УДК 519.6

БЫСТРЫЙ МЕТОД МУЛЬТИПОЛЕЙ*

© Н.А. Гумеров,
Институт передовых компьютерных исследований,
Университет штата Мэриленд,
115 А.В. Уильямс Билдинг,
Колледж Парк, Мэриленд, 20742, США
эл. почта: gumerov@umiacs.umd.edu

Цель данной статьи – дать представление о быстром методе мультиполей, позволяющем качественно ускорить численное решение ряда задач, возникающих в различных областях физики и химии, при моделировании разнообразных природных явлений и технологических процессов, в задачах дизайна, компьютерной графики, обработке сигналов и многих других. Метод представляет собой масштабируемый алгоритм, позволяющий эффективно решать многомерные задачи с миллионами переменных на персональных компьютерах и на порядки большим числом переменных на вычислительных кластерах, включая гетерогенные архитектуры. Формулируются основные идеи метода, говорится о существующих и перспективных приложениях и приводится краткий обзор литературы.

Ключевые слова: быстрый метод мультиполей, иерархические алгоритмы, масштабируемые алгоритмы, высокопроизводительные вычисления, параллельные вычисления, численные методы, задача N тел

© N.A. Gumerov

FAST MULTIPOLE METHOD

Institute for Advanced Computer Studies,
University of Maryland
115 A.V. Williams Building,
College Park, Maryland 20742, USA
e-mail: gumerov@umiacs.umd.edu

The goal of this paper is to give an idea about the fast multipole method, which enables qualitative acceleration of numerical solution of a number of problems appearing in different areas of physics and chemistry, in modeling of various natural phenomena and technological processes, in design, computer graphics, signal processing, and many other fields. The method itself is a scalable algorithm suitable for efficient solution of multidimensional problems with millions of variables on personal computers and with orders of magnitude larger number of variables on computing clusters including heterogeneous architectures. The paper provides basic ideas of the method, mentions existing and prospective applications, and provides a brief literature review.

Key words: Fast multipole method, hierarchical algorithms, scalable algorithms, high performance computing, parallel computing, numerical analysis, N-body problem

Введение. Стремительное развитие вычислительной техники, начавшееся в конце XX в., положило начало компьютерному моделированию больших и сложных систем. Это стало необходимым для решения задач, требующих значительных ресурсов памяти и высокой скорости вычислений. Усилия специалистов в этой области постепенно привели к возникновению новой научной концепции: «открытие посредством вычислений», «вычислительный эксперимент», дополняющей традиционные методы (исследование, натурный эксперимент, построение теории). Вычислительный эксперимент играет особую

* Работа выполнена при поддержке гранта Министерства образования и науки РФ 11.G34.31.0040 (руководитель проекта И.Ш. Ахатов). Автор выражает благодарность профессору университета штата Северная Дакота И.Ш. Ахатову за плодотворные обсуждения ряда научных проблем, связанных с применением БММ, профессору университета штата Мэриленд Рамани Дураисуами за многолетнее плодотворное сотрудничество в разработке и применении БММ, а также студентам, аспирантам и сотрудникам университета штата Мэриленд и вычислительной лаборатории Центра микро- и наномасштабной динамики дисперсных сред при Башкирском государственном университете за энтузиазм и поддержку в реализации различных проектов, связанных с БММ.

роль в случае, когда непосредственное экспериментальное исследование является дорогостоящим, трудно осуществимым, или когда исследуемая система зависит от многих параметров. Например, при исследованиях процессов в микро- и наномасштабах.

На сегодняшний день высокопроизводительные вычисления составляют неотъемлемую часть химии, биологии, медицины, экономики, многих междисциплинарных исследований. Но более всего они задействованы в механике и физике. Нобелевская премия 2013 г. по химии подтверждает важность и значимость компьютерного моделирования в современной науке.

Прогресс в вычислительной практике и возникновение нового метода исследований на основе расчета больших систем (состоящей из миллионов или миллиардов дискретных элементов) был бы невозможен, если бы зависел только от аппаратного обеспечения. Однако новые потребности создают и новые подходы к решению таких задач.

Разработка и совершенствование масштабируемых алгоритмов и эффективных методов с низкой вычислительной сложностью и низким потреблением памяти является неотъемлемой частью вычислительного прогресса. Примером такого алгоритма является быстрое преобразование Фурье (БПФ).

Хорошо известно, что произведение плотной $N \times N$ матрицы на вектор требует $O(N^2)$ операций и $O(N^2)$ памяти для хранения матрицы. Однако в случае преобразования Фурье (т.е. в случае матрицы Фурье) можно обойтись $O(N\,logN)$ операциями и $O(N)$ памятью для получения результата если использовать алгоритм БПФ. Таким образом, если N порядка миллионов или миллиардов, то того же порядка будет и алгоритмическое ускорение (пусть на порядок меньше за счет logN), что позволяет решать качественно другие научные проблемы.

Другим, менее известным, особенно в России, алгоритмом перемножения плотной матрицы специального типа на вектор является Быстрый метод мультиполей (БММ, по-английски, FMM, Fast Multipole Method), входящий наряду с БПФ в десятку лучших алгоритмов XX в. [1]. Вычислительная сложность этого алгоритма также $O(N\,logN)$ или $O(N)$ и алгоритм требует $O(N)$ памяти. Однако в отличие от БПФ, БММ может применяться к значительно более широкому классу матриц, в частности, к матрицам возникающим при решении классических задач физики, статистики, искусственного интеллекта и т.д.

БММ был разработан Владимиром Рохлиным (Vladimir Rokhlin), эмигрировавшим в США в середине 70-х прошлого века и в настоящее время являющимся профессором Йельского университета, и его учеником Лесли Грингардом (Leslie Greengard), в настоящее время являющегося профессором Института Куранта при университете Нью-Йорка, и впервые опубликован в середине 80-х [2; 3].

В оригинальном изложении метод был разработан для расчета гравитационного (или электростатического) потенциала взаимодействия N тел в двух- и трехмерном пространстве. Однако весьма скоро стало понятно, что метод гораздо более универсален; благодаря усилиям первооткрывателей и десятков исследователей (в основном из США и Японии) метод был применен к решению задач теплопроводности, электродинамики, акустики, упругости, течений жидкости, статистической обработки, компьютерного видения и графики, машинного обучения, томографии, определения молекулярных структур. Кроме того, данный метод был обобщен на расширенный класс математических задач, связанных с интерполяцией, неравномерного преобразования Фурье, преобразования Гаусса и т.д.

В настоящее время литература, относящаяся к БММ насчитывает сотни наименований, и исчерпывающий ее обзор вряд ли возможен в данной статье. Здесь дается общее представление о методе и приводятся основные ссылки, из которых заинтересованный читатель может почерпнуть более детальную информацию.

