

Заключение

Рассмотрен метод перекрёстной проверки предсказаний (скользящий контроль, *cross-validation*). Мы привели обобщённую формулу для вычисления погрешности предсказаний ϵ_m и сравнили её с известной оценкой погрешности воспроизведения с помощью среднеквадратичного отклонения ϵ_0 . Мы рассмотрели возможность построения модели, обладающей предсказательной силой, путём минимизации погрешности предсказаний, и предостерегли от слепого применения формулы ϵ_m к коррелированным данным.

ЛИТЕРАТУРА

1. Воронцов К. В. Комбинаторная теория надёжности обучения по прецедентам: Дис. док. физ.-мат. наук: 05–13–17. – Вычислительный центр РАН, 2010. – 271 с.
<http://www.machinelearning.ru/wiki/images/b/b6/Voron10doct.pdf>.
2. Stone M. Cross-validated choice and assessment of statistical predictions // J. Royal Stat. Soc. – 1974. – № 36 (2). – P. 111–147.
3. Zarkovich N.A., Tan T.L., Johnson D.D. First-principles prediction of phase-segregating alloy phase diagrams and a rapid design estimate of their transition temperatures // Phys. Rev. B. – 2007. – Vol. 75. – P. 104–203.
4. Zarkovich N.A., Johnson D.D. Reliable First-Principles Alloy Thermodynamics via Truncated Cluster Expansions // Phys. Rev. Letters. – 2004. – № 92 (25). – P. 255–702.

Н.А. ЗАРКЕВИЧ

МНОГОМАСШТАБНОЕ МОДЕЛИРОВАНИЕ МАТЕРИАЛОВ С ИСПОЛЬЗОВАНИЕМ РАСПРЕДЕЛЁННЫХ ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ

Рассмотрены методы моделирования материалов на распределённых компьютерных кластерах ограниченной временной доступности. Описан обобщённый алгоритм создания крупномасштабных моделей из множества результатов расчётов малого масштаба. Предложена оптимизация загрузки вычислительных мощностей.

Число компьютеров в мире уже перевалило за миллиард. С увеличением вычислительных мощностей и совершенствованием алгоритмов растёт и число задач, которые можно численно решить. Проектирование и моделирование реальных материалов с заданными свойствами – очень сложная задача. Свойства материалов определяются их электронной структурой. Без использования приближений моделирование макроскопических материалов было бы многочастичной задачей для $>10^{23}$ частиц на моль вещества. Такую задачу не смогли бы решить даже все компьютеры мира (10^9 – 10^{10} штук) за время, равное времени жизни Вселенной (10¹⁰ лет).

Совершенствующиеся вычислительные методы позволяют приближённо решать сложные задачи на доступных ресурсах. Для моделирования материалов созданы многомасштабные методы, в которых параметры модели одного масштаба определяются из результатов расчётов другого (как правило, меньшего) масштаба [1].

Особый интерес представляет сочетание расчётов электронной структуры, энергии и свойств нано-систем с предсказанием свойств макроскопических материалов, поскольку это позволяет указать на происхождение макроскопических свойств прямо из квантово-механической электронной структуры. В качестве примеров можно привести методы, позволяющие предсказывать свойства однородных материалов и/или исследовать локальные дефекты. Так, моделирование твёрдых тел с помощью разложения по взаимодействиям [2–5] сочетает расчёты электронной структуры на квантовом уровне с классическими методами Монте-Карло или молекулярной динамики атомного масштаба [6]. Для моделирования дефектов в кристаллах используется метод QM/MM [7], сочетающий квантовую механику и молекулярную механику. Особое внимание следует уделить многомасштабности [1], позволяющей использовать результаты меньшего масштаба в качестве исходных данных для построения модели более крупного масштаба.

