при различни стойности на N .

Таблица 3.3: Приблизени стойности за дяла на обяснената вариация (FVE_1), получени при $k = 8$ за $PMC_{(MT)}$, $k = 11$ за $PQMC_{(S)}^{(d)}$ и $k = 13$ за $PQMC_{(S)}^{(s,l)}$ за корелационната матрица \hat{R} с размерност $n = 32$. Референтната стойност е $FVE_1 = 56.71\%$.

N	$PMC_{(MT)}$	$PQMC_{(S)}^{(d)}$	$PQMC_{(S)}^{(s,l)}$
256	56.76%	56.77%	56.70%
2048	56.71%	56.71%	56.70%

точност при голям брой симулации N в сравнение с алгоритъм $PQMC_{(S)}^{(s,l)}$. При малък брой симулации $N \leq 512$, СКМК алгоритми с използване на параметрите *skip* и *leap* са препоръчителни. Тези резултати потвърждават, че *почти оптималните* стохастични алгоритми представляват надежден инструмент за оценяване на максималната собствена стойност на корелационна матрица с неголяма размерност. Получените приближени оценки за λ_1 позволяват да се изчисли с висока прецизност дялът на обяснената вариация FVE_1 , използвайки формула (3.12).

В Таблица 3.3 са представени приближени стойности за FVE_1 , като са използвани оценките за $\lambda_1^{(k)}$, получени чрез $PMC_{(MT)}$, $PQMC_{(S)}^{(d)}$ и $PQMC_{(S)}^{(s,l)}$ с балансиращите стойности за k и $N = 256$ и $N = 2048$. Референтната стойност за дяла на обяснената вариация за корелационната матрица \hat{R} с размерност $n = 32$ е $FVE_1 = 56.71\%$. Резултатите показват, че при по-малка извадка N използването на $PQMC_{(S)}^{(s,l)}$ за намиране FVE_1 е за предпочитане. При по-големи стойности на N се препоръчва да се използват $PMC_{(MT)}$ и $PQMC_{(S)}^{(d)}$ за оценяване на дяла на обяснената вариация за корелационната матрица, като по-точни и с по-малка изчислителна сложност сравнение с *почти оптималните* СКМК алгоритми с параметри *skip* и *leap*.

В допълнение, стойността на FVE_1 за корелационната матрица \hat{R} с размерност $n = 32$ показва, че е налице факторна концентрация в първия компонент. За инвеститора това означава, че диверсификацията може да бъде по-ограничена, защото активите реагират сходно на доминиращи шокове. От гледна точка на мениджмънта, това е индикатор за повишена системност на риска.

3.4.2 Случай (ii): корелационна матрица с размерност $n = 490$

В Таблица 3.4 са представени приближените оценки за най-голямата собствена стойност на корелационната матрица \hat{R} при размерност $n = 490$. Матрицата е изградена на базата на доходите на $n = 490$ компании, включени в индекса $S\&P500$, оценяващ състоянието на американския фондов пазар за петгодишния период от 2021г. до 2025 г. (виж Приложение). За десет компании липсваха пълни данни за този период и поради тази причина бяха изключени при конструирането на корелационната матрица.

Таблица 3.4: Резултати за максимални собствени стойности за $\lambda_{max}^{(k)}$ и абсолютна грешка при $PMC_{(MT)}$, $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ при различни N и k , когато $\lambda_{max} = 201.090512$.

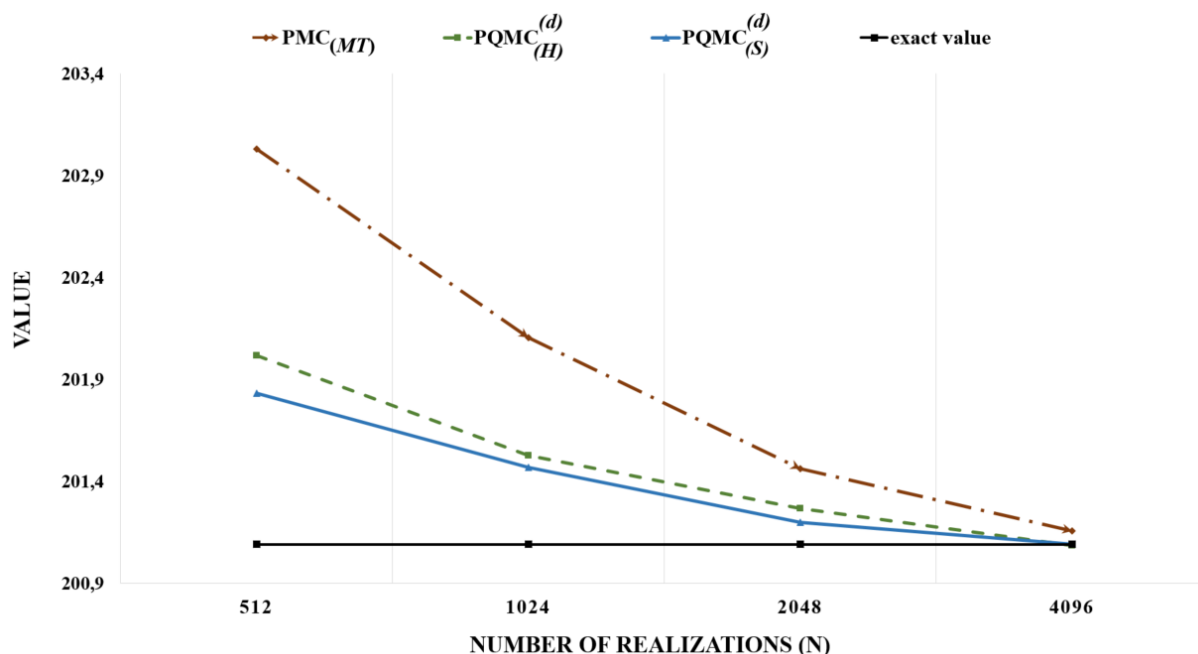
N	k	$\lambda_1^{(k)}$ using $PMC_{(MT)}$	Absolute Error $PMC_{(MT)}$	k	$\lambda_1^{(k)}$ using $PQMC_{(H)}^{(d)}$	Absolute Error $PQMC_{(H)}^{(d)}$	$\lambda_1^{(k)}$ using $PQMC_{(S)}^{(d)}$	Absolute Error $PQMC_{(S)}^{(d)}$
512	6	191.261222	9.8293	8	192.052259	9.0383	192.056796	9.0337
	7	197.764391	3.3261	9	193.305250	7.7853	194.4876692	6.6028
	8	203.030089	1.9396	10	196.545330	4.5452	196.561636	4.5289
	9	206.001210	4.9107	11	202.016871	0.9264	201.830255	0.7397
	10	209.202619	8.1121	12	207.216480	6.1260	206.994716	5.9042
	1024	6	194.244715	6.8458	8	191.604757	9.4858	191.662089
7		199.671479	1.4190	9	194.355045	6.7355	194.443329	6.6472
8		202.103664	1.0132	10	198.362863	2.7276	196.623242	4.4673
9		204.706063	3.6156	11	201.526659	0.4361	201.467671	0.3772
10		211.940453	10.8499	12	204.493635	3.4031	204.718275	3.6278
2048		6	196.531375	4.5591	8	191.522259	9.5683	191.737033
	7	199.092487	1.9980	9	194.216871	6.8736	194.383618	6.7069
	8	201.462168	0.3717	10	198.552479	2.5380	198.866204	2.2243
	9	204.787620	3.6971	11	201.267671	0.1772	201.199260	0.1087
	10	206.264686	5.1942	12	207.590516	6.5000	206.670141	5.5796
	4096	6	195.762474	5.3280	8	194.293635	6.7969	195.846272
7		197.402088	3.6884	9	198.498616	2.5919	198.498988	2.5915
8		201.154640	0.0641	10	199.459470	1.6310	199.802438	1.2881
9		202.887637	1.7971	11	201.083321	0.0072	201.091681	0.0012
10		207.343376	6.2529	12	206.448996	5.3585	205.624451	4.5339

*Маркираните редове представляват оптималните k стойности за всяко N .

