

メニーコア型スーパーコンピュータにおける

100 ナノメートル電子状態計算の強スケーリング性

星健夫・角田皓亮¹

鳥取大学

1. はじめに

スーパーコンピュータを利用した大規模電子状態計算は、次世代材料設計分野の基盤をなす。そのためには、アーキテクチャとアルゴリズムの両者を見据えての手法開発（コデザイン）が重要となる。応用分野として筆者らは、有機デバイス材料系に注力している。有機デバイスは、フレキシブル（薄くて柔らかく曲げたり巻いたりできる）光電子デバイスとして、トランジスタ・太陽電池・ディスプレイ・生体センサーなどが、次世代 Internet-of-Things (IoT) の中核をなす [1-3]。白川英樹 (2000 年ノーベル化学賞) らの研究など、伝統的に日本が強い。フレキシブルデバイスは、有機（高）分子凝集体であり、構造乱れ (disorder) が本質である。こうした系を計算するには、100 ナノメートルスケール電子状態計算が必須となる。上記を実現するためには、従来型（第一原理）計算を超えた、大規模計算手法を開発する必要がある。筆者らは数理系研究者と共同で、大規模電子状態計算法の基盤となる独自大規模疎行列計算アルゴリズムを複数開発してきた [4-16]。基本的戦略は、目的物理量ごとに、計算アルゴリズムを設計することである。分子とデバイスのスケールについて、図 1 に模式図の例を示した。

本記事では、シフト型クリロフ部分空間理論を用いた超並列型電子状態計算について、Oakforest-PACS 全ノード計算 (大規模 HPC チャレンジ) の結果を中心に報告する。また、関連した数理手法についても、合わせて報告する。

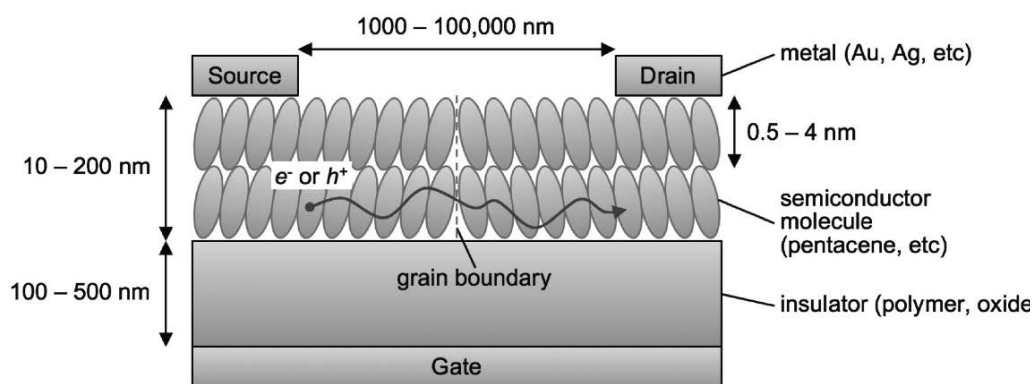


図 1 有機デバイスの例（トランジスタ、模式図） [12].

2. 大規模電子状態計算の数理的基礎とこれまでの経緯

電子状態計算の基礎方程式は、シュレーディンガー方程式に由来する一般化固有値問題

$$A \mathbf{y}_k = \lambda_k B \mathbf{y}_k \quad (1)$$

¹ 現所属：昭和電工株式会社

が典型である．ここで行列 A, B は $N \times N$ エルミート行列(以下計算では実対称行列)であり， B は正定値である．行列 A, B はハミルトン行列，重なり行列，とそれぞれ呼ばれる．固有値(λ_k)は電子 1 つ 1 つのエネルギーであり，固有ベクトル(\mathbf{y}_k)は電子波動関数の情報である．式(1)に密行列ソルバーを用いると，演算コストは N^3 に比例し， $N = 100$ 万次元程度(以下の計算における，25 万原子系程度に相当)が限界となる[9]．原子基底などの局在基底を用いた場合は，行列 A, B は疎行列として扱え，疎行列(クリロフ部分空間)ソルバーが利用できる．具体的計算手法の 1 つは，一般化固有値問題の代わりに，一般化シフト型線形方程式

$$(zB - A) \mathbf{x} = \mathbf{b} \quad (2)$$

を基礎方程式とすることである．式(2)に，疎行列ソルバーの一種であるシフト型クリロフ部分空間ソルバーを用いることで，大規模計算向けソルバーが構築できる[4-7, 9]．ここで z は(複素数としての)エネルギーであり， \mathbf{b} は与えられたベクトルである．グリーン関数 $G \equiv (zB - A)^{-1}$ を用いると，解ベクトル \mathbf{x} は形式的に $\mathbf{x} = G\mathbf{b}$ と書ける．そのため，式(1)に基づく計算法を波動関数描像，式(2)に基づく計算法をグリーン関数描像，と考えることができる．グリーン関数と波動関数には，

