課題番号　jh130011-NA07

# 超大規模超並列電子状態計算を中核とした物理・数理・HPC の融合研究

星健夫（鳥取大学, JST-CREST(PostPeta)）

概要
並行している「京」利用プロジェクトなどと連携し、独自開発の超大規模超並列電子状態計算コード(ELSES)を当座の対象として、物理分野（ニーズ）と数理・HPC 分野（シーズ）の融合研究を行った。主な成果は下記である；(1)超大規模超並列計算：「京」全コアでの 1 億原子(世界最大、１００ナノスケール)電子状態計算が初めて実現され、高い並列効率(strong scaling)を得た。(2) 中規模計算：ナノ構造研究者での需要が高い中規模計算(FX10 での 1000 コア計算を想定)を対象に、「準最速」計算手法（最速性を多少犠牲にしても、簡単に使えることを優先させる手法）を開発した。(3)新数理ソルバー：電子物性に重要な内部固有値問題に対して、シルベスタ慣性定理にもとづく新しい数理ソルバーを提案した。(5)ナノ物質応用研究：手法の特徴を生かし、ナノ複合カーボン、アモルファス状共役高分子、電池関連物質(イオン液体・リチウム系)、の研究研究を行った。今後の展望として、量子電気伝導などへの理論拡張を目指しつつ、100 ナノスケール系の各種デバイス研究へとつなげていきたい。

## 1.　研究の目的と意義

電子１つ１つを「波」としてあつかう量子物質計算（電子状態計算）は、量子力学にもとづく盤石な微視的理論基盤をもつ。ナノ物質系での応用は、材料・物理・化学・電気・機械など、従来学問分野を超えて広く普及している。理工学・産業にあらわれる種々の問題を解決するには、電子状態計算を基盤とするボトムアップ型アプローチとしての大規模系計算が必須となる。

近年筆者らは、計算科学と計算機科学の融合として、超大規模超並列電子状態計算コード ELSES（注 1）の開発と応用に取り組んでいる。

本プロジェクトの概要を図 1 に示した。本プロジェクトは、産業応用まで含めた Application 分野と Algorithm 分野と Architecture 分野の垂直統合(co-design)プロジェクトである。アイディアの根源は、「行列生成＋大行列数理ソルバー＋物理量計算」（の繰り返し）を一体と考え、アルゴリズムを最適デザインする点にある。個々の技術は汎用なものであり、他分野実アプリへの「水平的」波及も期待できる。

特に今年度は、ナノ物質応用研究としては世界最大となる、1 億原子(100 ナノスケール)電子状態計算が達成されている(図 2)[5] (注２)。こうした計算は、「オンリーワン」としての意義をもっている。

計算手法は、超並列オーダー$N$（$0(N)$）型計算である(注３)(図 3(a))。中核となるのは数種類の革新的大行列数理(クリロフ部分空間)ソルバー(注 4)であり、金属・半導体など幅広い物質に適用できる。これらは、第一原理にもとづくモデル化(tight-binding 型, TB 型)理論を用いている。