Алгоритмите, които са използвани за оценяване на максималната собствена стойност на корелационната матрица са: (1) $PMC_{(MT)}$ с използван *Mersenne Twister* (MT) генератор на псевдослучайни числа и (2) вариантите на *почти оптималния* СКМК алгоритъм

$PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$, използващи разбъркани редици на Холтгън и Соболев с настройки *по подразбиране*. Максималната собствена стойност $\lambda_1 = \lambda_{\max}$ е оценявана отново при различен брой преходи k във веригата на Марков и различен брой реализации N на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$. В Таблица 3.4 са включени и абсолютните грешки на съответните оценки.

Резултатите показват, че броят на преходите k във веригата на Марков и броят на реализациите N на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$ отново влияят върху баланса между статистическа и систематична грешка. Независимо, че корелационната матрица с размерност $n = 490$ е с порядък по-голям от матрицата с размерност $n = 32$, балансиращата стойност при броя на преходите във веригата на Марков за $PMC_{(MT)}$ е $k = 8$, а за $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ е $k = 11$. С увеличаването на броя реализации N , при фиксиран брой преходи k във веригата на Марков, абсолютната грешка намалява монотонно и при трите алгоритъма. В Таблица 3.4 минималните абсолютни грешки и съответните решения за всеки от алгоритмите $PMC_{(MT)}$, $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ са маркирани с удебелен шрифт.



Фигура 3.4: Сравнение между максимална собствена стойност λ_{\max} и оценените собствени стойности $\lambda_1^{(k)}$, получени при $k = 8$ за $PMC_{(MT)}$ и $k = 11$ за $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ за различни стойности на N

На Фигура 3.4 са сравнени приблизителните оценки $\lambda_1^{(k)}$ получени за $PMC_{(MT)}$, $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ с референтната стойност λ_{\max} за различен брой реализации N (по оста X) и при балансиращите стойности за k (по оста Y). При графичното представяне, трите алгоритъма демонстрират устойчива сходимост към λ_{\max} и висока точност при голям брой симулации N . Прави впечатление, че сходимостта при версиите с *почти оптималните* СКМК алгоритми е по-добра отколкото с *почти оптималния* СК алгоритъм.

Получените приближени оценки за λ_1 с $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ позволяват да се изчисли с по-висока прецизност дялът на обяснената вариация FVE_1 в сравнение с $PMC_{(MT)}$. Резултатите за приближени стойности на FVE_1 за корелационната матрица \hat{R} с размерност

Таблица 3.5: Приблизени стойности за дяла на обяснената вариация (FVE_1), получени при $k = 8$ за $PMC_{(MT)}$ и при $k = 11$ за $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ за корелационната матрица \hat{R} с размерност $n = 490$. Референтната стойност е $FVE_1 = 41.04\%$.

N	$PMC_{(MT)}$	$PQMC_{(H)}^{(d)}$	$PQMC_{(S)}^{(d)}$
1024	41.25%	41.13%	41.12%
2048	41.11%	41.08%	41.06%
4096	41.05%	41.04%	41.04%

$n = 490$ са представени в Таблица 3.5. За тяхното пресмятане са използвани приближените оценки за $\lambda_1^{(k)}$ получени чрез $PMC_{(MT)}$, $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ с балансиращи стойности $k = 8$ и $k = 11$. Референтната стойност за дяла на обяснената вариация е $FVE_1 = 41.04\%$. Резултатите в таблицата потвърждават предположението, че при увеличаване на симулациите N използването на $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ са за предпочитане пред $PMC_{(MT)}$ за изчисляване на FVE_1 .

В допълнение, стойността на FVE_1 за корелационната матрица \hat{R} с размерност $n = 490$ е около 41%. Този факт показва, че вариацията се разпределя между множество компоненти (т.е. не е съсредоточена само в първата компонента), което предполага по-сложна и по-разнообразна структура на зависимостите. Това обикновено се свързва с по-силна потенциална диверсификация, но и с необходимост от повече фактори при моделиране на риска (например, многофакторни модели вместо еднофакторни).

3.5 Обобщение

Получените резултати в настоящата глава потвърждават както практическата приложимост, така и числената ефективност на разработените *почти оптимални* стохастични алгоритми за оценяване на максималната собствена стойност на корелационни матрици в задачи, свързани с измерването на пазарния риск на инвестиционни портфейли. Изследваните матрици са конструирани въз основа на финансови данни, извлечени от два основни източника: (1) уеб платформата *Portfolio Visualizer* и (2) базата данни на *Yahoo Finance*, (виж Приложение).

Анализът на двете разгледани корелационни матрици с размерности съответно $n = 32$ и $n = 490$ показва, че максималната собствена стойност притежава ясна интерпретация като индикатор за наличието и интензитета на доминиращ общ пазарен фактор. Чрез дяла на обяснената вариация FVE_1 може да се получи количествена оценка както за степента на систематичен риск, така и за ограниченията пред диверсификацията.

Числените експерименти показват още, че при *почти оптималните* СКМ алгоритми оптималният баланс между стохастичната и систематична грешка се достига при $k = 8$. При *почти оптималните* СКМК алгоритми, използващи разбъркани редици на Собол и Холтън с настройки по *подразбиране*, този баланс се достига при $k = 11$. При включването на операторите *skip* и *leap* съответната стойност нараства до $k = 13$. Резултатите също така показват, че с увеличаване на броя реализации N абсолютната грешка при *почти оптималните* СКМ алгоритми и при *почти оптималните* СКМК алгоритми с настройки

по *подразбиране* намалява монотонно.

Съществен резултат е, че при оценяване на максималната собствена стойност на матрицата с размерност $n = 490$ *почти оптималните* СКМК алгоритми с настройки по *подразбиране* демонстрират по-добра сходимост и по-висока прецизност в сравнение с *почти оптималните* СК алгоритми. От друга страна, използването на операторите *skip* и *leap* е оправдано при по-малък брой симулации, тъй като при по-големи стойности на N тяхното предимство отслабва и не винаги води до допълнително намаляване на грешката. Част от представените резултати в тази глава са публикувани в [44].

В обобщение, разработените стохастични степенни алгоритми могат да бъдат разглеждани като надежден подход за числено оценяване на спектрални характеристики на корелационни матрици. Това от своя страна позволява задълбочен анализ на систематичния риск, колективната динамика на активите и ефективността на диверсификацията. В бъдещи изследвания тези алгоритми могат да бъдат прилагани в задачи, свързани с финансово моделиране, управление на портфейли и подпомагане на инвестиционни решения в условия на нарастваща размерност.

Заклучение

Настоящата дисертация е посветена на разработването, анализа и приложението на стохастични числени методи от типа МК и рандомизирани КМК методи за оценяване на екстремални собствени стойности на симетрични матрици и използването им при практически задачи, свързани с оценка на риска на инвестиционни портфейли. Основната цел е изграждането на стабилна и изчислително ефективна методологична рамка, която осигурява управляем баланс между систематичната и стохастичната грешка и е приложима при задачи с висока размерност.