Более десяти лет БММ находится в фокусе исследовательских интересов нашего коллектива, которые поддерживались грантами

Национального фонда науки и Национального аэрокосмического агенства США. В 2002 г. был разработан первый в мире курс *«Быстрый метод мультиполей: основы и приложения»*, который на протяжении почти десяти лет читался студентам, аспирантам и сотрудникам Университета штата Мэриленд в Колледж Парке.

Избранные лекции по этой теме прочитаны в 2011 г. аспирантам и сотрудникам Вычислительной лаборатории Центра микро- и наномасштабной динамики дисперсных сред при Башкирском государственном университете, организованного профессором Искандером Шаукатовичем Ахатовым при поддержке мегагранта Министерства образования и науки РФ.

Исследовательская и преподавательская работа по БММ тесно связана с моим коллегой по Институту передовых компьютерных исследований университета штата Мэриленд, профессором Рамани Дураисуами (Ramani Duraiswami), с которым в соавторстве было опубликовано более 40 работ и выступлений на международных конференциях, включая большую монографию [4].

ОСНОВНЫЕ ИДЕИ БММ

Собственно, БММ – это метод вычисления сумм типа

$$\phi_j = \phi(\mathbf{y}_j) = \sum_{i=1}^{N} q_i K(\mathbf{y}_j, \mathbf{x}_i), \quad j = 1,..., M, \qquad (1)$$

где $\boldsymbol{x}_i$ и $\boldsymbol{y}_j$ – так называемые источники и приемники, представляющие собой точки в d-мерном пространстве, q_i – интенсивности источников, K – заданная функция, называемая ядром. Очевидно, что вычисление таких сумм есть не что иное, как произведение матрицы с элементами K_{ji} на вектор с компонентами q_i, что в общем случае требует $O(MN)$ операций. Уменьшение этого числа становится возможным благодаря следующим основным идеям.

Факторизация. Обычно, курс БММ начинается с простого вопроса, можно ли вычислить сумму

$$\phi_j = \sum_{i=1}^{N} q_i (y_j - x_i)^2, \quad j = 1,..., M, \qquad (2)$$

за $O(M+N)$ операций? Ответ положительный и одна строчка дает решение:

$$\phi_j = \sum_{i=1}^{N} q_i (y_j^2 - 2x_i y_j + x_i^2) = c_0 y_j^2 - 2y_j c_1 + c_2,$$

$$c_0 = \sum_{i=1}^{N} q_i, c_1 = \sum_{i=1}^{N} q_i x_i, c_2 = \sum_{i=1}^{N} q_i x_i^2. \qquad (3)$$

Действительно, коэффициенты c_k могут быть вычислены за $O(N)$ операций независимо от приемников. Далее, нужно $O(M)$ операций, чтобы получить окончательный результат. В данном случае решение дала факторизация ядра, т.е. представление его в виде суммы произведений функций, зависящих только от источников и только от приемников. В более общем случае эта идея представляется так. Пусть ядро K факторизуемо в виде

$$K(\mathbf{y}, \mathbf{x}) = \sum_{n=1}^{\infty} C_n(\mathbf{x}) F_n(\mathbf{y}) = \sum_{n=1}^{P} C_n(\mathbf{x}) F_n(\mathbf{y}) + \varepsilon_P, \qquad (4)$$

где F_n – некоторые базисные функции и C_n – коэффициенты разложения. Предполагается, что ряд сходится и может быть усечен до первых P членов, при этом произойдет ошибка приближения ε_P. Подставив (4) в (1) и изменив порядок суммирования, не трудно видеть, что трюк (3) работает и в этом случае, и все φ_j можно вычислить за $O(PM+PN)$ операций. Если число усечения P гораздо меньше, чем M и N, например, $P = O(1)$, то $O(M+N)$ вычислительная сложность достигается и в этом случае. Если бы все этим ограничивалось, то метод быстрого суммирования был бы готов.

Контроль погрешности. Уже на этом этапе видно, что предлагаемый метод обладает погрешностью, и это — важная особенность БММ, в отличие, например, от БПФ. С точки зрения математика, неточность метода является неким дефектом. Однако с практической точки зрения, возможность ускорить вычисления за счет погрешности, которая может быть строго контролируемой, вероятно, является положительной чертой.

Более того, оказывается, что ошибку метода можно сделать, например, меньше точности машинных вычислений с двойной точностью. В этом случае результаты, полученные с помощью БММ, не менее точны, чем, скажем, вычисления синуса или квадратного корня. Примечательно, что из-за ошибок округления результаты, полученные с помощью БММ для больших матриц, могут быть более точными по сравнению с непосредственным суммированием (метод «грубой силы»).

БММ является спектральным методом, используемые ряды достаточно быстро сходятся, и обычно $P = O(\log\varepsilon_p)$. В зависимости от нормы контроля погрешности P может быть постоянным или увеличиваться с $N \sim M$ как $\log N$. В любом случае возникает концепция обмена точности вычислений на скорость.

Мультипольные и локальные разложения. Основной же проблемой, из-за которой простой глобальной факторизации недостаточно и метод должен быть усложнен, является неравномерная сходимость сумм (4), что характерно для сингулярных ядер, таких, как функции Грина уравнений Лапласа и Гельмгольца в трех измерениях:

$$K(\mathbf{y},\mathbf{x}) = \frac{1}{4\pi|\mathbf{y}-\mathbf{x}|}, \quad K(\mathbf{y},\mathbf{x}) = \frac{\exp(ik|\mathbf{y}-\mathbf{x}|)}{4\pi|\mathbf{y}-\mathbf{x}|} \quad . (5)$$

В последнем случае k – волновое число и ядро комплексно ($i^2 = -1$).

Идея факторизации в этом случае спасается посредством построения разложений для различных областей, в которых обеспечивается сходимость. В частности, вместо одной области и одного глобального разложения (4) рассматривается много областей. Для каждой рассматриваемой области пространства с центром $\mathbf{x}^*$ строится два разложения (ограничимся случаем ядер зависящих от $\mathbf{y} - \mathbf{x}$):

Здесь первое разложение, называемое мультипольным (сокращенно, *M*-или *S*- разложением), справедливо в бесконечной области удаленной от центра разложения, в то время как второе разложение, называемое локальным (сокращенно, *L*- или *R*- разложением), справедливо в конечной области около центра (примером является разложение функций в ряд Тейлора). Базисные функции S_n и R_n выбираются из соображений удобства и быстроты схождения рядов. Если БММ применяется для решения уравнений математической физики, то обычно эти функции удовлетворяют этим уравнениям, например, Лапласа или Гельмгольца. Типичными для трехмерного пространства являются сферические базисные функции, пропорциональные сферическим гармоникам, образующим полную ортогональную систему на сфере.