応用研究としては、sp2-sp3 ナノ複合カーボン固体・アモルファス状共役高分子・電池関連物質(イオン液体, リチウム系固体)、に注力している。

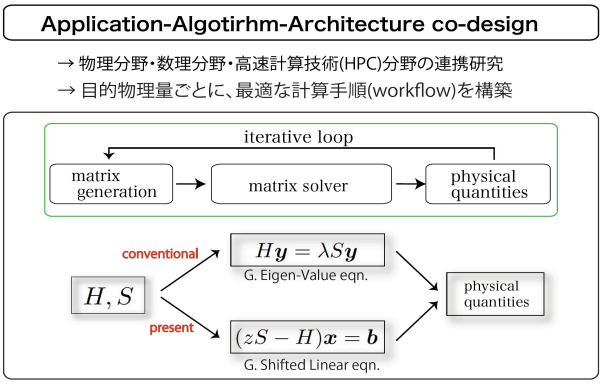



図 1:本プロジェクトの概要。

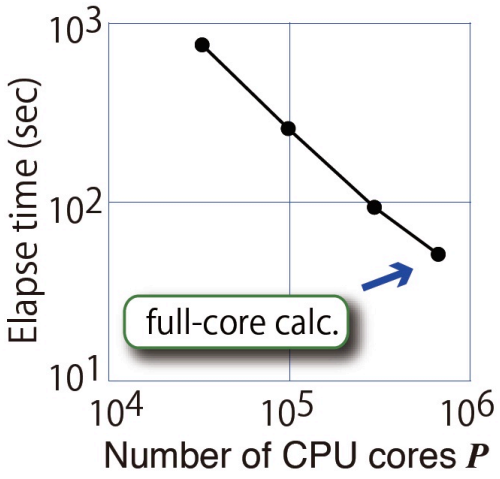
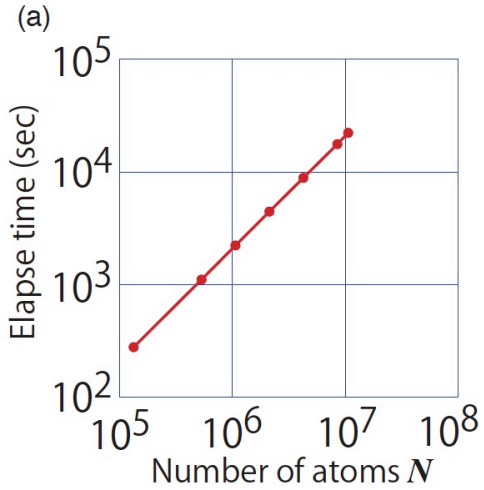



図 2:「京」での 103,219,200(約 1 億,世界最大)電子状態計算における並列効率(strong scaling) [5]。全コア(663,552 コア)までの計算。利用コア数 P=32,768 を基準とした並列効率 α が、P=98,304 で α=0.98、P=294,912 で α=0.90、P=663,552(「京」全コア)で α=0.73、であった。物質は、ナノ多結晶ダイヤモンド研究における、sp2-sp3 ナノ複合カーボン固体。



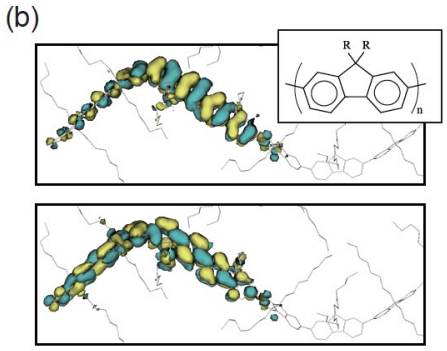




図 3(a)オーダー$N$(O($N$))スケーリング(注 2)。物質はアモルファス状共役高分子(aCP)(poly-(9,9 dioctyl fluorene)(注 4)。(b)aCP における波動関数の例[5]。最高占有(HOMO)波動関数(上段)、最低非占有(LUMO)波動関数(下段)、分子式(右上)。

本プロジェクトでは、以下をテーマとして取り組んだ。当初計画案での項目番号(i)~(iv)との対応を付して、項目をあげる。各項目詳細は「3. 研究成果の詳細」で述べる。

(A) 超大規模超並列計算(「京」も含めて)(注 5)
(B) 中規模計算での「準最速」計算(FX10 での 1000 コア計算を中心に)
(C) 独自開発疎行列ソルバーアルゴリズムの比較・実装・最適化(当初計画案での(i))
(D) FX10「ランクディレクトリ」機能と分割 XML ファイルを基盤とした、最適な並列化ファイル I/O 機構(当初計画案での(iii))
(E) 新数理ソルバー研究(当初計画案での(iv))
(F) ナノ物質応用研究(当初計画案での(ii))
(G) その他 (行列データベース公開など)

テーマ(A)(B)は、テーマ(C)が予想以上によかったために達成された発展的テーマである。テーマ(G)は、プロジェクト進展とともに発生した発展的テーマである。中間報告(2013 年 12 月)以降に発展が大きかったのは、テーマ(E)[14][15]である。