В рамките на изследването са обосновани и изследвани Степенният (Резолвентният) Монте Карло метод и рандомизираният Степенен (Резолвентен) Квази-Монте Карло метод за оценяване на екстремални собствени стойности. Съществен резултат е прецизирането на конструкцията на случайните величини при използването на почти оптимална преходна плътност във веригата на Марков, както и идентифицирането на класове симетрични матрици, при които тази конструкция се опростява. На тази основа са разработени почти оптимални алгоритми (Степенен и Резолвентен МК и КМК), описани чрез псевдокод, за които са дадени оценки за изчислителната сложност и условия за балансиране на статистическата и систематичната грешка. Теоретичните аргументи и експерименталните резултати потвърждават, че предложените почти оптимални конструкции водят до намаляване на дисперсията и повишаване на точността при същия или по-малък изчислителен разход спрямо класическите стохастични аналози.

Експерименталната верификация е проведена както с генератори на псевдослучайни числа (MT и MS), така и с редици с малък дискрепанс (редици на Соболев и Холтън) и техни разбъркани варианти. Числените изследвания показват измерим ефект от избора на генератор и типа редица върху точността и стабилността на оценките, като при разглежданите тестове генераторът на псевдослучайни числа MT демонстрира предимство спрямо генератора MS, а редиците на Соболев спрямо редиците на Холтън. Установено е също, че за постигане на сравним баланс на грешките рандомизираният Степенен (Резолвентен) КМК алгоритъм може да изисква малко по-голям брой преходи във веригата на Марков, което се компенсира от по-добрата сходимост при ограничени изчислителни ресурси.

Разработените алгоритми са приложени към задача от финансовата математика, свързана с управлението на риска на инвестиционен портфейл. Максималната собствена стойност на корелационната матрица на активи е използвана като индикатор за концентрация на риска, като оценяването ѝ е реализирано чрез предложените СМК и СКМК алгоритми. При реализиране на СКМК алгоритмите с разбъркани редици с малък дискрепанс

(редици на Соболев и Холтър) са изследвани варианти и с използване на операторите *skip* и *leap*.

В обобщение, дисертационният труд демонстрира, че прилагането на разработените алгоритми за получаване на спектрални оценки с целенасочено балансиране на грешките води до по-висока точност, стабилност и изчислителна ефективност. Очертани са и перспективи за бъдещо развитие, включително разширяване към устойчиви (робастни) модели, динамични ковариационни структури и допълнителни изчислителни оптимизации за разредени симетрични матрици с големи размерности от порядъка на милиони.

Списък на публикациите по дисертацията

За написването на дисертацията са използвани четири публикации, от които:

- Една е публикувана в списание *Contemporary Mathematics*, която е видима в **Web of Science (WoS): Q1, IF:2.5 (2025)**, <https://wos-journal.info/journalid/23187>);
- Една е публикувана в *Lecture Notes in Computer Science (LNCS)*);
- Една е публикувана като научно съобщение в национално академично издателство;
- Една е приета за публикуване в *Lecture Notes in Computer Science (LNCS)*.

Основна част от представените резултати в дисертационния труд са публикувани в издания, индексирани в *Scopus* и *Web of Science*.

1. **Gurova, S.-M.**, Karaivanova, A., *Quasi-Monte Carlo Algorithms for Eigenvalue Problems, 4th International Conference Numerical Methods for Scientific Computations and Advanced Applications, NMSCAA'24, 18-22, (2024)* (short communication)
<https://parallel.bas.bg/~elilkova/nmscaa24.pdf>
2. **Gurova, S.-M.**, Atanassov, E., Karaivanova, A., *A Resolvent Quasi-Monte Carlo Method for Estimating the Minimum Eigenvalues Using the Error Balancing*, In: Lirkov, I., Margenov, S. (eds), *Large-Scale Scientific Computing, LSSC'23, LNCS*, vol. **13952**, 394-403, (2024), (SJR:0.352, Q2, 2024)
https://doi.org/10.1007/978-3-031-56208-2_40
3. **Gurova, S.-M.**, Gurov, T., Karaivanova, A., *On the Stochastic Power Algorithms for Estimating the Maximum Eigenvalue of Symmetric Matrices*, **Contemp. Math.**, Universal Wiser Publisher Pte. Ltd, **6(5):7223-7249,(2025)**, doi.org/10.37256/cm.6520257781 (**IF:2.5, 2025, WoS: Q1**, <https://wos-journal.info/journalid/23187>)
4. **Gurova, S.-M.**, Gurov, T., Karaivanova, A., *Eigenvalue Estimation in Portfolio Risk: The Role of Skipping and Leaping in Sobol and Halton Sequences*, In: Lirkov, I., Margenov, S. (eds), *Large-Scale Scientific Computing, LSSC'25, LNCS*, vol. **16061**, (accepted)

Апробация на резултатите

Резултатите, които са включени в дисертацията, са докладвани на семинари на секцията "Скалируеми алгоритми и киберсигурност с Център по високопроизводителни пресмятания" към ИИКТ-БАН, както и на международни конференции. Изнесени са следните доклади, които са свързани с темата на дисертацията:

1. Gurova, S.-M., Monte Carlo Method for Estimating Eigenvalues Using Error Balancing, The 13th International Conference on Large-Scale Scientific Computations, LSSC'21, 07-11.06.2021, Sozopol, Bulgaria. <https://parallel.bas.bg/Conferences/SciCom21/program.html>
2. Gurova, S.-M., A Quasi-Monte Carlo method for estimation of eigenvalues using error balancing, The 15th International Conference on Monte Carlo and Quasi-Monte Carlo methods in Scientific Computing, MCQMC'22, 17-22.07.2022, Linz, Austria. https://www.ricam.oeaw.ac.at/events/conferences/mcqm2022/schedule/MCQMC2022_book_final_version.pdf
3. Gurova, S.-M., Power Monte Carlo method using randomized low discrepancy sequences, the 17th Annual Meeting of the Bulgarian Section of SIAM, 20-22.12.2022, BGSIAM'22, Sofia, Bulgaria. http://www.math.bas.bg/bgsiam/docs/bgsiam_2022_program.pdf
4. Gurova, S.-M., A Resolvent Quasi-Monte Carlo Method for Estimating the Minimum Eigenvalues Using the Error Balancing, The 14th International Conference on Large-Scale Scientific Computations, LSSC'23, 05-09.06.2023, Sozopol, Bulgaria. <https://parallel.bas.bg/~elilkova/lssc23.pdf#page=17>
5. Gurova, S.-M., Quasi-Monte Carlo Algorithms for Eigenvalue Problems, Fourth International Conference Numerical Methods for Scientific Computations and Advanced Applications, NMSCAA'24, 17-21.06.2024, Sozopol, Bulgaria. <https://parallel.bas.bg/Conferences/nmscaa24/program.html>

Статиите представят резултати, получени в рамките на един международен и три национални научни проекта, като единият от тях е разделен на две фази, а именно:

1. Международен научен проект

- Нови скалируеми алгоритми с използване на Монте Карло и Квази-Монте Карло методи и тяхното приложение в областта на машинното обучение, финансиран от CAF America, 2021-2024, ръководител: проф. д-р Емануил Атанасов.

2. Национални научни проекти

- "Център за върхови постижения по Информатика и информационни и комуникационни технологии" (фаза 1) по договор BG05M2OP001-1.001-0003, финансиран от оперативна програма "Наука и образование за интелигентен растеж" 2014-2020, съфинансиран от Европейския съюз чрез Европейските структурни и инвестиционни фондове, ръководител: чл.-кор. д.м.н. Светозар Маргенов.