Многомасштабное моделирование

В основе многомасштабного моделирования лежит постулат о том, что сложную систему можно рассматривать как совокупность подсистем меньшего масштаба. Предполагается, что система описывается параметрической моделью с конечным числом изменяемых параметров. При этом модель с одними и теми же параметрами должна *достаточно хорошо* описывать как сложную систему, так и каждую из подсистем. Сочетание расчётов разного масштаба включает следующие этапы:

- 1) свойства относительно простых систем (подсистем) малого масштаба рассчитываются маломасштабным (как правило, более точным) методом;
- 2) используя известные свойства малых систем (подсистем), выбираются параметры модели, *достаточно хорошо* описывающей каждую из этих систем;
- 3) выбранные параметры используются для моделирования более сложной крупномасштабной системы.

На начальном этапе рассчитываются свойства большого количества малых систем. Эти расчёты взаимно независимы и могут быть распределены между имеющимися вычислительными ресурсами. На промежуточном этапе возникают важные вопросы: как выбирать параметры и как количественно измерить, достаточно ли хороша выбранная модель. Если качество модели можно количественно измерить (например, с помощью оценки погрешностей), то ответ на первый вопрос очевиден: чем меньше погрешность, тем лучше модель; следовательно, наилучший выбор параметров модели определяется минимизацией погрешности. Мы используем погрешность предсказаний, оцененную с помощью перекрёстной проверки [8].

Распределённые расчёты

Первый этап многомасштабного моделирования является самым ресурсозатратным. На этом этапе свойства множества M относительно простых малых подсистем рассчитываются одним и тем же маломасштабным методом. Поскольку расчёт каждой подсистемы не зависит от результатов расчётов остальных подсистем, эти M расчётов можно проводить как последовательно, так и параллельно. Иными словами, большое число M взаимно-независимых расчётов можно распределить между имеющимися вычислительными ресурсами. Только на факультете информатики ППИ имеется более 50 компьютеров, пригодных для таких расчётов, которые выключены в ночное и нерабочее время. Распределённое моделирование позволит производить научные предсказания и оптимизировать загрузку вычислительных мощностей.

Число M простых подсистем можно поэтапно увеличивать для улучшения точности выстраиваемой модели. Как правило, рассчитывают все подсистемы, которые меньше заранее выбранного предельного размера R ; число $M(R)$ таких подсистем конечно и заранее известно. Увеличение предельного размера приводит к скачкообразному увеличению числа M : для $r < R$, $M(r) \leq M(R)$. Таким образом, используя имеющиеся компьютеры, можно быстро делать предсказания с малой точностью и затем постепенно её увеличивать по мере поступления новых расчётных данных. Имеющиеся вычислительные ресурсы вполне достаточны для успешного моделирования материалов и предсказания их свойств, а также для проектирования и дизайна новых материалов и нано-систем.

Усовершенствование численных методов и создание новых программ

На разных пространственно-временных масштабах используются различные методы. Начиная с малого, мы уже упоминали квантовые расчёты электронной структуры, изучение дефектов с помощью сочетания квантовой механики и молекулярной механики, моделирование кристаллических твёрдых тел с использованием разложения по взаимодействиям, методы Монте–Карло и молекулярной динамики атомного масштаба. Мы уделяли внимание многомасштабности и использованию результатов меньшего масштаба в качестве входных параметров для построения новой модели, используемой для крупномасштабных расчётов. В частности, для сочетания квантово-механических и термодинамических вычислений разработан алгоритм [3]; для сохранения результатов маломасштабных расчётов с целью их многократного использования создана база данных [1].

Вышеперечисленные непрерывно совершенствующиеся методы (и соответствующие программы с открытым кодом) применяются на квантово-механическом и атомном

нано-масштабах. Для моделирования на мезо, микро-, и макромасштабах применяются другие, более крупномасштабные методы. Для изучения сплошных сред с распределёнными полями часто используется метод конечных элементов. Разрабатываемой новой уникальной программе, позволяющей методом конечных элементов решать связанные задачи со многими взаимодействующими распределёнными полями и разнообразной физикой, посвящена недавняя статья [9].