$$G(z) = \sum_k^N \frac{\mathbf{y}_k \mathbf{y}_k^T}{z - \lambda_k} \quad (3)$$

の関係がある．式(2)に対するシフト型クリロフ部分空間ソルバーとして，以下では，超並列型の多重(multiple)アーノルディ法[6, 9]を用いている．具体的には，独自計算コード ELSSES (Extra-Large-Scale Electronic Structure calculation)[17]に実装にされている．ELSES では，第一原理に基づくモデル化(強束縛型)電子状態理論が用いられている．プログラムは，MPI/OpenMP のハイブリッド並列を行っている．数理的概要は，目的物理量がある行列 X のトレース

$$\text{Tr}[X] = \sum_i^N X_{ii} \quad (4)$$

で書ける場合に，各ノードが式(4)右辺における部分和を計算することにより，並列計算が達成できる(アルゴリズム詳細は[6, 9])．したがって原理的な最大並列度数は行列次元数 N となる．以下の計算では，行列次元数 N は原子数 N_a の数倍である．現状ではプログラムコードの仕様で，最大 MPI 並列度数は原子数 N_a となっている．上記より，MPI 並列度数 P_{MPI} より原子数 N_a が十分大きい場合($N_a \gg P_{MPI}$)，高い強スケーリング性が期待できる．ただし，通信時間など，他要員もあり，実際に強スケーリング性が達成されるかどうかは，実装にも依存する．これまで，京全体を使った計算として，最大で，1 億原子系の電子状態計算が達成されている[9]．構造乱れのある有機高分子 poly-(phenylene-ethynylene) 集合体に対する約 200 ナノメートル立方領域に相当する系であり，100 ナノメートルスケール計算が達成されたことになる．これら成果は，HPCI を通じた住友化学との産学共同研究に発展した(最近の課題：HPCI 産業利用実証課題 hp200049「大規模量子化学計算によるデータ生成と機械学習を統合した有機半導体材料設計」，利用計算機：Oakforest-PACS)．

なお，ELSES で生成された行列 A, B の一部データは，数値計算(疎行列ソルバー)研究者を想定して，ELSES matrix library として公開している[12, 18]．

3. メニーコア CPU 向けのプログラム改良

従来は京でプログラム開発を主に行っていたが，2018 年度より Oakforest-PACS に移行した．

アルゴリズムに変更はないが、実装において、メニーコア CPU 向け改良を行なった。京は 1CPU が 8 コアであったのに対し、Oakforest-PACS は 68 コア (計算利用は 64 コア) からなるメニーコア CPU 機である。Oakforest-PACS は flat モードで動作しているノードと、cache モードで動作しているノードがあるが、主には flat モードで利用している。メニーコア型 CPU では、CPU 内でコアが複数ドメインに分かれるなどの事情があり、キャッシュ有効利用が (京以上に) 重要と予想された。これに対応するために、OpenMP thread 内部で working array を用意するなどして、(メモリ量の増加になるが) キャッシュ有効利用化を促進させた。

ノード内 MPI 並列度数 Q を変えてのベンチマークを、表 1 にまとめた。原子座標が与えられてから電子状態計算が完了するまでの実行時間 T を記した。計算系は、ダイヤモンド系であり、周期セル内に $N_a = 524,288$ 原子系 (約 52 万原子系) を含むスーパーセル系である。行列次元は $N = 4N_a = 2,097,152$ (約 200 万) である。 $P = 4$ ノードを用い、ノード内 MPI 並列度数 Q を、 $Q = 1, 4, 16, 64$ と変化させ、経過時間を比較した。それぞれの計算で、MPI 並列度は $P_{MPI} = PQ$ に、OpenMP スレッド数は $R = 64/Q$ に設定した。 $Q = 64$ はフラット MPI に相当する。結果として実行時間 T は、ノード内 MPI 並列度数 Q にほとんどよらなかった。このことから、キャッシュ効率化が達成されていると考えられる。

表 1 実行時間 T のノード内 MPI 並列数 Q 依存性。ダイヤモンド約 52 万原子系。
 $P = 4$ ノード計算。OpenMP スレッド数 R は、 $R = 64/Q$ に設定した。

