

## ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

кандидата физико-математических наук Черных Игоря Геннадьевича на диссертационную работу ГАРЕЕВА Романа Альбертовича «Методы оптимизации выполнения тензорных операций на многоядерных процессорах», представленную на соискание ученой степени кандидата физико-математических наук по специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.

**Актуальность темы диссертационной работы.** Для реализации численных решений большого количества задач особую важность приобретает сокращение времени выполнения тензорных операций. Как правило, готовые реализации таких операций распространяются в виде библиотек, созданных с использованием ручной настройки и автонастройки. Такой подход неприменим, если на создание новых реализаций отводится ограниченное время, а также недоступна целевая аппаратная платформа. В указанных случаях можно смоделировать выполнение операции на целевой архитектуре процессора с целью получения значений параметров реализации, позволяющих достичь наилучшей производительности. В следствие этого *актуальными* являются задачи создания методов сокращения времени выполнения тензорных операций на многоядерных процессорах и их автоматического распараллеливания, не использующих ручную настройку и автонастройку.

Для достижения поставленной *цели*, состоящей в разработке эффективных методов обработки информации, получаемой со сложных систем, путем сокращения времени выполнения многопоточных реализаций тензорных операций на многоядерных процессорах общего назначения без ручной настройки и автонастройки, соискатель формулирует и рассматривает четыре задачи:

1. Создание модели целевой архитектуры процессора для сокращения времени выполнения матрично-векторных операций и их обобщений.
2. Создание алгоритмов, сокращающих время выполнения тензорных операций на многоядерных процессорах общего назначения.
3. Применение созданных алгоритмов и модели для разработки программной системы, используемой для автоматической оптимизации времени выполнения тензорных операций и их автоматического распараллеливания при компиляции программ.
4. Проверка эффективности разработанной программной системы с помощью вычислительных экспериментов.

Актуальность и практическая значимость диссертации подтверждается наличием разработанной программной системы, части которой используются в основном коде Polly проекта Low Level Virtual Machine.

**Обоснованность и достоверность результатов.** Диссертация Р.А. Гареева обладает смысловым единством, достоверность представленных в ней результатов подтверждена экспериментальными данными, наличием публикаций и докладов по теме диссертации на различных специализированных конференциях. Таким образом, работа носит законченный характер.

**Научная новизна.** Основная научная новизна полученных в работе результатов состоит в следующем:

1. Разработано новое расширение модели целевой архитектуры процессора Лоу, а также оригинальные алгоритмы выполнения тензорных операций константной сложности относительно размерности тензоров для сокращения времени выполнения таких операций.
2. Создана оригинальная программная система для автоматической оптимизации времени выполнения тензорных операций и их автоматического распараллеливания в процессе компиляции программ на многоядерных процессорах общего назначения.

Полученные теоретические и практические результаты, а также разработанные программные средства дают специалистам в области математического моделирования инструмент для эффективной организации параллельных вычислений на многоядерных процессорах.

Все основные результаты, выносимые на защиту, получены впервые и представляют несомненный интерес для специалистов в области математического и программного обеспечения вычислительных машин, комплексов и компьютерных сетей.

**Теоретическая и практическая ценность.** Сформулированные в диссертации научные положения и основные результаты представляются обоснованными. Так Р.А. Гареевым создано расширение модели целевой архитектуры процессора Лоу для сокращения времени выполнения матрично-векторных операций и их обобщений на замкнутые полукольца; разработаны оригинальные алгоритмы выполнения тензорных операций константной сложности относительно размерности тензоров для сокращения времени выполнения таких операций. Результаты диссертации имеют практическую важность и состоят в разработке программной системы автоматической оптимизации тензорных операций, позволяющей автоматически оптимизировать время выполнения тензорных операций и осуществить их автоматическое распараллеливание в процессе компиляции программ для многоядерных процессоров.

**Рекомендации по использованию полученных результатов.** Полученные в диссертационной работе результаты могут применяться для реализации численных решений задач математической физики, математической геофизики и механики на многоядерных процессорах общего назначения.

**Публикации и апробации.** Основные результаты по теме диссертации в полном объеме изложены в 7 публикациях, в том числе в 2 статьях в журналах, входящих в Перечень рецензируемых научных изданий ВАК; в 2 публикациях в изданиях,

проиндексированных в Web of Science; и в 2 публикациях в изданиях, проиндексированных в Scopus. Результаты диссертации докладывались и обсуждались на 5 международных и всероссийских научных конференциях.

**Соответствие диссертации паспорту специальности.** Содержание диссертации и полученные результаты соответствуют специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей (п.8).

**Основное содержание диссертации.** Диссертация состоит из введения, четырех глав, заключения и списка литературы. Общий объем диссертации составляет 179 страниц, включая 37 рисунков, 34 таблицы и список литературы из 177 наименований.

Во введении отмечена актуальность темы исследований, сформулирована цель работы и полученные результаты, описана новизна и практическая ценность.