Разбиение вычислительного пространства. Теперь возникает вопрос, как выбрать тип разложения, центры разложений и соответствующие области, чтобы покрыть все вычислительное пространство. В БММ это делается просто. Все вычислительное пространство заключается в куб (для конкретности рассмотрим трехмерный случай), который затем разбивается на кубики или ящички меньшего размера. Кубики, содержащие источники или приемники, соответственно называются кубами-источниками или кубами- приемниками.

Ясно, что могут присутствовать кубики, одновременно содержащие и источники, и приемники, так и не содержащие интересующих точек, или «пустые». В БММ, который является адаптивным алгоритмом, пустые кубики полностью исключаются из рассмотрения, что весьма ускоряет вычисления для неравномерных распределений (например, для распределений источников и приемников по

$$\begin{aligned} K(\mathbf{y}-\mathbf{x}) &= \sum_{n=1}^{P} C_n(\mathbf{x}-\mathbf{x}_*)S_n(\mathbf{y}-\mathbf{x}_*)+\varepsilon_P, \quad |\mathbf{y}-\mathbf{x}_*| > \alpha|\mathbf{x}-\mathbf{x}_*|, \quad \alpha>1, \\ K(\mathbf{y}-\mathbf{x}) &= \sum_{n=1}^{P} D_n(\mathbf{x}-\mathbf{x}_*)R_n(\mathbf{y}-\mathbf{x}_*)+\varepsilon_P, \quad |\mathbf{y}-\mathbf{x}_*| < \beta|\mathbf{x}-\mathbf{x}_*|, \quad \beta<1. \end{aligned} \qquad (6)$$

двумерной поверхности в трехмерном пространстве, что характерно для краевых задач).

Для каждого источникового кубика строится M-разложение. Это разложение описывает эффект всех источников в кубике на любой приемник, находящийся вне окрестности этого куба-источника. Действительно, если представить, что источник находится близко к углу своего кубика с центром $\mathbf{x}^*$, то для применимости M-разложения (6) нужно, чтобы приемник находился дальше от центра, что не выполняется для всех точек соседнего кубика.

Таким образом, с помощью факторизации можно ускорить суммирование вкладов источников, только «хорошо разделенных» с приемниками, вклады же «плохо разделенных» источников приходится вычислять, что называется, в лоб (прямым суммированием).

Как бы там ни было, но, оказывается, что даже оптимизируя разбиение, в общем случае нельзя добиться искомой $O(N)$ вычислительной сложности алгоритма (здесь и далее для простоты полагается $M \sim N$), и максимум, на что можно расчитывать, — это $O(N^{3/2})$. Трансляция разложений, описанная ниже, позволяет улучшить асимптотику до $O(N^{4/3})$, а ее сочетание с иерархической структурой позволяет достичь $O(N)$.

Трансляция разложений. Если у нас имеется M- или L-разложение справедливое в области Ω с центром $\mathbf{x}^*$, а нам нужно иметь разложение, применимое в ее подобласти Ω' с центром $\mathbf{x}'^*$, вовсе не обязательно заново вычислять коэффициенты разложения, используя координаты и интенсивности источников. Для этого достаточно применить оператор трансляции $T(\mathbf{x}'^*-\mathbf{x}^*)$ к коэффициентам известного разложения, чтобы получить коэффициенты искомого разложения.

Оператор трансляции является линейным, и если речь идёт о нахождении P неизвестных коэффициентов по имеющимся P данным, оператор трансляции может быть представлен $P\times P$ матрицей. Умножение матрицы на вектор «стоит» $O(P^2)$ операций, хотя, в связи со специальной структурой этих матриц для уравнений Лапласа и Гельмгольца, трансляция может быть выполнены за $O(P^{3/2})$ и даже за $O(P)$ операций.

Последнее достигается при использовании базисов, в которых операторы трансляции представляются диагональными матрицами. К сожалению, такие базисы бесконечномерны, и корректные операции усечения с контролем точности привносят дополнительную сложность, поэтому, например, для решения уравнения Лапласа с умеренной точностью $O(P^{3/2})$ оказываются более практичными.

В многоуровневом (иерархическом) БММ встречается три типа операторов трансляции, *M2M, M2L, и L2L,* где первая буква означает тип начального разложения, а вторая — тип конечного разложения.

Иерархический алгоритм. Как отмечалось выше, простого разбиения пространства недостаточно для достижения алгоритмической сложности $O(N)$. В трехмерном случае используется разбиение с помощью восьмиричного дерева. Начальный куб, которому присваивается уровень 0, делится на 8 дочерних кубов в два раза меньшего размера, которым присваивается уровень 1.

Процесс восьмеричного деления продолжается до некоторого уровня l_{max}, определяемого из соображений оптимизации. На каждом уровне определяются кубики приемников и источников, соседи, а также «дети» и «родители» кубиков. M-разложения на уровне $l < l_{max}$ получаются путем сложения результатов $M2M$ трансляций разложений с уровня $l+1$ (этот процесс инициализируется прямым вычислением коэффициентов на уровне l_{max}). L-разложения находятся более изощренно — каждый кубик получает информацию о дальних источниках посредством $L2L$ трансляции коэффициентов L-разложения для родительского куба.

К этой информации добавляется информация об источниках в промежуточной области, лежащих вне окрестности кубика, но внутри окрестности его родителя. Осуществляется это посредством сложения результатов $M2L$ трансляций. Здесь можно было бы употребить слово «фрактал», поскольку структура

областей повторяется на каждом уровне иерархии. Таким образом, алгоритм состоит из «восхождения» по иерархии источников от уровня lmax до верхнего уровня (оказывается, достаточно $l = 2$), когда каждый куб-источник обеспечивается M-разложением, и «нисхождения» по иерархии приемников от верхнего уровня $l = 2$ до l_{max}, когда каждый куб-приемник обеспечивается L-разложением. L-разложение учитывает вклад в сумму всех источников за исключением источников в непосредственной окрестности куба-приемника.

Их вклад, как отмечалось выше, учитывается прямым суммированием.

Рисунок 1 иллюстрирует этапы алгоритма для двух измерений.

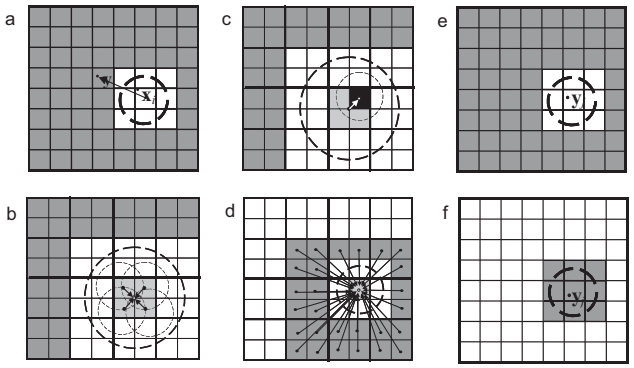

Рис. 1. Этапы алгоритма БММ: a) построение M-разложений; b) M2M-трансляции; c) L2L-трансляции; d) M2L-трансляции; e) оценка L-разложений; f) прямое суммирование в окрестности

Структура данных. Требуемая быстрота алгоритма предполагает быстрые и эффективные методы построения нетривиальной структуры данных БММ, требующей адаптивных восьмеричных деревьев для источников и для приемников с определением соседей и т.п. Основу этих методов составляет так называемая кривая, заполняющая пространство, которая в определенной последовательности проходит через все кубики иерархии (индексация Мортона).