- ”Национален геоинформационен център (НГИЦ) за мониторинг, оценка и прогнозиране на природни и антропогенни рискове и бедствия” обект от ”Националната пътна карта за научна инфраструктура 2020-2027” (договори: Д01-321/30.11.2023, Д01-164/28.07.2022), финансиран от Министерство на образованието и науката, ръководител: доц. д-р Петя Трифонова.
- ”Национален център за високопроизводителни и разпределени пресмятания” (НЦВРП) обект от ”Националната пътна карта за научна инфраструктура 2020-2027” (договори: Д01-98/26.06.2025, Д01-325/01.12.2023), финансиран от Министерство на образованието и науката, ръководител: проф. д-р Анета Караиванова.
- ”Център за върхови постижения в областта на информатиката и информационните и комуникационни технологии” (фаза 2) по договор BG16RFPR002-1.014-0018-C01, финансиран по програмата за научни изследвания, иновации и дигитализация за интелигентна трансформация 2021–2027 и съфинансиран от Европейския съюз, ръководител: чл.-кор. д.м.н. Светозар Маргенов.

Основни научни и научно-приложни приноси

Получените научни и научно-приложни приноси напълно съответстват на поставените цели и задачи на дисертационния труд. Те пряко отразяват обосноваването, разработването и експерименталното изследване на Степенен МК/КМК алгоритми и Резолвентен МК/КМК алгоритми за оценка на екстремни собствени стойности на симетрични матрици. Освен това, приносите са потвърдени и в реална задача от финансовата математика чрез оценяване на пазарния риск на портфейл от глобални активи.

1. Основните научни приноси на настоящата дисертация са:

- Обосновани и изследвани са Степенният (Резолвентният) Монте Карло метод и рандомизираният Степенен (Резолвентен) Квази-Монте Карло метод за оценяване на екстремални собствени стойности на симетрични матрици. Прецизирана е конструкцията на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$ при използване на почти оптимална преходна плътност във веригата на Марков. Формулирани са класове симетрични матрици, за които се опростява конструкцията на случайните величини в стохастичните степенни алгоритми. Получени са теоретични оценки за зависимостта между степента на резолвентната матрица и дължината на веригата на Марков в зависимост от ускоряващия параметър и нормата на матрицата.
- Разработени са почти оптимални СМК и СКМК алгоритми, както и почти оптимални РМК и РКМК алгоритми, базирани на предложената прецизирана конструкция на случайните величини, които са представени чрез псевдокод. Дадени са оценки за изчислителната сложност на алгоритмите и условията за

баланс между стохастичната и систематична грешка. Аргументирано е, че предложените алгоритми водят до намаляване на дисперсията спрямо класическите аналози.

- Предложена е практическа задача от областта на финансовата математика за оценяване на пазарния риск на инвестиционни портфейли. Конструираната корелационна матрица за съответния портфейл е използвана за оценяването на максималната собствена стойност, която е индикатор за концентрация на пазарния риск. Оценяването се извършва чрез разработените в Глава 1 почти оптимален Степенен МК и Степенен КМК алгоритми, като допълнително Степенният КМК алгоритъм е модифициран чрез използване на параметрите *skip* и *leap*.

2. Основните научно-приложни приноси на настоящата дисертация са:

- Предложените почти оптимални СМК алгоритми са реализирани с МТ и MS генератори на псевдослучайни числа, а СКМК алгоритмите, с разбъркани редици на Соболев и Холтън. Числените експерименти за оценяване на максималната собствена стойност на плътни симетрични матрици с големи размерности демонстрират превъзходството на почти оптималните алгоритми над класическите им аналози както по точност, така и по изчислителна сложност.
- Числено е изследван балансът между стохастичната и систематичната грешка, който зависи съществено от броя на преходите k във веригата на Марков и от броя реализации N на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$. Резултатите показват, че СКМК алгоритмите се нуждаят от 2-3 стъпки повече във веригата на Марков в сравнение със СМК алгоритмите. Изборът на генератор на псевдослучайни числа или на разбърканата редица с малък дискрепанс има измерим ефект върху точността и стабилността. В разгледаните експерименти МТ е предпочитан пред MS, а редиците на Соболев пред редиците на Холтън.
- Почти оптималният РМК алгоритъм е реализиран с МТ генератор на псевдослучайни числа, а почти оптималният РКМК алгоритъм с разбъркана редица на Соболев чрез BRODA's Sobol RSG генератор. Числените експерименти, проведени на НРС клъстер с NVIDIA Tesla V100 (32 GB) GPU карти, за оценяване на минималната собствена стойност, показват, че балансът между стохастичната и систематичната грешка зависи от броя на преходите k във веригата на Марков, броя реализации N на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$, степента m на резолвентната матрица и параметъра за ускорение q . Резултатите, получени с рандомизирана редица на Соболев и с МТ генератор за оценяване на минималната собствена стойност на симетрични матрици, демонстрират един и същ порядък. От съществено значение е възможността да се намери константа, която отделя минималната собствена стойност от нулата.
- Проведени са числени експерименти за оценяване на максималната собствена стойност на две корелационни матрици, която е индикатор за концентрация на

пазарния риск. Първата корелационна матрица е конструирана въз основа на доходите на 32 глобални актива, докато при втората са използвани доходите на 490 компании от *S&P500*, за конкретни времеви периоди (виж Приложение). Оценяването е извършено чрез почти оптималния СКМ алгоритъм с МТ/MS генератори на псевдослучайни числа и чрез почти оптималния СКМК алгоритъм, както без допълнителни настройки (*default* вариант), така и с използване на параметрите *skip* и *leap* при разбъркани редици на Соболев и Холтън.

Декларация за оригиналност

Декларирам, че настоящият дисертационен труд съдържа оригинални резултати, получени при проведени от мен научни изследвания с подкрепата и съдействието на научния ръководител и съавторите. Резултатите, които са получени, описани и/или публикувани от други учени, са надлежно и подробно цитирани в библиографията.

Настоящата дисертация не е прилагана за придобиване на научна степен в друго висше училище, университет или научен институт.

Подпис:

Благодарности

*"Умът е не само в знанието, но и
умението то да се прилага."*

Аристотел

Изказвам най-искрена благодарност към моя научен ръководител и към колегите от Института по информационни и комуникационни технологии - БАН (ИИКТ-БАН), които създадоха среда, в която научното любопитство и колегиалността вървят ръка за ръка. Техните насоки, критични бележки и ежедневна подкрепа превърнаха този труд от идеен проект в завършена дисертация. Особено признателна съм на екипа на секцията „Скалируеми алгоритми и киберсигурност с Център по високопроизводителни пресмятания“ за споделените знания и за конструктивните дискусии по време на семинарите ни. Именно там се родиха голяма част от идеите, прераснали в статии и в разработените в тази работа алгоритми. Благодарности и към съавторите и рецензентите по публикациите, свързани с дисертацията. Техните бележки и въпроси помогнаха да се изчистят детайлите и да се структурират по-ясно методологичните избори и изводи, които да очертаят мястото на приноса в съвременната литература.

Сърдечни благодарности към моето семейство и приятели за търпението, разбирането и вярата в мен през целия период на докторантурата. Вашето търпение даде време на идеите да узреят, а вашата вяра - силата да бъдат доведени докрай. Този труд е толкова и техен, колкото и мой, защото ако знанието е мощ, както казва Аристотел, то приложението му е дело на общност.

Приложение

За провеждането на числените експерименти са използвани генерирани симетрични матрици (Глава 1, Глава 2) и реални финансови данни (Глава 3). Симетричните матрици с размерности 100 и 500, използвани за числените експерименти в Глава 1, са генерирани в софтуерната среда МАТЛАВ, докато тези с размерности 500 и 1000 от Глава 2 са част от алгоритмичната реализация, разработена на програмния език С++.