В первой главе проведен детальный аналитический обзор литературы по теме исследований, на основании которого в базис созданной в диссертации модели целевой архитектуры процессора выбрана модель целевой архитектуры процессора Лоу. Рассматриваются способы сведения сокращения времени выполнения тензорных операций к сокращению времени выполнения матрично-векторных операций.

Вторая глава посвящена описанию расширения модели целевой архитектуры процессора Лоу. Представлены алгоритмы для вычисления обобщений матрично-векторных операций на замкнутые полукольца. Описан алгоритм автоматического сокращения времени выполнения тензорных операций.

В третьей главе изложена архитектура программной системы автоматической оптимизации тензорных операций, и описаны ее возможные реализации.

Четвертая глава посвящена оценке эффективности разработанной программой системы автоматической оптимизации тензорных операций. Подтверждена применимость созданной программной системы. Рассмотрены эксперименты при решении обратной задачи гравиметрии, общей задачи о путях, оптимизации матрично-векторных операций и тензорных сверток.

В заключении сформулированы основные выводы и основные результаты диссертации, приведены рекомендации по использованию результатов диссертации и перспективы дальнейшей разработки темы.

**Замечания по диссертации.** Вместе с тем, диссертационная работа не свободна от недостатков. Основными недостатками диссертации являются:

1. В рамках диссертации разработаны модели и алгоритмы для сокращения времени выполнения тензорных операций на многоядерных процессорах общего назначения. Несмотря на это, существенной значимостью для проведения вычислительных экспериментов обладают специализированные процессоры такие как, например, Intel Xeon Phi. Как показывают приведенные в диссертации результаты экспериментов разработанные методы требуют доработки в случае специализированного процессора Intel Xeon Phi 7210.
2. Соискателем разработаны новые алгоритмы выполнения тензорных операций. С помощью моделирования выполнения созданных алгоритмов на представленном расширении модели целевой архитектуры процессора Лоу выведены формулы, позволяющие получить значения параметров указанных алгоритмов в зависимости от характеристик процессоров общего назначения. Представленное расширение и оригинальная модель целевой архитектуры процессора Лоу не рассматривают время доступа к данным, хранящимся в различных уровнях кэш-памяти.
3. Линейку процессоров Intel Xeon Phi перестали выпускать, однако, набор векторных команд перешел на серию процессоров Intel Xeon Scalable

первого и второго поколения. Было бы интересно проверить материалы диссертации на этих процессорах с учетом возможности программно включить и отключить понижение тактовой частоты в зависимости от количества задействованных ядер процессоров и векторных функциональных блоков.

4. На странице 23 термин «обрезку весов» нужно взять в кавычки. На странице 92 вместо слова «реализованную» нужно использовать «реализованной». Имеются опечатки в словах «программная» (стр. 103), «1970-х» (стр. 109). На странице 136 у второго графика, изображенного на Рис. 4.28, вместо «ГФлоп/сек» должно быть указано «Количество промахов».

Указанные замечания не снижают общей значимости и высокого научного уровня исследования, выполненного Р.А. Гареевым.

В автореферате обоснована актуальность исследования, его цели и задачи, научная новизна, практическая ценность и значимость научных результатов, выносимых на защиту. В целом автореферат полно и правильно отражает содержание диссертации.

**Заключение.** Диссертационная работа Гареева Романа Альбертовича «Методы оптимизации выполнения тензорных операций на многоядерных процессорах» представляет собой законченную научно-квалификационную работу. На основе модели и алгоритмов, предложенных в диссертации, решены задачи автоматического распараллеливания и сокращения времени выполнения тензорных операций на многоядерных процессорах. Указанные задачи имеют существенное значение в области создания программ и программных систем для параллельной обработки данных. Диссертационная работа соответствует всем требованиям, предъявляемым к диссертациям на соискание ученой степени кандидата наук согласно п. 9 «Положения о присуждении ученых степеней», утвержденного постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842 «О порядке присуждения ученых степеней» с изменениями

постановления Правительства Российской Федерации от 21 апреля 2016 г. № 335 «О внесении изменений в Положение о присуждении ученых степеней», а соискатель Гареев Р.А. заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

Официальный оппонент:

Я, Черных Игорь Геннадьевич, даю согласие на включение моих персональных данных в документы, связанные с работой Диссертационного совета ЮУрГУ 212.298.18 и их дальнейшую обработку.

Черных Игорь Геннадьевич  
« 02 » марта 2021 г.

Кандидат физико-математических наук,  
заведующий лабораторией суперкомпьютерного моделирования Федерального государственного бюджетного учреждения науки «Институт вычислительной математики и математической геофизики Сибирского отделения Российской академии наук», г. Новосибирск.