На самом деле, эта индексация ведет начало от известной задачи Пеано, в XIX в. построившего взаимно-однозначное отображение точек d-мерного куба на отрезок прямой. Практический алгоритм воплощает идею Пеано путем масштабирования вычислительной области и операции битового перемежения декартовых координат источников (то же самое для иерархии приемников). Особенностью индексации является возможность быстрого определения «родственников» и «соседей» кубика. Кроме того, индексация позволяет ввести упорядочивание трехмерных (многомерных) данных, следовательно, осуществлять сортировку, применять быстрые алгоритмы поиска, и т. д.

Теоретически построение структуры данных занимает $O(N \log N)$ операций. Однако на практике эта подготовительная часть БММ гораздо быстрее, чем, собственно, его исполнительная часть с формальной сложностью $O(N)$ и большой асимптотической константой. Кроме того, оказывается, что размеры памяти обычных современных компьютеров таковы, что неполную сортировку, достаточную для БММ, реально осуществить за $O(N)$ операций.

БММ ДЛЯ РАЗНЫХ ЯДЕР

Публикация БММ вызвала бурную реакцию в среде исследователей, занимающихся вычислительной физикой, численными методами, параллельными и высокопроизводительными вычислениями. Незадолго до БММ был опубликован иерархический алгоритм Барнса и Хата (Burnes & Hut [5], «Tree-code»), существенно ускоряющий расчеты динамики звездных скоплений и идеологически близкий к БММ.

Однако, в отличие от БММ, этот алгоритм осуществляет группировку источников по другому принципу и не имеет строгого контроля точности. Естественно, что первыми БММ заинтересовались астрофизики. Расчет кулоновских взаимодействий в молекулярной динамике (МД) происходит по той же схеме, поэтому БММ начал цитироваться в МД литературе практически сразу после своего возникновения.

Численные методы в целом. Осознание БММ как метода гораздо более общего, чем метода суммирования гравитационных или

кулоновских потенциалов, происходило в основном в 90-е годы. Высококачественные интерполяция, дифференцирование и интегрирование одномерных функций, заданных значениями в неравномерно распределенных точках, которые необходимы для успешной реализации многих численных методов, оказались естественными задачами для решения с помощью БММ [6]. Основу здесь играет матрица Коши с ядром $K(y,x) = (y - x)^{-1}$ и степенные ряды, для которых операторы трансляции схожи с соответствующими операторами для уравнений Лапласа в двух и трех измерениях.

Неравномерное преобразование Фурье. Суммирование с периодическим аналогом этого ядра, $K(y,x) = \mathrm{ctg}[(y - x)/2]$, возникает в алгоритме так называемого неравномерного быстрого преобразования Фурье (НБПФ) и успешно осуществляется с помощью БММ [7].

Хорошо известно, что широко распространенное БПФ применимо только к данным на равномерных сетках. Создание эффективного с вычислительной точки алгоритма НБПФ сулит существенный прогресс для технологий, связанных с обработкой сигналов, например, для компьютерной томографии, а также для вычислительных методов, оперирующих с неструктурированными данными. Последующие работы, посвященные НБПФ, но на основе спектрально аккуратной интерполяции с гауссовым ядром, $K(y,x) = \exp(-|y-x|^2)$ [8], подтверждают актуальность темы.

Преобразование Гаусса. БММ для гауссового ядра в многомерном пространстве, вообще, заслуживает особого разговора. Во-первых, ряд важных физических задач сводится непосредственно к суммированию одно-, двух- или трехмерных гауссиан, которые связаны с фундаментальными решениями задачи теплопроводности. Во-вторых, гауссианы также являются «размазанными» дельта-функциями, что позволяет аппроксимировать непрерывные поля дискретно распределенными источниками, а также находить осредненные параметры неструктурированных данных. Преобразование Гаусса применяется не только для решения задач физики, но и во многих других областях, например, в компьютерном видении для сегментации изображений (см. рис. 2) (здесь для трехмерных цветных изображений размерность пространства увеличивается до шести), в статистической обработке данных и в обучении компьютеров, где размерность пространства может быть очень велика. Стандартные версии БММ для гауссового ядра (быстрое преобразование Гаусса) опубликованы в начале 90-х [9], однако существенное ускорение, особенно для пространств большой размерности, было достигнуто в 00-х [10], что связано с более компактным представлением функций и использованием разбиения пространства с помощью известного в вычислительной геометрии алгоритма k-центров.

РБФ-интерполяция. В этом контексте уместно упомянуть важность БММ для задач многомерной интерполяции на неструктурированных множествах, осуществляемых с помощью радиальных базисных функций (РБФ). РБФ-интерполяция считается качественной, например, она признана приемлемой для алгоритмов, используемых в медицинской практике. Кроме того, ускоренный с помощью БММ метод РБФ может использоваться для неявного представления и визуализации поверхностей сложной геометрии и топологии, что является важной задачей компьютерной графики [11], а также для компрессии и восстановления изображений (см. рис. 2). РБФ-интерполянт – это не что иное, как ряд (1), где ядро K зависит только от расстояния между источником и приемником, $r = |\mathbf{y} - \mathbf{x}|$.

Ядра (5) являются примерами РБФ, однако практически не используются. Распространенными функциями являются ядра с глобальным носителем $K(r) = r$, $K(r) = (r^2 + c^2)^{1/2}$ и с компактным носителем (включая гауссово распределение). Основная проблема здесь в уравнении (1) определить коэффициенты разложения q_i по данным φ_j, или, что то же самое, решит большую линейную систему.

Имея основанный на БММ быстрый метод матрично-векторного умножения, естественно использовать итеративные методы решения. Проблемой здесь является плохая

обусловленность матрицы системы, что влечет большое количество итераций. Этот дефект можно исправить, используя эффективные предобуславливатели, построение которых является нетривиальной, но выполнимой задачей, решение которой приводит к $O(N \log N)$ алгоритму РБФ-интерполяции [12].