Финансовите данни, използвани за конструирането на корелационната матрица с размерност $n = 32$ в Глава 3, са извлечени от уеб платформата *Portfolio Visualizer* (www.portfoliovisualizer.com). Изследваният времеви обхват обхваща периода от януари 2007 г. до март 2024 г., като са използвани реални пазарни показатели. Въз основа на тази емпирична информация е конструиран инвестиционен портфейл. Портфейлът се състои от тридесет и два глобални актива, класифицирани според техния инвестиционен тип, както следва:

1. **Осемнадесет акции** - „US Stock Market“, „US Large Cap Value“, „US Large Cap Growth“, „US Mid Cap“, „US Mid Cap Value“, „US Mid Cap Growth“, „US Small Cap Value“, „US Small Cap Growth“, „US Large Cap“, „US Small Cap“, „US Micro Cap“, „Global ex-US Stock Market“, „European Stocks“, „Pacific Stocks“, „Emerging Markets“, „Intl Developed ex-US Market“, „International ex-US Small Cap“, „International ex-US Value“;
2. **Девет облигации** - „Short Term Treasury“, „Intermediate Term Treasury“, „10-year Treasury“, „Long Term Treasury“, „Total US Bond Market“, „Global Bonds (USD Hedged)“, „Short-Term Investment Grade“, „TIPS (Treasury Inflation-Protected Securities)“, „Corporate Bonds“;
3. **Четири алтернативни инвестиции** - „REIT (Real Estate Investment Trusts)“, „Gold“, „Precious Metals“, „Commodities“;
4. **Един паричен актив и парични еквиваленти** - „Cash“.

Финансовите данни, използвани за конструирането на корелационната матрица с размерност $n = 490$ в Глава 3, са доходите на 490 компании от *S&P500* (*Standard&Poor's 500*), индекс за оценка на състоянието на фондовия пазар в САЩ, за 5-годишен период от 2021 до 2025 година. Данните за тяхната доходност са изтеглени от *Yahoo Finance* (<https://finance.yahoo.com/>) посредством програмната библиотека *yFinance*, достъпна

чрез хранилището на *PyPI* (<https://pypi.org/project/yfinance/>). Чрез тази библиотека се осъществява автоматизирано извличане на исторически ценови серии и пазарни индикатори чрез интерфейс към уеб страницата на *Yahoo Finance*, което позволява програмно изтегляне на данни за конкретни финансови инструменти в зададен времеви интервал и последващата им обработка в аналитична среда.

1. Генерираните симетрични матрици с размерности 100, 500, използвани за числените експерименти в Глава 1, са публикувани в платформата **GitHub** и са достъпни на следния адрес:

- https://github.com/smgurova/Symmetric-matrices_chapter-1.git

2. Финансовият отчет, предоставен от *Portfolio Visualizer*, е наличен в платформата **GitHub** на следния адрес:

- <https://github.com/smgurova/AssetAllocation.git>

3. Корелационната матрица на тридесетте и два глобални актива, използвана за числените експерименти в Глава 3, е достъпна в платформата **GitHub** на следния адрес:

- https://github.com/smgurova/correlation-matrix_chapter-4.git

4. Данните за доходността на 490 компании от *S&P500* (*Standard&Poor's 500*) за времеви период от 2021 до 2025 година и образуваната от тях корелационна матрица, използвана за числените експерименти в Глава 3, са достъпни в платформата **GitHub** на следния адрес:

- <https://github.com/smgurova/Data-S-P500.git>

Библиография

- [1] Alexandrov, V., Atanassov, E., Dimov, I., Branford, S., Thandavan, A., Weihrauch, C., Parallel Hybrid Monte Carlo Algorithms for Matrix Computations. In: Sunderam, V.S., van Albada, G.D., Sloot, P.M.A., Dongarra, J. (eds.) *Comput. Science - ICCS 2005*, LNCS, vol. **3516**, Springer, 752–759, (2005), https://doi.org/10.1007/11428862_102
- [2] Alexandrov, V., Davila, D., Esquivel-Flores, O., Karaivanova, A., Gurov, T., Atanassov, E., On Monte Carlo and quasi-Monte Carlo Methods for Matrix Computations, *Lecture Notes in Computer Science*, **10665**, Springer-Verlag, 249-257, (2018), https://doi.org/10.1007/978-3-319-73441-5_26
- [3] Alexandrov, V., Esquivel-Flores, O., Ivanovska, S., Karaivanova, A., On the Preconditioned Quasi-Monte Carlo Algorithm for Matrix Computations. In: Lirkov, I., Margenov, S., Waśniewski, J. (eds.) *Large-Scale Scientific Computing - LSSC 2015*, *Lecture Notes in Computer Science*, vol. **9374**, Springer, Cham., 163-171, (2015), https://doi.org/10.1007/978-3-319-26520-9_17
- [4] Alexandrov, V., Esquivel-Flores, O.A., Towards Monte Carlo preconditioning approach and hybrid Monte Carlo algorithms for matrix computations, *Computers & Mathematics with Applications*, vol. **70** (11), 2709-2718, (2015), <https://doi.org/10.1016/j.camwa.2015.08.035>
- [5] Almulih A.H. H., High-Performance Quasi-Monte Carlo Integration and Applications, *Dissertations*. 3452, Western Michigan University, (2019), <https://scholarworks.wmich.edu/dissertations/3452>, (Last accessed: January 2026)
- [6] Antonov, I., Saleev, V., An economic method of computing $LP\tau$ -sequence, *USSR Computational Mathematics and Mathematical Physics*, **19**, 252-256, (1979)
- [7] Atanassov, E., Gurov, T., Karaivanova, A., Ivanovska, S., Durchova, M., Dimitrov, D., On the parallelization approaches for Intel MIC architecture, *AIP Conference Proceedings*, vol. **1773**, 070001-1, (2016), <https://doi.org/10.1063/1.4964983>
- [8] Atanassov, E., On the discrepancy of the Halton sequence, *Mathematica Balkanica. New Series*, **18(1-2)**: 15-32, (2004)
- [9] Atanassov, E., Durchova, M., Generating and Testing the Modified Halton Sequences. In: Dimov, I., Lirkov, I., Margenov, S., Zlatev, Z. (eds.) *Numerical Methods and Applications*

- NMA 2002, Lecture Notes in Computer Science, vol. **2542**, Springer, Berlin, Heidelberg, 91-98, (2003), https://doi.org/10.1007/3-540-36487-0_9
- [10] Atanassov, E., Gurov, T., Ivanovska, S., Karaivanova, A., Parallel Monte Carlo on Intel MIC Architecture, *Procedia Computer Science*, Vol. **108**, 2017, 1803-1810, (Open Access), (2017), <https://doi.org/10.1016/j.procs.2017.05.149>
- [11] Branford, S., Hybrid Monte Carlo Methods for Linear Algebra Problems, Ph.D. dissertation, School of Systems Engineering, The University of Reading, (2009)
- [12] BRODA's Sobol RSG: broda.co.uk/software.html (Last accessed: 2026/01/11)
- [13] Brouwer, A.E., Haemers, W.H., *Spectra of Graphs*, Springer, (2012) <https://doi.org/10.1007/978-1-4614-1939-6>
- [14] Caflisch, R. E., Monte Carlo and quasi-Monte Carlo methods, *Acta Numerica*, Cambridge University Press, 1-49, (1998) <https://www.math.ucla.edu/~caflisch/Pubs/Pubs1995-1999/actaNumerica1998.pdf> (Last accessed: 2026/01/11)
- [15] Chi, H. Scrambled Quasirandom Sequences and Their Applications, Florida State University, (2004) http://purl.flvc.org/fsu/fd/FSU_migr_etd-3823, (Last accessed: 2026/01/11)
- [16] Chi, H, Beerli, P., Evans, D. W., Mascagni, M., On the Scrambled Sobol Sequence, LNCS 3516, Springer, 775-782, (2005)
- [17] Chung, F., *Spectral Graph Theory*, CBMS Regional Conference Series in Mathematics 92, AMS, (1997)
- [18] Dick, J., Pillichshammer, F., *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration*, Cambridge University Press (2010), <https://doi.org/10.1017/CB09780511761188>
- [19] Dekking, M., *A Modern Introduction to Probability and Statistics*. Springer, 181–190, (2005), ISBN: 9781852338961
- [20] Dimov, I., Monte Carlo algorithms for linear problems, *Pliska Stud. Math. Bulgar.*, **13**, 57-77, (2000)
- [21] Dimov, I., Dimov, T., Gurov, T., A new iterative Monte Carlo approach for inverse matrix problem, *Journal of Computational and Applied Mathematics*, **92**(1), 15-35, (1998)
- [22] Dimov, I., Tonev, O., Monte Carlo algorithms: performance analysis for some computer architectures, *Journal of Computational and Applied Mathematics*, **48**, 253-277, (1993)
- [23] Dimov, I., Alexandrov, V., Karaivanova, A, Implementation of Monte Carlo Algorithms for Eigenvalue Problem Using MPI, PVM/MPI'98, LNCS, **1497**, 346-353, (1998)