ノード内 MPI 並列数 Q	1	4	16	64 (フラット MPI)
実行時間 T (s)	371.13	371.10	370.25	371.06

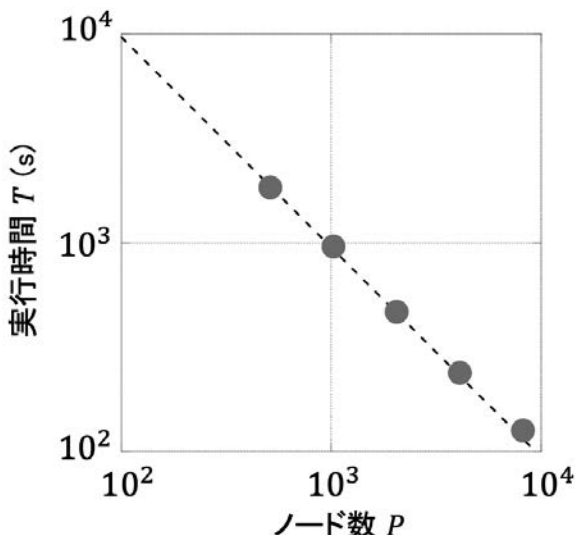


図 2 $P=512, 1024, 2048, 4096, 8192$ ノードでの強スケーリングテスト。ダイヤモンド約 3 億原子系に対して、実行時間 T (s) をノード数 P の関数として描いた。理想スケーリング線 (破線) も描いた。

4. Oakforest-PACS でのベンチマーク

大規模 HPC チャレンジ制度を利用して、Oakforest-PACS 全ノード (8192 ノード) 計算を行った。これまでの計算実績上限である 1 億原子系 [9] の 3 倍に当たる約 3 億原子系計算を行った。

図 2 に、ダイヤモンド約 3 億 ($N_a = 314,136,576$) 原子系での $P=512, 1024, 2048, 4096, 8192$ ノードでの強スケーリングベンチマークテスト結果を示した。1 ノードは 1MPI プロセス・64 OpenMP スレッドに設定した。原子座標が与えられてから電子状態計算が完了するまでの実行時間 $T = T(P)$ を測定した。強スケーリング型並列効率 $\alpha(P)$

$$\alpha(P) \equiv \frac{T(512)/T(P)}{P/512} \quad (5)$$

は、 $\alpha(1024) = 0.958$, $\alpha(2048) = 0.987$, $\alpha(4096) = 0.966$, $\alpha(8192) = 0.910$, であり、非常に高い強スケーリング性が得られた。

5. 関連した数理手法

最後に、我々が開発してきた関連手法について概要を報告する。これらは電子状態計算以外にも有用である；(i) シフト型クリロフ部分空間手法は複数あるが、その一部(シフト型 conjugate-gradient 法, シフト型 conjugate-orthogonal conjugate-gradient 法, シフト型 bi-conjugate gradient 法)について、汎用数値計算ライブラリ $K\omega$ を開発・公開した [14, 19]。多体量子模型ソルバー $H\Phi$ [20, 21] での動的グリーン関数の計算機能などに利用されている。(ii) 式 (2) に基づくグリーン関数計算法を補完する手法として、選択的中間固有対計算手法を開発し、数値計算ライブラリ $k\text{-ep}$ として公開した [8, 11, 22]。グリーン関数計算法では、固有対 (固有値・固有ベクトル) を計算することができない。しかし、デバイス性能 (電気伝導・発光特性など) には、一部の (最高占有エネルギー・最低非占有エネルギー付近の) 固有対が必要であり、これらを選択的に計算する手法が当該手法である。(iii) 実測前に強スケーリング性能を予測する手法を開発した [13]。ベイズ推定と性能モデルに基づき、実行時間 T をノード数 P の関数 ($T = T(P)$) として外挿する ($P \rightarrow$ 大) 手法である。本研究では Oakforest-PACS 全ノードまでの強スケーリング性ベンチマークを行ったが、このような大規模計算資源が利用できる機会は限られている。そのため、少ないノード数でのベンチマークから、計算時間の外挿できる手法への需要が大きい。

6. まとめと展望

独自数理アルゴリズムを用いた大規模電子状態計算により、Oakforest-PACS 全ノードに至る強スケーリング性能が確認された。今後は、他の関連手法と合わせて、有機デバイス材料系の産学連携研究を推進していきたい。また、開発された数理手法は汎用であり、電子状態計算以外の計算科学分野への波及も期待できる。

謝 辞

大規模 HPC チャレンジ (Oakforest-PACS 全ノード計算) 実施に際し、東京大学情報基盤センター・富士通株式会社の関係者各位に、情報提供・サポートを受けた。