(注 1) ELSES ( = Extra Large-Scale Electronic Structure calculation) : http://www.elses.jp/

(注 2) シリコン単結晶では、5,000 万原子系が 100 ナノメートル立方領域に相当するため、本研究では 5,000 万以上の原子系を「100 ナノスケール」と呼ぶ。

(注 3) オーダー$N$ (O($N$))法とは、計算時間が原子数 $N$ に比例する手法の総称である。同様に、O($N^k$)法($k$ は自然数)は、計算時間が原子数 $N$ の $k$ 乗に比例する手法をさす。通常の電子状態計算は、O($N^3$)あるいはより高次の計算量がかかる。そのため原子数を 10 倍にすると計算時間が 1000 倍以上に増加し、計算機が高速になっても大規模系への発展が難しい。一方、オーダーN 法では計算量は 10 倍

にしか増えない(図３(a)参照)。

(注4) 図2,図3(a)では、一般化シフト型線形方程式((*zS*-*H*)*x*=*b*,図 1 参照)に対する多重アーノルディ(multiple Arnoldi)法を用いている；Takeo Hoshi, Susumu Yamamoto, Takeo Fujiwara, Tomohiro Sogabe, Shao-Liang 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 (2012).

(注 5) 「京」での計算は、提案番号 hp120170, hp120280, hp130052を通じて行われた。

## 2. 当拠点公募型共同研究として実施した意義

### (1) 共同研究を実施した拠点名および役割分担

拠点名：東京大学(計算機：Oakleaf-FX)

役割分担：

(a) 大規模電子状態計算コードELSESの超並列化・データベース構築（テーマ(A)-(D),(G)）・・・・星健夫（鳥取大、代表)、井町宏人(東大)、秋山洋平・山﨑渓太（鳥取大）

(b) 次世代大規模電子状態にむけた新しいアルゴリズム研究および関連する数理的研究（テーマ(E)）・・・曽我部知広(愛知県立大,副代表)、 宮田考史・張紹良(名古屋大)、星健夫(鳥取大, 代表)

(c) ナノ物質応用研究（テーマ(F)）・・・西野信也（東大)、山元進（東京工科大）星健夫(鳥取大, 代表)

### (2) 共同研究分野

超大規模数値計算系応用分野

### (3) 当公募型共同研究ならではという事項など

(a) 応用(電子状態計算)分野研究者と数理・HPC 分野研究者の融合チームであり、基盤的数理研究から超並列化、実用研究まで、全てを垂直統合したプログジェクトが実施できた。

(b) 実施マシン(富士通FX10)は「京」との親和性が高く、「京」向けのソフト開発を行う目的として、理想的であった。例えば、ランクディレクトリ機能を用いた入出力ファイル(XML形式)の並列化(テーマ(D))は本研究の大規模計算にとって必須であったが、これは FX10上でのソフト開発の貢献が非常に大きい。

## 3. 研究成果の詳細と当初計画の達成状況

### (1) 研究成果の詳細について

「1. 研究の目的と意義」にあげたテーマ(A)-(G)を順番に説明する。

(A) 超大規模超並列計算

テーマ(C)(アルゴリズムの比較・実装・最適化)をFX10で行い、その成果と知見を基礎に、「京」全コア計算での1億原子(世界最大,100ナノスケール)電子状態計算を行い、高い並列効率(strong scaling)を得た[5](図２)。昨年度(注6)の10倍の原子数であり、テーマ(C)を基礎としたコード発展による大きな成果と言える。

(B) 中規模計算での「準最速」計算(FX10での1000コア計算を中心に)

ナノ物質研究対象は多岐にわたるため、テーマ(A)のような世界最大規模の計算だけでなく、中規模計算(1000コア程度)が重要であり、これをターゲットとして「準最速」計算手法を開発した。

「準最速」計算とは、最速性を多少犠牲にしても、簡単に使えることを優先させる手法である。電子状態計算のユーザーには、手法や計算機の非専門家（例えば、実験系研究者や企業系研究者）も想定される。そのため、使いやすい「準最速」計算への需要が極めて高い。