- [24] Dimov, I., Philippe, B., Karaivanova, A., Weihrauch, C., Robustness and applicability of Markov chain Monte Carlo algorithms for eigenvalue problems, *Applied Mathematical Modelling*, **32**(8), 1511-1529, (2008)
- [25] Dimov, I, Alexandrov, V., Karaivanova, A., Parallel resolvent Monte Carlo algorithms for linear algebra problems, *Mathematics and Computers in Simulation*, **55**(1), 25-35, Published by Elsevier BV., (2005)
- [26] Dimov, I., Karaivanova, A.: A Power Method with Monte Carlo Iterations, *Proceedings Recent Advances in Numerical Methods and Applications II*, World Scientific, 239-247, (1999), https://doi.org/10.1142/9789814291071_0022
- [27] Dimov, I., *Monte Carlo Methods for Applied Scientists*, World Scientific Publishing Co. Pte. Ltd., (2008), ISBN: 13-978-981-02-2329-8
- [28] Dimov, I., Karaivanova, A., Iterative Monte Carlo algorithms for linear algebra problems, *Springer Lecture Notes in Computer Science*, **1196**, 66-77, (1996)
- [29] Dimov, I., Karaivanova, A., Parallel computations of eigenvalues based on a Monte Carlo approach, *Monte Carlo Methods and Applications*, Vol. **4**, Num. 1: 33-52, (1998)
- [30] Dimov, I.T., et al.: Monte Carlo Algorithms for Calculating Eigenvalues. In: Niederreiter, H., et.al. (eds.) *Monte Carlo and Quasi-Monte Carlo Methods*, 1996, *Lecture Notes in Statistics*, vol. **127**, Springer, New York, NY., (1998)
- [31] Eichenauer-Hermann, J., Statistical independence of a new class of inversive congruential pseudorandom numbers, *Math. Comp.* **60**: 375-384, (1993)
- [32] Erkan, U., A precise and stable machine learning algorithm: eigenvalue classification (EigenClass), *Neural Comput & Applic*, **33**, 5381-5392, (2021), <https://doi.org/10.1007/s00521-020-05343-2>
- [33] Fathi, B., Liu, B., Alexandrov, V., Mixed MC parallel algorithms for Matrix Computation, *LNCS*, **2330**, 609-618, (2002)
- [34] Faure H., Lemieux Ch.: Generalized Halton sequences in 2008: A comparative study. *ACM Transactions on Modeling and Comp. Sim.* Vol. **19**, Issue 4, 1-31, (2009) <https://doi.org/10.1145/1596519.1596520>
- [35] Faure, H., Discrepance de suites associées à un système de numération (en dimension s), *Acta Arithmetica* **41**, 337-351, (1982)
- [36] Faure, H., Good permutations for extreme discrepancy, *Journal of Number Theory*, Vol. **42**, 47-56, (1992)
- [37] Fischer, H., *A History of the Central Limit Theorem: From Classical to Modern Probability Theory*, Sources and Studies in the History of Mathematics and Physical Sciences. New York: Springer, (2011), <http://dx.doi.org/10.1007/978-0-387-87857-7>

- [38] Fishman, G.S., Multiplicative congruential random number generators, *Math. Comp.*, **54**, 331-344, (1990)
- [39] Golub, G., Loan, Ch. F. Van, *Matrix computation*, 4th edition, The Johns Hopkins University Press, (2013), ISBN: 13-978-1-4214-0794-4, <https://math.ecnu.edu.cn/~jypan/Teaching/books/2013%20Matrix%20Computations%204th.pdf>
- [40] Gilboa, G., *Nonlinear Eigenproblems in Image Processing and Computer Vision*, Publisher: Springer Cham, ed.number 1, (2018), eBook ISBN: 978-3-319-75847-3
- [41] Gurova, S.-M., Atanasov, E., Karaivanova, A., A Resolvent Quasi-Monte Carlo Method for Estimating the Minimum Eigenvalues Using the Error Balancing, In: Lirkov, I., Margenov, S. (eds.), *Large-Scale Scientific Computing, LSSC 2023, Lecture Notes in Computer Science*, vol. **13952**, 394-403, (2024)
- [42] Gurova, S.-M., Karaivanova, A., Quasi-Monte Carlo Algorithms for Eigenvalue Problems, Fourth International Conference Numerical Methods for Scientific Computations and Advanced Applications, NMSCAA'24, 17-21, (2024) (short communication) <https://parallel.bas.bg/~elilkova/nmscaa24.pdf>
- [43] Gurova, S.-M., Gurov, T., Karaivanova, A., On the Stochastic Power Algorithms for Estimating the Maximum Eigenvalue of Symmetric Matrices, *Contemp. Math.*, Universal Wiser Publisher Pte. Ltd., **6**(5): 7223-7249, (2025), <https://doi.org/10.37256/cm.6520257781>
- [44] Gurova, S.-M., Gurov, T., Karaivanova, A., Eigenvalue Estimation in Portfolio Risk: The Role of Skipping and Leaping in Sobol and Halton Sequences, In: Lirkov, I., Margenov, S. (eds), *LSSC'25, LNCS*, **16061**,(accepted for publication)
- [45] Halton, J., On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals, *Numerische Mathematik*, **2**, 84-90, (1960), <https://doi.org/10.1007/BF01386213>
- [46] Halton, J., Smith, GB, Algorithm 247: Radical-Inverse Quasi-Random Point Sequence, *Communications of the ACM*, Vol. **7**, 701-702, (1964)
- [47] Hammersley, J.M., Handscomb, D.C., *Random, Pseudorandom, and Quasirandom Numbers*. In: *Monte Carlo Methods. Monographs on Applied Probability and Statistics*. Springer, Dordrecht, (1964), https://doi.org/10.1007/978-94-009-5819-7_3
- [48] Hellekalek, P., Inversive pseudorandom number generators: concepts, results and links, *Proceedings of the Winter Simulation Conference*, 255-262, (1995)
- [49] Higham, N. J., Computing the nearest correlation matrix—a problem from finance, *IMA Journal of Numerical Analysis*, **22**(3): 329–343, (2002), <https://doi.org/10.1093/imanum/22.3.329>