Рис. 2. Использование БММ в компьютерной графике: a) сегментация цветного изображения с помощью преобразования Гаусса [10]; b) восстановление изображения с помощью РБФ-интерполяции [12] (изображение в центре получено из изображения справа (Леонардо да Винчи, портрет Джиневры Бенчи, Национальная картинная галерея, Вашингтон) при потере 86% информации; изображение слева восстанавливает картину интерполяции данных с изображения в центре)

Уравнение Гельмгольца. Пожалуй, наиболее обширная литература по БММ посвящена уравнению Гельмгольца и связанным с ним уравнениям и системам, например, уравнениям Максвелла, описывающим распространение электромагнитных волн (включая монографии [4; 13]). Разработка методов ультразвуковой томографии, качественного воспроизведения аудио реальности, дизайна мобильных телефонов, бесшумного интерьера автомобилей или военных самолётов-«невидимок» связана с решением задач акустического или электромагнитного рассеяния в сложных трехмерных областях. И оно может быть существенно ускорено с помощью БММ. Принципиальным отличием уравнения Гельмгольца от уравнения Лапласа является наличие «внутреннего» масштаба длины (длины волны). Это приводит к тому, что количество членов в разложениях функций зависит не только от точности, которую желательно достигнуть, но и размера области, для которой такое внешнее или внутреннее разложение строится. Это приводит к модификации БММ, связанной с изменением числа P для каждого уровня пространственной иерархии. Более того, оказывается, что для достижения $O(N \log N)$ сложности алгоритма операторы трансляции должны быть высокоэффективны, и сложность одной трансляции не может быть больше чем $O(P^{3/2})$ для объемных распределений источников и приемников и $O(P \log^{\alpha} P)$ для поверхностных распределений [4]. Последнее оказалось возможным после публикации работы Рохлина [14], где построены диагональные формы операторов трансляции сложности $O(P)$. Учитывая, что для многоуровневого БММ процедура контроля точности может быть реализована за $O(P \log P)$ операций с помощью алгоритма «быстрого сферического фильтра» [15], опять-таки осуществляемого с помощью одномерного БММ для ядра Коши, $O(N \log N)$ сложность БММ для трехмерного уравнения Гельмгольца может быть достигнута. Практическая реализация алгоритмов может быть найдена в [16; 17].

Независимый от ядра БММ. Весьма условно, для большинства задач, имеющих практический интерес, можно выделить два типа ядер K в (1): неосциллирующие и осциллирующие. К первому типу относится ядро Лапласа, ко второму ядро Гельмгольца. В случае, когда возникает задача с ранее неисследованным ядром, проявляется один из недостатков БММ: необходимость существенной теоретической работы по построению базисных функций, операторов трансляций, оценки погрешности и т. п., как, например, для уравнений Стокса [18]. Попытка устранить этот дефект для неосциллирующих ядер предпринята в [19], где предложен алгоритм НЯ БММ

(независимый от ядра БММ, от "KI FMM – Kernel Independent FMM"). Следует, однако, отметить, что, несмотря на $O(N)$ сложность, присущую БММ, и избавление разработчика от большой теоретической работы, производительность НЯ БММ, как правило, оказывается ниже, чем у БММ, специализированного для определенного ядра (например, Лапласа), что связано с использованием более эффективных методов трансляции в специальных случаях, не допускающих обобщения. НЯ БММ для осциллирующих ядер был предложен недавно в [20].

Ядра других физических задач. Среди других подходов для построения БММ с ядрами для задач, имеющих практический интерес, можно выделить два, которые можно назвать методом прямой факторизации и методом скалярных потенциалов. Оба метода сводят задачу суммирования с нужным ядром к нескольким задачам суммирования с известными ядрами, если такое возможно. Первый метод проще, второй, возможно, более эффективен. Так, например, суммирование фундаментальных решений бигармонического уравнения в первом методе сводится к пяти суммированиям решений уравнения Лапласа [21], а во втором – к двум [22], для уравнений Стокса эти числа равны четырем [23; 24] и трем, соответственно. Последний метод также успешно применялся для компактного представления и быстрого построения решения векторных уравнений Лапласа для солеиноидальных полей [25] и Максвелла [26].

ПРИМЕНЕНИЕ БММ ДЛЯ РЕШЕНИЯ УРАВНЕНИЙ В ЧАСТНЫХ ПРОИЗВОДНЫХ

Трактовка БММ как метода ускоренного матрично-векторного произведения может вызывать вопрос, каким же образом задачи физики и механики, сформулированные в терминах уравнений в частных производных, сводятся к проблеме решаемой БММ? Кроме того, иногда можно услышать ошибочное мнение, что БММ пригоден только для решения линейных задач. Среди разнообразия подходов здесь можно выделить два следующих.

Лагранжевы методы. При использовании лагранжевого подхода непрерывная система моделируется как эволюционирующая совокупность дискретных объектов (частиц). Частицы могут занимать произвольное положение в пространстве и образуют неструктурированное множество источников, что отличает лагранжевы методы от эйлеровых, где пространство обычно дискретизируется с помощью равномерных или неравномерных сеток или конечных элементов. Поля частиц и их взаимодействие могут быть расчитаны на каждом шаге по времени при помощи БММ. Для ряда задач лагранжевы методы могут быть существенно быстрей и точней, чем эйлеровы, а также потреблять гораздо меньший объем машинной памяти.

Хорошим примером использования лагранжевого подхода является решение существенно нелинейных уравнений Навье-Стокса методом вихревых элементов. Здесь движение жидкости представляется как движение N вихрей в поле скоростей, собственно, создаваемом этими вихрями (источниками). Это поле является суперпозицией N полей, и его расчет в центре каждого элемента является наиболее трудоемкой процедурой метода. БММ позволяет выполнить эту процедуру за $O(N)$ операций, вместо $O(N^2)$, и тем самым ускорить все решение в $O(N)$ раз (а N может быть порядка миллионов и больше). Эффективная версия БММ для этого случая описана в недавно опубликованной работе [25].

Методы граничных элементов. Эти методы (МГЭ) применяются для решения краевых задач в областях со сложной границей и основаны на представлении решения в форме поверхностных интегралов (например, потенциалов простого и двойного слоя). Поверхность здесь дискретизируется N элементами (обычно сеткой), по которым приближенно вычисляются интегралы, что сводит интегральные уравнения к системе линейных уравнений. Несмотря на хорошее качество численного решения, применение МГЭ к задачам, требующим высокой дискретизации границы, до недавнего

времени было практически невозможным. Это связано с тем, что основная вычислительная сложность метода заключается в решении $N \times N$ алгебраической системы с плотной матрицей, что требует $O(N^3)$ операций и $O(N^2)$ памяти. Использование итеративных методов с хорошо подобранным предобуславливателем позволяет добиться $O(N_{it}N^2)$ сложности с относительно небольшим числом итераций N_{it}. Однако этого недостаточно для решения задач с N порядка сотен тысяч, или миллионов.