参 考 文 献

- [1] T. Sekitani, T. Someya, Human-friendly organic integrated circuits. *Mater. Today* 14, 398-407 (2011).
- [2] 安達千波矢, 有機半導体のデバイス物性, 講談社 (2012).
- [3] H. Matsui, Y. Takeda, S. Tokito, Review: Flexible and printed organic transistors: From materials to integrated circuits, *Organic Electronics* 75 105432/1-17 (2019).
- [4] H. Teng, T. Fujiwara, T. Hoshi, T. Sogabe, S.-L. Zhang, S. Yamamoto, Efficient and accurate linear algebraic methods for large-scale electronic structure calculations with nonorthogonal atomic orbitals, *Phys. Rev. B* 83, 165103/1-12 (2011).
- [5] T. Sogabe, T. Hoshi, S.-L. Zhang, T. Fujiwara, Solution of generalized shifted linear systems with complex symmetric matrices, *J. Comp. Phys.* 231, 5669-5684 (2012).
- [6] T. Hoshi, S. Yamamoto, T. Fujiwara, T. Sogabe, S.-L. Zhang, An order- N electronic structure theory with generalized eigenvalue equations and its application to a ten-million-atom system, *J. Phys.: Condens. Matter* 24, 165502/1-5 (2012).
- [7] T. Hoshi, Y. Akiyama, T. Tanaka, T. Ohno, Ten-million-atom electronic structure calculations on the K Computer with a massively parallel order- N theory, *J. Phys. Soc. Jpn.* 82, 023710/1-4 (2013).
- [8] D. Lee, T. Miyata, T. Sogabe, T. Hoshi, S.-L. Zhang, An interior eigenvalue problem from electronic structure calculations, *Jpn. J. Ind. Appl. Math.* 30, 625-633 (2013).
- [9] T. Hoshi, H. Imachi, K. Kumahata, M. Terai, K. Miyamoto, K. Minami, F. Shoji, Extremely scalable algorithm for 10^8 -atom quantum material simulation on the full system of the K computer, *Proc. ScalA16 in SC16*, 33-40 (2016).
- [10] H. Imachi, T. Hoshi, Hybrid numerical solvers for massively parallel eigenvalue computations and their benchmark with electronic structure calculations, *J. Inf. Process* 24, 164-172 (2016).
- [11] D. Lee, T. Hoshi, T. Sogabe, Y. Miyatake, S.-L. Zhang, Solution of the k -th eigenvalue problem in large-scale electronic structure calculations, *J. Comp. Phys.* 371, 618-632 (2018).
- [12] T. Hoshi, H. Imachi, A. Kuwata, K. Kakuda, T. Fujita, H. Matsui, Numerical aspect of large-scale electronic state calculation for flexible device material *Jpn. J. Indust. Appl. Math* 36, 685-698 (2019).
- [13] K. Tanaka, H. Imachi, T. Fukumoto, A. Kuwata, Y. Harada, T. Fukaya, Y. Yamamoto, T. Hoshi, EigenKernel: A middleware for parallel generalized eigenvalue solvers to attain high scalability and usability, *Jpn. J. Ind. Appl. Math.* 36, 719-742 (2019).
- [14] T. Hoshi, M. Kawamura, K. Yoshimi, Y. Motoyama, T. Misawa, Y. Yamaji, S. Todo, N. Kawashima, T. Sogabe, $K\omega$ - Open-source library for the shifted Krylov subspace method of the form $(zI - H)\mathbf{x} = \mathbf{b}$, *Comp. Phys. Comm.* 258, 107536/1-10 (2021).
- [15] T. Hoshi, Y. Yamamoto, T. Sogabe, K. Shimamura, F. Shimojo, A. Nakano, R. Kalia, P. Vashishta, Numerical methods for large scale electronic state calculation on supercomputer, Chap. 15 of *21st Century Nanoscience - A Handbook: Nanophysics Sourcebook (Volume One)*, ed. K Sattler, CRC Press (2019).
- [16] T. Hoshi, S. Itoh, Recent progress in large-scale electronic state calculations and data-driven sciences, *Chap. 14 of Handbook of Silicon Based MEMS Materials and Technologies 3rd. Ed.*, ed. M. Tilli, M. Paulasto-Kröckel, M. Petzold, H. Theuss, T. Motooka, V. Lindroos, Elsevier (2020).
- [17] <http://www.elses.jp/>
- [18] <http://www.elses.jp/matrix/>
- [19] <https://github.com/issp-center-dev/Komega/>
- [20] M. Kawamura, K. Yoshimi, T. Misawa, Y. Yamaji, S. Todo, N. Kawashima, *Comp. Phys. Commun.* 217, 180-192 (2017).
- [21] <http://www.pasums.issp.u-tokyo.ac.jp/hphi/>
- [22] <https://github.com/lee-djl/k-ep/>