

# РЕЦЕНЗИЯ

по процедура за защита на дисертационен труд на тема:

*"Симулация на течения в порести среди чрез масивно паралелен Многонилов Монте Карло алгоритъм"*

за придобиване на образователна и научна степен Доктор

кандидат: **Николай Георгиев Шегунов,**

Област на висше образование: **4. Природни науки, математика и информатика**

Професионално направление: **4.6. Информатика и компютърни науки**

Докторска програма: **Информационни системи,** катедра: **Компютърна информатика,**

**Софийски университет „Св. Климент Охридски“, Факултет по математика и информатика**

Рецензията е изготвена от: **проф. д-р Васил Георгиев Цунижев, ФМИ – СУ «Св. Кл. Охридски»** в качеството на член на научното жури, съгласно Заповед № РД-38-258/7. 06. 2021 г. на Ректора на Софийския университет.

## **1. Обща характеристика на дисертационния труд и представените материали**

Представеният за рецензиране дисертационен труд с горесцитираното заглавие и с авторски превод на заглавието „Симулация на течения в порести среди чрез масивно паралелен Многонилов Монте Карло алгоритъм“ представлява монография на английски език, посветена на решаването на научно-приложен проблем от областта на симулационното моделиране на процесите на взаимодействие на материали в течна и твърда фаза. Независимо от приложната област обаче, авторът по същество си е поставил задачата да оптимизира процеса на машинна числова обработка на такъв клас симулационни модели чрез числово решаване на съответната система от равновесни уравнения по метода на случайните процеси, известен като Монте Карло метод. Двете главни особености на избрания подход са

- (1) приложимост на „многонивов“ Монте Карло метод за моделиране на проблемната област, свързана с уравнението на конвекция-реакция-дифузия и също така
- (2) приложимост и резултати от паралелната числова обработка на модела.

При това предвид показаните резултати става дума за машинен и оттам за софтуерен паралелизъм, който действително може да се окачестви като масивен, тъй като при някои от докладваните експерименти е измерена скалируемостта при стойности на паралелизма 9600.

По същество приносят на дисертацията е именно модел за високопаралелна числова обработка на „многонивов“ Монте Карло алгоритъм. Разбира се в хода на разгръщане и експериментално изследване на този модел са решени и редица съпътстващи въпроси, включително свързани с първоначалната приложна област на симулационния числов модел, а именно поръозността на твърдите среди.

Основния труд е представен на 129 страници, съдържащи 5 глави и кратко заключение, което включва списък на приносите и списък от пет публикации по дисертацията. В този текст са включени и няколко приложения относно използваните съкращения, списък на фигурите и таблиците, кратък опис на използваната хардуерна инфраструктура (2 различни споделени клъстера), както и списък на източниците с 62 заглавия.

Към този основен текст се добавя Автореферат на български от 32 страници и допълнения към тях, представящо основно списъка източници, списъка авторски публикации и списък на приносите.

Като цяло основният труд и съпътстващите го документи създават впечатлението за компетентност, дълбоко познаване на изследваната проблематика и свободно навлизане в интердисциплинарни области. Описанието е прецизно и лаконично, с правилно подбрани подтеми за обсъждане в отделните глави. Все пак могат да се идентифицират известни слабости и възможности за подобрене, на които ще се спра в т. 6.

## **2. Данни и лични впечатления за кандидата**

Дисертантът Николай Шегунов познавам по повод изследователската му и преподавателска работа във ФМИ. Неговата подготовка и дейност се изпълняват под наблюдението на колегите от катедра Компютърна информатика, на която са членове заедно с научния си ръководител доцент Армянов. Това, с което се характеризира и е известен в своята професионална общност колега Николай, е задълбочената работа, високите познания в интересующите го проблемни области, успешното му и търсено участие в образователния процес, интеграцията му в международни и мултидисциплинарни изследователски колективи и при това проявената лоялност да не се откъсва от основното изследователско и образователно звено.

### **3. Съдържателен анализ на научните и научно-приложните постижения на кандидата, съдържащи се в представения дисертационен труд и публикациите към него, включени по процедурата**

Научните и научно-приложни приноси на дисертанта са отразени коректно в приложените документи по процедурата.

Дисертацията започва с обзор и сравнителен анализ на съществуващите решения на поставените проблеми. Оценени са предимствата и недостатъците на съществуващите решения за генериране на стохастични полета и семплиращи алгоритми. Част от Глава 1. и почти цялата Глава 2. на дисертацията са посветени на този обзор. В края на Секция 1.3 са разгледани различните подходи, прилагани за симулиране на течения в порести среди. Показани са предимства и слабости за всеки от тях. Секция 1.4 разглежда технически средства и софтуерни библиотеки, които могат да се използват за решаването на поставените проблеми. В точка 2.1 са разгледани и съпоставени начините за генериране на стохастични полета за целите на поставената задача. В секция 2.2 са разгледан метода на крайните обеми, използван за моделиране на поставения проблем. В секция 2.4 е описан в детайли използвания в дисертацията модел, базиран на „многогониовия“ Монте Карло метод. В Глава 3. са сравнени и анализирани са различни подходи за приближение на стохастичното ниво за задачата на Лаплас.