結果を図 4 に示した。厳密な最速計算である「micro-cell」法と、「準最速」(一部を計算である「one-cell」法を比較した。両者の結果は数学的に等価であり、計算時間は変わっても計算結果は全く変わらない。「micro-cell」法は厳密な O(*N*) 計算であり最速であるが、設定項目が多く、非専門家には利用が難しい。一方、「one-cell」法では一部の計算を O($N^2$) で行うため計算コスト・メモリコストが増えるが、設定項目が少なく、非専門家でも容易に扱える。

図 4 に示した計算(1000 コア程度・1 万原子程度)では、両者の違いは実用上ないに等しく、「準最速」計算法が有用であることが分かった。特に、スパコンユーザー人口を増やすという意味で、「準最速」計算法は必須である。

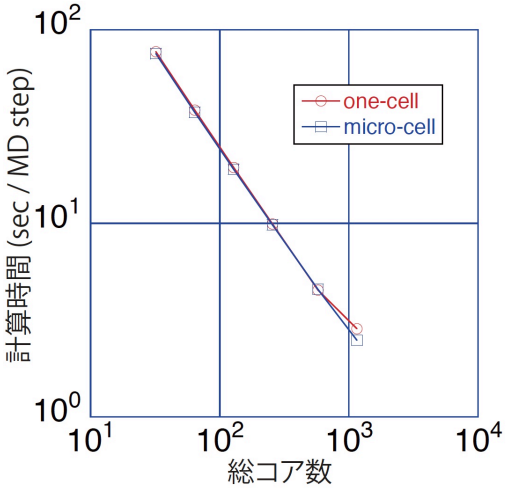



図 4:中規模計算における「準最速」計算手法の研究。FX10 における並列効率(最大 1152 コアまで)。物質は液体カーボン 13824 原子系。厳密な最速計算(厳密な O(*N*)計算)である「micro-cell」法と、「準最速」(一部を(O($N^2$)にした)計算である「one-cell」法の比較。

(C) 独自開発疎行列ソルバーアルゴリズムの比較・実装・最適化

昨年度までに開発した、疎行列ソルバーアルゴリズム（一般化シフト型線形方程式にたいするクリロフ部分空間法）を比較・実装・最適化した。これら成果がテーマ(A)(B)へと発展していった。具体的には、(a)演算コスト・(b)通信コスト・(c)メモリコスト・(d)ファイル IO コスト・(e)ロードバランス・(f)人的コスト（「手間」)、の競合が論点とった。

本年度は量子分子動力学を行うことを想定し、多重アーノルディ法(注 4)が上記にあげた論点から総合的に見て良いアルゴリズムであり、これにしぼる結論となった。

コード発展は、計算コストとメモリコストのバランスの調整を中心に行った。計算内容は数学的に等価であるため、計算結果を全く変えず、計算コスト・メモリコストだけを変える。特にメインメモリ容量の限界を回避するために、同じ量を封数回（冗長に）計算することで、メモリコストの節約を図ったワークフローを構築している(「省メモリ型ワークフロー」)（注 7)。こうしたワークフローはキャッシュメモリの効率的活用という点からも有用であり、次世代のスパコンにおいても重要となろう。

(D) FX10「ランクディレクトリ」機能と分割 XML ファイルを基盤とした、最適な並列化ファイル I/O 機構

分割 XML ファイルを基盤とした、最適な並列化ファイル IO 取り組んだ。XML ファイルにおける記述の例として、原子 1 個の座標データは下記のようなものである；

```
<atom element="C">
  <position unit="angstrom"> 1.0d0 0.0d0 0.0d0 </position>
</atom>
```

XML パーサーとしては、代表的なオブジェクト型(DOM)・ストリーム型(SAX)、の実装と最適化を行った。特に、10 万原子以上では、原子構造データが大きすぎるため SAX 型でないと実用的でなくなるため、原子構造データについては SAX のみとした。DOM, SAX 以外のパーサーも検討したが、差し迫った課題でないと判断し、実装はしなかった。