- [50] Holian, B., et al., Pseudorandom number generator for massively parallel molecular-dynamics simulations, *Phys. Rev. E* **50**, 1607, (1994), <https://doi.org/10.1103/PhysRevE.50.1607>
- [51] Isaacson, E., Keller, H. B., *Analysis of Numerical Methods*, Dover Publications, Inc., N.Y., Wiley, (1996), ISBN: 0-486-68029-0
- [52] Ivanovska, S., Karaivanova, A., Manev, N., Numerical Integration Using Sequences Generating Permutations. In: Lirkov, I., Margenov, S., Waśniewski, J. (eds) *Large-Scale Scientific Computing. LSSC 2011. Lecture Notes in Computer Science*, vol. **7116**, Springer, Berlin, Heidelberg, 455-463, (2011), https://doi.org/10.1007/978-3-642-29843-1_51
- [53] Joe, S., Sloan, I., Implementation of a lattice method for numerical multiple integration, *ACM Transactions on Mathematical Software (TOMS)*, Volume **19**, Issue 4, 523 – 545, (1993), <https://doi.org/10.1145/168173.168425>
- [54] Joe S., Kuo F.Y.: Constructing Sobol Sequences with Better Two-Dimensional Projections, *SIAM Journal on Scientific Computing*, **30**(5): 2635-2654, (2008), <https://doi.org/10.1137/070709359>
- [55] Johnstone, I. M., On the distribution of the largest eigenvalue in principal components analysis, *Ann. Statist.* **29**(2), 295-327, (2001), <https://doi.org/10.1214/aos/1009210544>
- [56] Kalos, M. and Whitlock, P., *Monte Carlo Methods*, 2nd ed., Wiley-VCH, (2009), ISBN: 978-3-527-62621-2
- [57] Karaivanova, A., Quasi-Monte Carlo Methods for Some Linear Algebra Problems, Convergence and complexity, *Serdica J. Computing*, **4**, 57-72, (2010)
- [58] Kocis, L., and W. J. Whiten, Computational Investigations of Low-Discrepancy Sequences, *ACM Transactions on Mathematical Software*, **23**(2), 266–294, (1997)
- [59] Kolda, T. and Bader, B., Tensor Decompositions and Applications, *SIAM Review*, **51**(3), 455-500, (2009) <https://doi.org/10.1137/07070111X>
- [60] Kroese, D.P., Taimre, T., Botev, Z. I., *Handbook of Monte Carlo Methods*, John Wiley & Sons Inc., New York, NY., (2011), ISBN: 9780470177938
- [61] Laloux, L., Cizeau, P., Bouchaud, J.-P., Potters, M., Noise Dressing of Financial Correlation Matrices, *Phys. Rev. Lett.*, **83**(7), 1467-1470, American Physical Society, (1999), <https://doi.org/10.1103/PhysRevLett.83.1467>
- [62] Lemieux, C., *Monte Carlo and Quasi-Monte Carlo Sampling*. Springer, (2009), <https://doi.org/10.1007/978-0-387-78165-5>

- [63] L'Ecuyer, P., Randomized Quasi-Monte Carlo: An Introduction for Practitioners, 12th International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing (MCQMC 2016), <https://inria.hal.science/hal-01561550v1>
- [64] L'Ecuyer, P. Uniform random number generation. *Ann Oper Res*, **53**, 77–120, (1994), <https://doi.org/10.1007/BF02136827>
- [65] L'Ecuyer, P., Efficient and portable combined random number generators, *Communications of the ACM*, vol.**31**, Issue 6, 742-751, (1988), <https://doi.org/10.1145/62959.62969>
- [66] Lobo, P.D.C., Mikhailov, M., Ozisik, M.N., On the complex eigenvalues of Luikov system of equations, *Drying Technology*, **5**(2), 273-286, (1987), <https://doi.org/10.1080/07373938708916540>
- [67] Loperfido, N., Skewness-based projection pursuit: A computational approach, *Computational Statistics and Data Analysis*, **120**, 42-57, (2018), <https://doi.org/10.1016/j.csda.2017.11.001>
- [68] Macdonald, R.R., Correlation and Covariance Matrices. In *Wiley StatsRef: Statistics Reference Online* (eds. N. Balakrishnan, T. Colton, B. Everitt, W. Piegorisch, F. Ruggeri and J.L. Teugels), (2014), <https://doi.org/10.1002/9781118445112.stat06481>
- [69] Marsaglia, G., Diehard: A battery of tests of randomness, Florida State University, (1995)
- [70] Marsaglia, G., Random Number Generators, *Journal of Modern Applied Statistical Methods: Vol. 2*, Issue 1, Article 2, (2003), <https://doi.org/10.22237/jmasm/1051747320>
- [71] Mascagni, M., SPRNG: A Scalable Library for Pseudorandom Number Generation, in: O. Iliev et al., eds., *Recent Advances in Numerical Methods and Applications II*, Proceedings of NMA'98 (WS, Singapore), 284-295, (1999).
- [72] Mascagni, M., Karaivanova, A., A Parallel Quasi-Monte Carlo Method for Solving Systems of Linear Equations, P.M.A. Sloot et al. (eds.): *ICCS 2002, LNCS 2330*, Springer-Verlag Berlin Heidelberg, 598-608, (2002)
- [73] Mascagni, M., Karaivanova, A., A Parallel Quasi-Monte Carlo Method for Computing Extremal Eigenvalues, K.-T. Fang et al. (eds.), *Monte Carlo and Quasi-Monte Carlo Methods*, Springer-Verlag Berlin Heidelberg, 369 – 380, (2000)
- [74] Mascagni, M., Karaivanova, A., Matrix Computations Using Quasirandom Sequences, *Lecture Notes in Computer Science*, (Wulkov, Yalamov, Wasniewsky eds.), Vol. **1988**, Springer, 552-559, (2001)
- [75] Mascagni, M., Karaivanova, A., A Monte Carlo Approach for Finding More than One Eigenpair. In: Dimov, I., Lirkov, I., Margenov, S., Zlatev, Z. (eds.) *Numerical Methods*

- and Applications - NMA 2002, Lecture Notes in Computer Science, vol. **2542**, Springer, 123-131, (2003), https://doi.org/10.1007/3-540-36487-0_13
- [76] Mascagni, M., Srinivasan, A., Parameterizing parallel multiplicative lagged-Fibonacci generators, *Parallel computing*, **30**(7), 899-916, (2004), <https://doi.org/10.1016/j.parco.2004.06.001>
- [77] Matousek's, J., On the L2-discrepancy for anchored boxes, *Journal of Complexity*, vol. **14**, 527–556, (1998)
- [78] Matsumoto, M., Nishimura, T., Mersenne Twister: A 623-dimensionally Equidistributed Uniform Pseudo-Random Number Generator, *ACM Transactions on Modeling and Computer Simulation*, vol. **8** (1), 3–30, (1998), <https://doi.org/10.1145/272991.272995>
- [79] Mehrdoust, F., Algorithm for Estimating the Condition Number of Matrices, *Math. Reports*, **15(65)**(3), 203-210, (2013)
- [80] Muminov M.I., Rasulov T.H.: Embedded Eigenvalues of a Hamiltonian in Bosonic Fock Space. *Commun. Math. Anal.*, **17**(1), 1 - 22, (2014)
- [81] Mushtaq, A., Noreen, A., Olaussen, K., Numerical Solutions of Quantum Mechanical Eigenvalue Problems, *Front. Phys.*, **8**:390, (2020), <https://doi.org/10.3389/fphy.2020.00390>
- [82] Neumann, J., Various techniques used in connection with random digits, A.S. Householder, G.E. Forsythe, and H.H. Germond, eds., *Monte Carlo Method*, National Bureau of Standards Applied Mathematics Series, vol. **12** (Washington, D.C.: U.S. Government Printing Office, 36–38, (1951)
- [83] Niederreiter, H., *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM (1992), <https://doi.org/10.1137/1.9781611970081>
- [84] Niederreiter, H., Nets, (t,s) -sequences, and algebraic curves over finite fields with many rational points, *Doc. Math. (Bielefeld) Extra Vol. ICM Berlin*, vol. **III**, 377–386, (1998).
- [85] Niederreiter, H., Constructions of (t,m,s) -nets and (t,s) -sequences, *Finite Fields and Their Applications* Vol. **11**, Issue 3, 578-600, (2005), <https://doi.org/10.1016/j.ffa.2005.01.001>
- [86] Owen, A., Randomly Permuted (t,m,s) -Nets and (t, s) -Sequences. In: Niederreiter, H., Shiue, P.J.S. (eds.) *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*. Lecture Notes in Statistics, vol. **106**. Springer, New York, 299-317, (1995), https://doi.org/10.1007/978-1-4612-2552-2_19
- [87] Owen, A., Monte Carlo variance of scrambled net quadrature, *SIAM Journal on Numerical Analysis*, Vol. **34**, No. 5, 1884-1910, (1997), <https://www.jstor.org/stable/2952022>