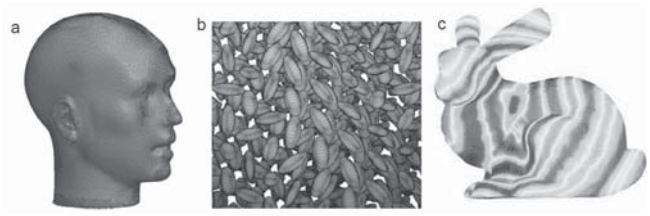

Рис. 3. Примеры задач, решенных с помощью метода граничных элементов, ускоренных с помощью БММ: a) сетка высокого разрешения (N > 105), моделирующая голову человека для расчета восприятия звука в диапазоне до 20 кГц [27]; b) фрагмент дисперсной системы (каждая частица покрыта сеткой с N ~ 103, число частиц ~ 103) [29]; c) расчитанное акустическое давление на поверхности объекта в задаче рассеяния (сетка с N > 105) [17]. N – число граничных элементов

Для современных технологий такие расчеты оказываются важными. Например, при моделировании формы головы человека, что необходимо для задачи акустической ориентации в пространстве, N может составлять сотни тысяч (см. рис. 3). Еще большие значения нужны для моделирования акустики концертных залов или рассеяния высокочастотных электромагнитных волн от самолетов, поскольку размер элементов в МГЭ должен быть существенно меньше длины волны. Другим примером является расчет динамики капельных систем, где несмотря на то, что поверхность каждой деформируемой капли можно дискретизировать, скажем, ста элементами, - число капель в системе может быть огромным (см. рис. 3). И для таких больших задач БММ воистину революционизирует МГЭ, поскольку позволяет получить решение за $O(N_{it}N)$ операций, используя минимум памяти ($O(N)$). Интереснымздесьявляется,чтосамБММможет использоваться для эффективного предобуславливания. Это основано на свойстве БММ, что суммирование можно ещё больше ускорить за счет точности. Для предобуславливания высокая точность не требуется и за время одной «внешней» итерации можно осуществить десяток «внутренних» итераций, существенно улучшающих сходимость. Более детально с использованием БММ в МГЭ можно ознакомиться в публикациях [17; 23; 28; 29].

БММ И ВЫСОКОПРОИЗВОДИТЕЛЬНЫЕ ВЫЧИСЛЕНИЯ

Рассказ о БММ был бы далеко не полным, если бы не была упомянута связь метода с высокопроизводительными вычислениями, проводимыми на устройствах с параллельной архитектурой (многоядерные процессоры, вычислительные кластеры, графические карты, гетерогенные архитектуры и т.д.). Практически сразу после создания БММ стало ясно, что метод обладает очень высоким параллелизмом [30]. В частности, несмотря на то, что серийный (непараллельный) алгоритм БПФ на практике в несколько раз быстрее, чем БММ для матриц того же размера, для больших параллельных систем БММ может оказаться быстрее, что связано с меньшим числом и более простой топологией коммуникаций между процессорами.

Пожалуй, наибольшей по величине проблемой N тел, решенной с помощью БММ, на настоящее время является проблема с $N \sim 3 \cdot 10^{12}$, для чего в 2011 г. было задействовано порядка $3 \cdot 10^5$ ядер суперкомпьютера JUGENE, построенного IBM в Юлихе, Германия [31]. Этот рекорд может быть вписан в книгу рекордов Гиннесса, однако вряд ли имеет большое практическое значение, поскольку использовалось очень низкое $P = 4$, что говорит о плохой точности. Кроме того, производительность для такой большой системы оказалась низкой, 64 Tflops, а время расчета большим (порядка 1000 секунд). Все это исключает практическое использование подобной реализации для задач, требующих многих тысяч шагов по времени.

Производительность 700 Tflops, что является уже пета-масштабом, была достигнута при использовании НЯ БММ при решении задачи о течении крови ($2 \cdot 10^8$ деформируемых красных кровяных телец, $9 \cdot 10^{10}$ неизвестных) на суперкомпьютере Jaguar PF (Оак Ридж, США) [32], состоящем из ~ $2 \cdot 10^5$ ядер. Один временной шаг здесь занял 300 секунд, что позволило сделать несколько шагов.

Для практических задач более эффективным является подход, основанный на использовании графических процессоров. Первая реализация БММ на персональном компьютере с одной высокопроизводительной графической картой (NVIDIA 8800 GTX) была осуществлена нами в Мэриленде, США, в 2007 г. [33] и вызвала большой резонанс после детальной публикации [34]. Здесь одно матрично-векторное произведение для задачи N-тел с $N \sim 10^6$ выполняется с хорошей точностью ($P \sim 100$) за время порядка 1 секунды, что соответствует производительности примерно 200 Gflops. Это делает осуществимым решение достаточно больших эволюционных задач на гораздо более дешевом и доступном аппаратном обеспечении.

Дальнейшее развитие этот подход получил в работе исследовательской группы из Японии, удостоенной наиболее престижной премии за высокопроизводительные вычисления (Gordon Bell Prize) в 2009 г. [35]. Здесь БММ был применен для $N \sim 10^7$ и задействован кластер, состоящий из 128 центральных процессоров и 256 графических карт, и была достигнута производительность 28-42 Tflops. К сожалению, алгоритм метода вихревых элементов, использованный группой, был не очень эффективен и расчет для $N \sim 10^7$ ($P = 100$) занял порядка 1 с, что достаточно много для такого кластера. Подобный кластер (Lincoln, США, 256 графических карт) был использован для расчета динамики крови, где для $N \sim 2 \cdot 10^8$ алгоритм НЯ БММ потребовал порядка 2 с на один шаг по времени [32].

Наша исследовательская группа в Мэрилендском университете в 2011 г. разработала и опубликовала алгоритм БММ для гетерогенной архитектуры, или гетерогенного кластера, каждая нода которого состоит из многоядерного центрального процессора и нескольких графических карт [36]. В этом алгоритме центральные процессоры выполняют часть алгоритма БММ, связанную с трансляциями, а графические карты — прямое суммирование и построение структуры данных. Поскольку эти процессы в основном проходят параллельно, достигается высокая задействованность всего кластера, ускорение вычислений и экономия ресурсов. Испытания на CPU/GPU parallel region кластере Chimera, состоящем из 32 гетерогенных нод (8 ядер и две графических карты на ноду), показали существенное ускорение (примерно на порядок) по сравнению с [35], несмотря на несколько меньшую производительность 13—38 Tflops. Этот подход позволил достичь выполнения одного шага по времени для системы из $N \sim 10^9$ частиц за 20 с. Кроме того, здесь достигнут наилучший показатель экономической эффективности, 177 Mflops/dollar (количество операций в секунду в расчете на доллар, затраченный на покупку оборудования). В дальнейшем алгоритм был усовершенствован, что сократило время в два раза [37] (см. рис. 4).