По-нататък в Глава 3. е разкван и обоснован подходът за ренормализация на стохастичното поле за целите на „многониволия“ Монте Карло метод скаларни елиптически стохастични частни диференциални уравнения. Това е видно от сравнението на „многониволия“ Монте Карло метод с използване ренормализация спрямо класическия, които показват предимството на разглеждания подход са поместени в дисертацията на фигури: 3.6, 3.7 и 3.8 (41-43 страница). На тях са показани съответно броят експерименти, проведени на различните нива и намаляването на време за изчисление (Фигура 3.8). Таблицы 3.2 и 3.3 показват предимствата на приложената от автора *опростена ренормализация* спрямо *средно аритметично*.

По-нататък в Глава 4. е описано прилагането на „многониволия“ Монте Карло метод за уравнението на конвекция-реакция-дифузия. Разработен е адаптивен алгоритъм за разпределение на ресурсите върху различните нива на „многониволия“ Монте Карло алгоритъм. Приложени са резултати с различни параметри на симулация - различни числа на Пекле и Дамколер. Резултатите са поместени в дисертацията в таблица 4.1 и таблица 4.2.

По отношение на апробация и приложение на направените разглеждания в Глава 5. също има добри резултати. Направен е преглед, анализ и сравнение на 6 стратегии за паралелизация тъй като е разгледана динамична стратегия за паралелизация между различните нива като и за така предложената стратегия са анализирани три подхода за паралелизация на двата слоя - локално прекъсване, глобално прекъсване и „опашков“ паралелизъм.

Така описаният модел на паралелната обработка на числовия алгоритъм е изследван експериментално чрез програмна реализация и обработка в условията на системи с масивен паралелизъм. Докладваните тестови резултати за ускорението (нормализирано чрез величината ефективност) показват напълно удовлетворяващи резултати в условията на паралелизъм от порядъка на няколко хиляди. Използването на предложената от автора оптимизация с глобално прекъсване води до подобрене на ефективността при 4800 процесорни ядра с 15%, при 7200 ядра с 12% и при 9600 ядра със 7%. Разбира се, от никой паралелен алгоритъм не можем да очакваме безкрайна скалируемост. Достатъчно е, че за конкретния обхват на паралелизма виждаме, че оптимизацията „работи“. Намалението на относителната ефективност с нарастване на паралелизма авторът обяснява – и то с основание – с два фактора: фиксирана изчислителна сложност на заданието при нарастващ

обработващ капацитет и нарастващи комуникационни закъснения в по-големия сърверен клъстер.

#### 4. Аprobация на резултатите

Докладваните от кандидата пет публикации по дисертацията, както и самата дисертация и изготвения автореферат, се характеризират със следните черти:

- (1) пълно, точно и коректно отразяват представения дисертационен труд;
- (2) всички публикации по дисертацията са индексирани в Скопус и следователно може да се изчисли, че изискванията за публикационна активност са изпълнени почти двукратно:

по-точно за общия индекс на автора  $R_{NSH}$  по група показатели  $G$  от ЗРАС могат да се направят следните наукометрични изчисления (по [индексите] от списъка публикации):

$$R_{NSH} = 1/3 \times 30 \text{ т.}_{[1]} + 0.8^1 \times 30 \text{ т.}_{[2]} + 1/3 \times 30 \text{ т.}_{[3]} + 1/2 \times 30 \text{ т.}_{[5]} = 59 \gg 30.$$

Публикация [4] пренебрегвам поради големия брой съавтори.

Не се установяват засега цитирания в Скопус, ако не броим 3 самоцитирания и авторът не е приложил сведения за евентуално цитиране, но този показател не се изисква по текущата процедура;

- (3) всички публикации не са отчитани по предходна процедура;
- (4) публикациите на кандидата са свободни от плагиатство предвид извършените проверки в рецензираните и индексирани издания както и машинната проверка на дисертационния труд и автореферата.

#### 5. Качества на автореферата

Както отбелязах в 4.(1), авторефератът пълно, точно и коректно отразява дисертационния труд и получените резултати като се отчете езиковата разлика между двата документа.

<sup>1</sup> коефициент на база на приложен разделителен протокол между авторите.