Заключение

В заключение, имеющиеся распределённые вычислительные ресурсы ППИ технически позволяют осуществлять многомасштабное моделирование материалов. Такие расчёты помогают проектировать новые материалы (в том числе нано-материалы) с заданными свойствами и создавать инновационные нано-технологии. В дополнение к доступным программам нами разработаны алгоритмы и созданы программные продукты с элементами искусственного интеллекта для распределённых и взаимосвязанных расчётов нано- и макромасштаба.

ЛИТЕРАТУРА

1. Zarkevich N.A. Structural Database for reducing cost in materials design and complexity of multi-scale computations. – Complexity. – Wiley, 2006. – Vol. 11 (4). – P. 36–42.
2. Sanchez J.M., Ducastelle F., Gratias D. Generalized cluster description of multicomponent systems. – Physica, 1984. – Vol. 128A. – P. 334.
3. Zarkevich N.A., Johnson D.D. Reliable Alloy Thermodynamics from Truncated Cluster Expansions. – Phys. Rev. Lett. – 2004. – Vol. 92. – P. 255–702.
4. Ficsher C.C., Tibbetts K.J., Morgan D., Ceder G. Predicting crystal structure by merging data mining with quantum mechanics. – Nature Materials, 2006. – Vol 5. – P. 641–646.
5. Zarkevich N.A., Tan T.L., Wang L.L., Johnson D.D. Low- energy antiphase boundaries, degenerate superstructures, and phase stability in frustrated fcc Ising model and Ag-Au alloys. – Phys. Rev. B., 2008. – Vol. 77. – P. 144–208.
6. Multiscale Modeling: From Atoms to Devices // P. Derosa, T. Cagin, Editors. – CRC press, 2010. – 271 p.
7. Multiscale Simulation Methods in Molecular Sciences // J. Grotendorst, N. Attig, S. Blugel, D. Marx, Editors. – NIC Series, 2009. – Vol. 42. – 576 p.
8. Заркевич Н.А. Количественная оценка погрешности предсказаний. – Труды ППИ. – Псков, 2011.
9. Исаков А.Н., Андреев М.Л., Козырева О.И., Плохов И.В., Заркевич Н.А. Метод конечных элементов. Мультифизика. – Труды XII областной научно-практической конференции молодых учёных «Энергия и талант молодёжи – залог развития инновационных и наукоёмких производств». – Т. 1. – Псков, 2011.

Т.Н. МИХАЙЛУСОВА, А.Е. ЛУКИН

ЭКЗАМЕНАЦИОННОЕ ИНТЕРНЕТ-ТЕСТИРОВАНИЕ СТУДЕНТОВ ПО ФИЗИКЕ

Проведён анализ тематической структуры интернет-тестов по физике и методики оценки результатов. Рассматриваются некоторые аспекты целенаправленной подготовки студентов к экзаменационному тестированию.

Согласно решению Учёного Совета ППИ «О плане подготовки аккредитации института 2011 года» сотрудниками кафедры физики была проведена работа по подготовке студентов вторых курсов к интернет-тестированию по дисциплине «Физика». Аккредитация – процедура, периодически повторяющаяся и, на наш взгляд, некоторый анализ и обобщение итогов тестирования может быть полезным для использования в будущем.

Большинство учебных программ, в том числе и по физике, предусматривает изложение материала от простого к сложному. Замечено, что этот же принцип заложен в качестве основного при формировании тестов интернет-экзамена. Все тесты разбиты на шесть (или в некоторых случаях на семь) разделов, получивших название «Дидактическая единица» (ДЕ). Каждый такой раздел охватывает фундаментальные положения физики, с которыми студент обязан быть знаком. Ниже, в качестве примера, приводится таблица, содержащая семь ДЕ для специальности **190601.65 – «Автомобили и автомобильное хозяйство»**.