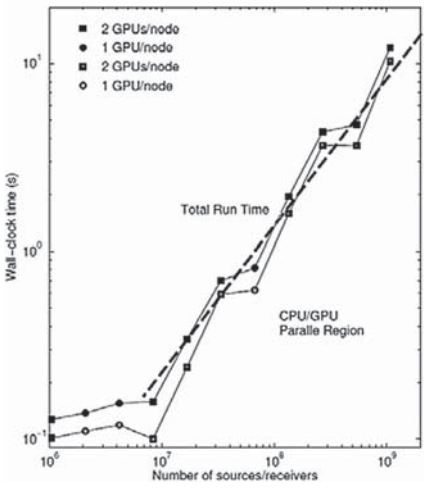

Рис. 4. Производительность гетерогенного БММ алгоритма, реализованного на кластере Chimera (32 ноды) Института передовых компьютерных исследований университета штата Мэриленд (зависимость времени решения задачи N тел от N для одного временного шага) [37]. Показано общее время решения для одного и двух графических процессоров на ноду и время, которое центральные и графические процессоры работают параллельно. Штриховая линия показывает линейную зависимость

Пожалуй, самая высокая опубликованная производительность, 1.08 Pflops, при использовании БММ для расчета изотропной турбулентности методом вихревых элементов была достигнута на японском кластере Цубами (TSUBAME-2.0), состоящем из 1408 нод с тремя графическими процессорами NVIDIA M2050 на ноду (всего 4224) [38]. Расчет одного шага по времени для $N \sim 7 \cdot 10^8$ здесь занял 108 с (заметим, что это лучше, чем 156 с требуемого для спектрального метода, основанного на БПФ). Можно отметить невысокую эффективность этой конкретной реализации и лучшие перспективы для БММ, масштабирование результатов [37] дает в несколько раз меньшее время.

Несмотря на упомянутые достижения, список которых неполон и не включает результатов, доложенных на последней конференции SC'13, можно отметить, что в настоящее время расчеты с $N > 10^8$ носят, скорее, демонстрационный характер, показывающий масштабируемость и потенциал БММ. Для исследовательской же работы в области физики, механики, химии, биологии, и т. д., связанной с моделированием больших систем, более эффективным инструментом являются персональные суперкомпьютеры, состоящие из многоядерных центральных и нескольких графических процессоров, а также мини-кластеры, состоящие из 2—4 подобных нод, позволяющие более удобное взаимодействие исследователя с компьютером. Такие системы обладают производительностью 1—8 Tflops и позволяют эффективно решать динамические задачи с $N \sim 10^5$—10^8. Во многих случаях этого вполне достаточно для получения научной информации. Примером такого подхода являются исследования, проводимые в Центре микро- и наномасштабной динамики дисперсных сред в Уфе, ряд из которых опубликован в России и за рубежом (последние публикации [39—42]).

Заключение. Быстрый метод мультиполей является современным численным методом, позволяющим качественно ускорить решение ряда основных задач математической физики, обработки сигналов, статистики, компьютерного видения и графики, а также применяться в других областях знаний, связанных с моделированием больших систем и обработкой больших потоков информации. БММ представляет собой алгоритм, хорошо параллелизуемый в разных масштабах, что позволяет эффективно использовать современные вычислительные архитектуры для аппаратного ускорения. Развитие и совершенствование БММ как алгоритма не завершено и представляет собой «горячую» тему для исследования. Его развитие связано как с теоретическими разработками в области представления функций и операторов, линейной алгебры, вычислительной геометрии, структур данных, и т.п., так и с развитием вычислительной техники, появлением новых вычислительных архитектур.

REFERENCES

1. Dongarra J.J., Sullivan F. The top 10 algorithms. *Comput. Sci.* Eng. 2000, vol. 2, no. 1, pp. 22–23.

2. Greengard L., Rokhlin V. A fast algorithm for particle simulations. *J. Comput.* Phys.1987, vol. 73, no. 2, pp. 325–348.

3. Greengard L. The Rapid Evaluation of Potential Fields in Particle Systems. Cambridge:MIT Press, 1998. 106 p.

4. Gumerov N.A., Duraiswami R. Fast Multipole Methods for the Helmholtz Equation in Three Dimensions. Oxford: Elsevier, 2005, 525 p.

5. Barnes J., Hut P. A hierarchical O(N log N) force-calculation algorithm. Nature. 1986, vol. 324, no. 4, pp. 446–449.

6. Dutt A., Gu M., Rokhlin V. Fast algorithms for polynomial interpolation, integration, and differentiation. *SIAM J. Numer. Analysis.* 1996, vol. 33, no. 5, pp. 1689–1711.

7. Dutt A., Rokhlin V. Fast Fourier transforms for nonequispaced data, II. *Appl. Comput. Harmon. Anal. Comput.* 1995, vol. 2, no. 1, pp. 85–100.

8. Greengard L., Lee J.–Y. Accelerating the nonuniform fast Fourier transform. *SIAM Review.* 2004, vol. 46, no. 3, pp. 443–454.

9. Greengard L., Strain J. The fast Gauss transform. SIAM J. *Sci. Stat. Comput.* 1991, vol.12, no. 1, pp. 79–94.

10. Yang C., Duraiswami R., Gumerov N.A., Davis L.S.

Improved fast Gauss transform. In: Proceedings of the International Conference on Computer Vision (ICCV), Nice, France, October, 2003, pp. 464–471.

11. Carr J.C., Beatson R.K., Cherrie J.B., Mitchell T.J., Fright W.R., McCallum B.C., Evans T.R. Reconstruction and representation of 3D objects with radial basis functions. In: Proceedings of the International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH), Los Angeles, CA, USA, August, 2001, pp. 67–76.

12. Gumerov N.A., Duraiswami R. Fast radial basis function interpolation via preconditioned Krylov iteration. SIAM *J. Sci. Comput.* 2007, vol. 29, no. 5, pp. 1876–1899.

13. Chew W.C., Jin J.M., Michielssen E., Song J. Fast and Efficient Algorithms in Computational Electromagnetics. Boston: Artech House, 2001. 950 p.

14. Rokhlin V. Diagonal forms of translation operators for the Helmholtz equation in three dimensions. *Appl. Comput. Harmon. Anal.* 1993, vol. 1, no. 1, pp. 82–93.

15. Jacob-Chien R., Alpert B.K. A fast spherical filter with uniform resolution. *J. Comput. Phys.* 1997, vol. 136, no. 2, pp. 580–584.

16. Cheng H., Crutchfield W.Y., Gimbutas Z., Greengard L., Ethridge J. F., Huang J., Rokhlin V., Yarvin N., Zhao J. A wideband fast multipole method for the Helmholtz equation in three dimensions. *J. Comput. Phys.* 2006, vol. 216, no. 1, pp. 300–325.