## 6. Критични бележки и препоръки

Формулировката на заглавието и по същество на цялото изследване не е фокусирана правилно с оглед на представяне на съдържанието на дисертационен труд: вместо да представи цялостно разгърнато изследване на *формулиран модел* относно изследвания клас изчислителни проблеми, заглавието извежда фокуса върху решението на приложен проблем в областта на поръзността. То съответства по-скоро на отчет на проведено приложно изследване, отколкото на дисертация, предполагаща моделиране на проблемната област и изследване на характеристиките на така полученото [многомерно] моделно пространство. Бих предпочел като по-съответстващо на общността на проблема заглавие от вида [*Модел на*] *масивен паралелизъм при обработката на MLMC [клас от] алгоритми*. А тестовите приложни случаи могат да се оставят в подзаглавие или друго ниво на представяне, макар и на видно място – най-вече в заключителната част с експерименталните резултати. Между другото първоначалния мотив да се разгърне предложената в дисертацията изчислителна програма дори и да е бил моделиране на поръзността, за целите на дисертацията той може спокойно и да бъде пропуснат, тъй като тематиката и претендираните приноси не са в областта на материалознанието. Разбра се, наличието на приложна област е важно предимство и нейното цитиране, даже изтъкване, е строго необходимо, но по-скоро в апробационната част на дисертацията.

За щастие това не влияе по същество на дълбочината и качеството на проведеното изследване: така например изследователският модел може да се разчете неявно формулиран в точки 5.1 до 5.4, а в тт. 5.5 и 5.6 е даден изследователски план за експериментални резултати по отношение на два характерни изследователски случая.

Дисертацията е представена на английски език, което маскира традиционните проблеми с терминологията в основния текст, но те се появяват в автореферата. Прекият превод на *multilevel* като *многонивов* – е дума, която звучи нагласено и двусмислено. Не е тук мястото и повода за терминологична дискусия, но в конкретния случай бих предложил на вниманието на специалистите по Монте Карло симулиране термина *многофазен* или *многостъпков*: най-напред трябва да бъдат генерирани междинни резултати, на базата на които [от следваща фаза и евентуално от друг алгоритъм] се генерират резултати от друго ниво или стъпка. Такъв е и разглеждания случай на изчисления: „Очакваната стойност на най-финото ниво се получава като сума от изчислената стойност за това ниво плюс корекции получени за стойностите на по-грубите нива“.

Друг повод за дискусия е употребата от дисертанта на термина *грануларност* и степените ѝ от *груба* (coarse – бих предпочел *едра*) към *фина*. Смесът на употребата на грануларността тук е в посока *прецизност* или *дискретизация* на числовия модел на съответната структура на твърдото тяло. Предимството на *прецизност* и на *дискретизация* е, че те не създават смислов конфликт с една друга – много важна при паралелната обработка с декомпозиция по данни – величина, която също се означава като грануларност [на декомпозицията]. Този конфликт остава скрит в рецензираната работа, но само защото проблемът с грануларността на декомпозицията не е адресиран явно в изследователския план.

Описанието на изпълнителната платформа/и за тестови експерименти – хардуерни клъстери от паралелни машини, както и поддържащата междинна платформа – са дадени твърде схематично. Междувременно отделни параметри на тази платформа биха могли да участват в Модела за високопаралелна обработка и да се провери ролята им за повишаване на ефективността – по-конкретно имам предвид адаптивност към кеша на отделните ядра, както и „цената“ на синхронизацията между локални и отдалечени процеси – т.е. като цяло на топологията на процесите и отношението ѝ към и топологията на възлите.

Също така известно е, че в областта на симулационното Монте Карло моделиране и числовите методи за неговата паралелна обработка в България имаме цяла изградена школа, но за съжаление това не личи от списъка с източници по дисертацията. От друга страна не виждам с какво такъв специализиран текст печели от източници с предимно справочен профил, като например [22, 23, 24].

*Efficiency* правилно е преведено като ефективност, но се среща в автореферата и като *ефикасност* (във Фиг. 4.7), вариант, по-подходящ за сферата на лекарствените средства.

## 7. Заключение

След като се запознах с представените в процедурата дисертационен труд и придружаващите го научни трудове и въз основа на направения анализ на тяхната значимост и съдържащи се в тях научни и научно-приложни приноси, **потвърждавам**, че представеният дисертационен труд и научните публикации към него, както и качеството и

оригиналността на представените в тях резултати и постижения, отговарят на изискванията на ЗРАСРБ, Правилника за приложението му и съответния Правилник на СУ „Св. Климент Охридски“ за придобиване от кандидата на образователната и научна степен *доктор* в научната област 4. Природни науки, математика и информатика и професионално направление 4.6. Информатика и компютърни науки. Потвърждавам, че кандидатът удовлетворява минималните национални изисквания в професионалното направление и не е установено плагиатство в представените по конкурса научни трудове.

Въз основа на гореизложеното, **препоръчвам** на научното жури да присъди на **Николай Георгиев Шегунов** образователна и научна степен *доктор* в научна област 4. Природни науки, математика и информатика и професионално направление 4.6. Информатика и компютърни науки и пожелавам на Николай успешна и мотивираща го кариера на млад учен.

8. август 2021 г.

Изготвил рецензията: .....

(проф. д-р Васил Цунижев)