Адрес организации: 630090, Россия, г. Новосибирск, пр. академика Лаврентьева, д. 6

Телефон: +7 (383) 330-61-49

Email: chernykh@parbz.sccc.ru

Подпись И.Г. Черных удостоверяю:

Ученый секретарь ИВМИГ СО РАН

Кандидат физико-математических наук



/ Вшивкова Л.В./



### СВЕДЕНИЯ ОБ ОФИЦИАЛЬНОМ ОППОНЕНТЕ

диссертации Гареев Р.А. «Методы оптимизации выполнения тензорных операций на многоядерных процессорах» на соискание ученой степени кандидата физико-математических наук по специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей

Фамилия, имя, отчество	Игорь Геннадьевич Черных
Ученая степень (с указанием номера и шифра специальности)	Кандидат физико-математических наук, 05.13.11 Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей
Ученое звание	
Организация основного места работы	Федеральное государственное бюджетное учреждение науки «Институт вычислительной математики и математической геофизики Сибирского отделения Российской академии наук»
Ведомственная принадлежность	
Занимаемая должность	Заведующий лабораторией суперкомпьютерного моделирования
Почтовый адрес	630090, г. Новосибирск, пр. академика Лаврентьева, д. 6
Телефон	+7 (383) 330-61-49
Адрес электронной почты	chernykh@parbz.sccc.ru


### Список основных публикаций по теме диссертации в рецензируемых научных изданиях

1. Kulikov I., Chernykh I., Karavaev D., Berendeev E. et al. HydroBox3D: Parallel & Distributed Hydrodynamical Code for Numerical Simulation of Supernova Ia // Lecture Notes in Computer Science. 2019. Vol. 11657. P. 187-198. DOI: 10.1007/978-3-030-25636-4\_15.



2. Kulikov I., Chernykh I., Tchernykh A. A Scalable Parallel Computing Framework for Large-Scale Astrophysical Fluid Dynamics Numerical Simulation // 2019 20th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT) (Gold Coast, QLD, December, 05-07, 2019). Massachusetts, IEEE Xplore Digital Library, 2019. P. 328–333. DOI: 10.1109/PDCAT46702.2019.00066.
3. Mironov V., Chernykh I., Kulikov I., Moskovsky A. et al. Performance Evaluation of the Intel Optane DC Memory With Scientific Benchmarks // 2019 IEEE/ACM Workshop on Memory Centric High Performance Computing (MCHPC) (Denver, CO, November, 18, 2019). Massachusetts, IEEE Xplore Digital Library, 2019. P. 1–6. DOI: 10.1109/MCHPC49590.2019.00008.
4. Kulikov I., Chernykh I., Protasov V., Gubaydullin I. Simulation of formaldehyde formation during a galaxy collision using vectorized numerical method on Intel Xeon Phi accelerators // Journal of Physics: Conference Series. 2019. Vol. 1368, no. 4. P. 1–10. DOI: 10.1088/1742-6596/1368/4/042023.
5. Kulikov I., Chernykh I., Tutukov A. A new hydrodynamic code with explicit vectorization instructions optimizations that is dedicated to the numerical simulation of astrophysical gas flow. I. Numerical method, tests, and model problems // The Astrophysical Journal Supplement Series. 2019. Vol. 243, no. 1. P. 1–15. DOI: 10.3847/1538-4365/ab2237.
6. Kulikov I., Chernykh I., Sapetina A., Prigarin V. A new MPI/OpenMP code for numerical modeling of relativistic hydrodynamics by means adaptive nested meshes // Journal of Physics: Conference Series. 2019. Vol. 1336, no. 1. P. 1–5. DOI: 10.1088/1742-6596/1336/1/012008.
7. Kulikov I.M., Chernykh I.G., Tutukov A.V. A new parallel intel xeon phi hydrodynamics code for massively parallel supercomputers // Lobachevskii Journal of Mathematics. 2018. Vol. 39, no. 9. P. 1207–1216. DOI: 10.1134/S1995080218090135.
8. Chernykh I., Kulikov I., Glinsky B., Vshivkov V. et al. Advanced vectorization of PPML method for Intel® Xeon® scalable processors // Russian Supercomputing Days. 2018. Vol. 965. P. 465–471. DOI: 10.1007/978-3-030-05807-4\_39.
9. Глинский Б.М., Черных И.Г., Куликов И.М., Снытников А.В. и др. Интегральный подход к разработке алгоритмического и программного обеспечения экзафлопсных суперЭВМ: некоторые результаты // Марчуковские научные чтения-2017 (Новосибирск, 25 июня – 14 июля 2017 г.). Новосибирск: Институт вычислительной математики и математической геофизики Сибирского отделения РАН, 2017. С. 204–210.

10. Kulikov I., Chernykh I., Glinskiy B., Shmelev A. et al. Vectorization of astrophysical code for massively parallel supercomputers // Russian Supercomputing Days 2016 (Moscow, Russia, September, 26 –27, 2016). Moscow, Moscow State University, 2016. P. 52–60.

 / Черных И.Г. /

Подпись И.Г. Черных удостоверяю:

Ученый секретарь ИВМиМГ СО РАН

Кандидат физико-математических наук



 / Вшивкова Л.В. /