17. Gumerov N.A., Duraiswami R. A broadband fast multipole accelerated boundary element method for the 3D Helmholtz equation. *J. Acoust. Soc. Am.* 2009, vol. 125, no. 1, pp.191–205.

18. Sangani A.S., Mo G. An O(N) algorithm for Stokes and Laplace interaction of particles. *Phys. Fluids.* 1996, vol. 8, no. 8, pp. 1990–2010.

19. Ying L., Biros G., Zorin D. A kernel-independent adaptive fast multipole algorithm in two and three dimensions. *J. Comput. Phys.* 2004, vol. 196, no. 2, pp. 591–626.

20. Messner M., Schanz M., Darve E. Fast directional multilevel summation for oscillatory kernels based on Chebyshev interpolation. *J. Comput. Phys.* 2012, vol. 231, no. 4, pp.1175–1196.

21. Fu Y., Klimkowski K.J., Rodin G.J., Berger E., Browne J.C., Singer J.K., van de Geijn R.A., Vemaganti K.S. A fast solution method for three-dimensional many-particle problems of linear elasticity. *Int. J. Numer. Meth. Engng.* 1998, vol. 42, pp. 1215–1229.

22. Gumerov N.A., Duraiswami R. Fast multipole method for the biharmonic equation. *J. Comput. Phys*. 2006, vol. 215, no. 1, pp. 363–383.

23. Wang H., Lei T., Li J., Huang J., Yao Z. A parallel fast multipole accelerated integral equation scheme for 3D Stokes equations. *Int. J. Num. Meth. Engng.* 2007, vol. 70, pp. 812–839.

24. Tornberg A.K., Greengard L. A fast multipole method for the three-dimensional Stokes equations. *J. Comput. Phys.* 2008, vol. 227, no. 3, pp. 1613–1619.

25. Gumerov N.A., Duraiswami R. Efficient FMM accelerated vortex methods in three dimensions via the Lamb-Helmholtz decomposition. *J. Comput. Phys.* 2013, vol. 240, pp. 310–328.

26. Gumerov N.A., Duraiswami R. A scalar potential formulation and translation theory for the time-harmonic Maxwell equations. *J. Comput. Phys.* 2007, vol. 225, no. 1, pp. 206–236.

27. Gumerov N.A., O'Donovan A.E., Duraiswami R., Zotkin D.N. Computation of the headrelated transfer function via the fast multipole accelerated boundary element method and. its spherical harmonic representation. *J. Acoust. Soc. Am.* 2010, vol. 127, no. 1, pp. 370–386.

28. Nishimura N. Fast multipole accelerated boundary integral equation methods. *Appl.Mech. Rev.* 2002, vol. 55, no. 4, pp. 299–324.

29. Gumerov N.A., Duraiswami R. FMM accelerated BEM for 3D Laplace and Helmholtzequations, In: Proceedings of the International Conference on Boundary Element Techniques, BETEQ–7, Paris, France, September, 2006, pp. 79–84.

30. Greengard L., Gropp W.D. A parallel version of the fast multipole method. *Comput. Math. App.* 1990, vol. 20, no. 7, pp. 63–71.

31. Kabadshow I., Dachsel H., Hammond J. Passing the three trillion particle limit with an error-controlled fast multipole method. In: Proceedings of the 2011 Companion on High Performance Computing Networking, Storage and Analysis (SC'11), Seattle, WA, USA, November, 2011.

32. Rahimian A., Lashuk I., Veerapaneni S., Chandramowlishwaran A., Malhotra D., Moon L., Sampath R., Shringarpure A., Vetter J., Vuduc R., Zorin D., Biros G., Petascale direct numerical simulation of blood flow on 200K cores and heterogeneous architectures. In: Proceedings of the 2010 ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis (SC'10), New Orleans, LO, USA, November, 2010.

33. Gumerov N.A., Duraiswami, R. Fast multipole methods on graphics processors. Presented on the Workshop on General Purpose Computation on Graphics Processing Units in Astronomy and Astrophysics (AstroGPU 2007), IAS, Princeton, NJ, USA, November, 2007.

34. Gumerov N.A., Duraiswami R. Fast multipole methods on graphics processors. *J. Comput. Phys.* 2008, vol. 227, no. 18, pp. 8290–8313.

35. Hamada T., Narumi T., Yokota R., Yasuoka K., Nitadori K., Taiji M. 42 TFlops hierarchical N-body simulations on GPUs with applications in both astrophysics and turbulence. In: Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis (SC'09), Portland, OR, USA, November, 2009.

36. Hu Q., Gumerov N.A., Duraiswami R. Scalable fast multipole methods on distributed heterogeneous architectures. In: Proceedings of the 2011 International Conference for High Performance Computing, Networking, Storage, and Analysis (SC'11), Seattle, WA, USA, November, 2011.

37. Hu Q., Gumerov N.A., Duraiswami R. Scalable distributed fast multipole methods. In: Proceedings of the 14th, International Conference on High Performance Computing and Communications (HPCC'12), Liverpool, UK, June, 2012.

38. Yokota R., Barba L.A., Narumi T., Yasuoka K. Petascale turbulence simulation using a highly parallel fast multipole method on GPUs. *Computer Phys. Comm.* 2013, vol. 184, pp. 445–455.

39. Abramova O.A., Itkulova Yu.A., Gumerov N.A. FMM/GPU accelerated BEM simulations of emulsion flow in microchannels. In: Proceedings of ASME 2013 International Mechanical Engineering Congress and Exposition IMECE 2013, San–Diego, CA, USA, November, 2013.

40. Moiseeva E.F., Maryin D.F., Malyshev V.L., Mikhaylenko K.M., Gumerov N.A., Akhatov I.S. Molecular dynamics simulations of nano-bubbles on liquid-solid interface. Presented on ASME 2013 International Mechanical Engineering Congress and Exposition IMECE 2013, San–Diego, CA, USA, November, 2013.

41. Abramova O.A., Itkulova Yu.A., Gumerov N.A., Akhatov I.Sh. *Trekhmernoe modelirovanie dinamiki deformiruemykh kapel emulsii metodom granichnykh elementov i bystrym metodom multipolei na geterogennykh vychislitelnykh sistemakh* [Threedimensional simulation of the dynamics of deformable droplets of an emulsion using the boundary element method and the fast multipole method on heterogeneous computing systems]. *Numerical Methods and Programming.* 2013, vol. 14, pp. 438–450 (in Russian).

42. Marin D.F., Malyshev V.L., Moiseeva E.F., Gumerov N.A., Akhatov I.Sh., Mikhailenko K.I. Uskorenie molekulyarno-dinamicheskikh raschetov c pomoschyu bystrogo metoda multipolei i graficheskikh protsesorrov [Acceleration of the molecular dynamics simulations using the fast multipole method and graphics processing units]. *Numerical Methods and Programming.* 2013, vol. 14, pp. 483–495 (in Russian).