- [88] Owen, A., Scrambling Sobol and Niederreiter-Xing Points, *Journal of Complexity*, **14**(4): 466-489, (1998)
- [89] Owen, A., Variance with alternative scramblings of digital nets, *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, Vol. **13**(4), 363–378, (2003), <https://doi.org/10.1145/945511.945518>
- [90] Padányi, V., Herendi, T., Generalized Middle-Square Method, *Annales Mathematicae et Informaticae*, **56**, 95-108, (2022), <https://doi.org/10.33039/ami.2022.12.003>
- [91] Park, S.K., Miller, K.W., Random Number Generators: Good Ones Are Hard To Find, *Communications of the ACM*, vol. **31**(10), 1192–1201, (1988), <https://doi.org/10.1145/63039.63042>
- [92] Rometsch, M., *Quasi-Monte Carlo Methods in Finance with Application to Optimal Asset Allocation*, Druck Diplomica Verlag GmbH, Hamburg, (2008)
- [93] Rotter, I., A non-Hermitian Hamilton operator and the physics of open quantum systems, *J. Phys. A: Math. Theor.*, Vol. **42**, No. 15, 153001, (2009), <https://doi.org/10.1088/1751-8113/42/15/153001>
- [94] Saito, M., Matsumoto, M., SIMD-Oriented Fast Mersenne Twister: a 128-bit Pseudorandom Number Generator, In: Keller, A., Heinrich, S., Niederreiter, H. (eds.) *Monte Carlo and Quasi-Monte Carlo Methods 2006*. Springer, Berlin, Heidelberg, 607–622, (2008), https://doi.org/10.1007/978-3-540-74496-2_36
- [95] Schlier, C., On scrambled Halton sequences, *Applied Numerical Mathematics*, **58**(10), 1467–1478, (2008), <https://doi.org/10.1016/j.apnum.2007.09.001>
- [96] Sobol I., Shukhman B., Integration with quasirandom sequences: Numerical experience., *Internat. J. Modern Phys. C*, **6**(2), 263-275 (1995)
- [97] Sobol I., Asotsky D., Kreinin A., Kucherenko S., Construction and Comparison of High-Dimensional Sobol' Generator, *WILMOTT magazine*, 64-79, (2012)
- [98] Sobol I., Uniformly distributed sequences with an additional property of uniformity, *USSR Computational Mathematics and Mathematical Physics*, **16**(5), 236–242, (1976), [https://doi.org/10.1016/0041-5553\(76\)90154-3](https://doi.org/10.1016/0041-5553(76)90154-3)
- [99] Spiegel, M., *Theory and Problems of Probability and Statistics*, 2nd ed. New York: McGraw-Hill, (1992)
- [100] Straßburg J, Alexandrov, V., Enhancing Monte Carlo preconditioning methods for matrix computations, *Procedia Computer Science*, **29**, 1580-1589, (2014), <https://doi.org/10.1016/j.procs.2014.05.143>
- [101] Straßburg, J., *On hybrid and resilient Monte Carlo methods for linear algebra problems*, Ph.D. dissertation, The University of Reading, (2014)

- [102] Sturmfels, B., Tensors and Their Eigenvectors, Notices of the AMS, **63**(6), 604-606, (2016) <http://dx.doi.org/10.1090/noti1389>
- [103] Sugita, H., Monte Carlo Method, Random Number, and Pseudorandom Number, MSJ Memoirs, **25**: 133, (2011), <https://doi.org/10.2969/msjmemoirs/025010000>
- [104] Taft K., Vajargah, F., MC method for solving system of linear algebraic equations with minimum markov chains, International Conference PDPTA'2000 Las Vegas, (2000)
- [105] Tezuka, S., On randomization of generalized Faure sequences, Research Report, RT0494:15 pages, (2002)
- [106] Tezuka, S., Faure, H., I-binomial scrambling of digital nets and sequences, Journal of Complexity, **19**(6): 744-757, (2003)
- [107] Tezuka, S., Linear Congruential Generators. In: Uniform Random Numbers. The Springer International Series in Engineering and Computer Science, Springer, Boston, MA., vol. **315**, 57–82, (1995), https://doi.org/10.1007/978-1-4615-2317-8_3
- [108] Vajargah, K., Kamalzadeh, F., Estimating of Eigenvalue with Monte Carlo Method and its Application in the Principal Components (PCA), Journal of Mathematics and Computer Science, **9**, 240–248, (2014)
- [109] Vajargah F., et al., Parallel Monte Carlo algorithms for matrix computations, Ph.D. dissertation, The University of Reading, (2001)
- [110] Vajargah, F., Hassanzadeh, Z., Improvements on the hybrid Monte Carlo algorithms for matrix computations, Sadhana, Indian Academy of Science, **44**(1), <https://doi.org/10.1007/s12046-018-0983-y>
- [111] Van der Corput, J., Verteilungsfunktionen (Erste Mitteilung), Proceedings of the Koninklijke Akademie van Wetenschappen te Amsterdam (in German), **38**: 813–821, (1935), Zbl 0012.34705
- [112] Wang, X., Hickernell, F., Randomized Halton sequences, Mathematical and Computer Modelling, **32**(7–8), 887-899, (2000) [https://doi.org/10.1016/S0895-7177\(00\)00178-3](https://doi.org/10.1016/S0895-7177(00)00178-3)
- [113] Weyl, H., On the uniform distribution of numbers modulo one, Mathematische Annalen (in German), **77**(3), 313-352, (1916), <https://doi.org/10.1007/BF01475864>
- [114] Боянов, Б., Лекции по Числени методи, изд. Дарба София, (2008), ISBN: 954-90126-1
- [115] Димитров, Б., Янев, Н., Вероятности и статистика, Софттех, София, (2007)
- [116] Караиванова, А., Стохастични числени методи и симулации, изд. Деметра ЕООД, (2012), ISBN: 978-954-9526-78-3

- [117] Сендов, Б., Попов, В., Числени методи, Университетско издание „Св. Климент Охридски“, първа част, (1996), ISSN: 954-07-0994-6
- [118] Стоянов, Й., Миразчийски, И., Игнатов, Ц., Танушев, М., Ръководство за упражнения по теория на вероятностите. Издателство „Наука и изкуство“, София, (1976).
- [119] Соболев, И. М., Численние методы Монте-Карло, изд. Москва, Наука, (1973)
- [120] MathWorks: haltonset - Create a Halton point set. <https://www.mathworks.com/help/stats/haltonset.html>, (Last accessed: 2026/01/11)
- [121] MathWorks: sobolset - Create a Sobol point set. <https://www.mathworks.com/help/stats/sobolset.html>, (Last accessed: 2026/01/11)