分割 XML ファイルを用いて、FX10(および「京」)の「ランクディレクトリ」機構を基盤とした最適な並列化ファイル I/O 機構を実現した。共有すべきデータ(例えば総原子数・分割情報)をファイルサイズの小さいファイルにいれこれを共通に読み込むファイルとすることで、利便性と高速性の両立がはかれた。

並列効率は「ノード数＝分割ファイル数」のときが最高であるが、上記以外の場合でも稼動するような実装とした。

(E) 新数理ソルバー研究

電子物性研究(光学特性や量子電気伝導)を想定して、内部固有値問題を対象とした、新しい数理ソルバー研究を行った(図 5)。

内部固有値問題とは、一部の（最低でも最高でもない）固有値や対応する固有ベクトルを計算する問題である。大行列計算では全固有値・固有ベクトルを計算することが（計算コストの点から）困難になり、このような計算法が重要になってくる。特に電子物性研究(光学特性や電気伝導)には、一部の（フェルミレベルと呼ばれる固有値やその近傍の）固有値・固有ベクトルだけの高精度計算が必要になる場合が多いため、大きな需要がある。

本研究では、$k$ 番目の固有値・固有ベクトル($k$ は与えられた自然数)を計算する、という「順番指定型」問題を対象に、手法研究を行った。

結果を図 5 に示した[7]。シルベスタ慣性定理を使うことで、繰り返し型計算により $k$ 番目の固有値を求めるアルゴリズムである。図では 3 つの手法を比較しながら、収束までの繰り返し回数を比較している[7]。固有ベクトル計算についての手法開発も行っている[10]。

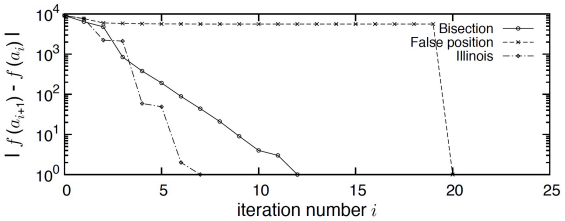



図 5: 内部固有値問題に対する、新しい数理ソルバー計算[7]。目的解に辿り着くまでの収束の振る舞いを、３種類の手法で比較。

また、大規模電子状態計算に関連する数理的研究もあわせて行った; (a) 大規模シフト非対称線形方程式に対す Restarted shifted GMRES 法の改良 [2]。(b) k-3 重対角 Toeplitz 行列の逆行列について、任意成分を容易に算出できる方法を提案[3]。(c) 電子状態計算(表面グリーン関数計算)に現れる特殊な固有値問題に対しする、高精度なアルゴリズム開発[4]。

(F) ナノ物質応用研究

手法の特徴を生かした、ナノ物質応用研究として、以下の３課題を行った。

(i)超強度材料であるナノ多結晶ダイヤモンドを対象とした sp2-sp3 ナノ複合カーボン固体のドメイン境界解析,(ii)オプトエレクトロニクスの基礎材料であるアモルファス状共役高分子系(poly-(9,9 dioctyl fluorine))に出現する局在した $\pi$ 電子波動関数[15]、(iii)電池関連物質研究(イオン液体分子[1](図 6)およびリチウム固体系[14]) (注 8)。

これらは年度後半からスタートしたものが多く、中間報告書と本最終報告書の主な差異である。

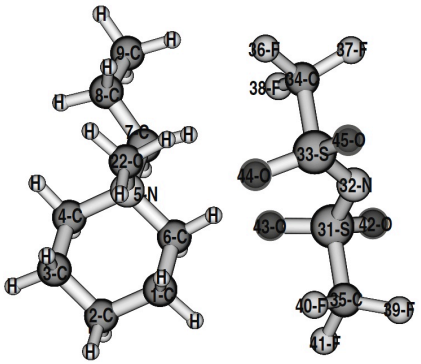


図 6: 論文[1]で取り上げた、電池関連物質であるイオン液体分子の原子構造(N-methyl-N-propylpiperidinium bis trifluoromethanesulfonyl imide, PP13-TFSI)。

(G) その他(行列データベース公開など)

研究に進展にしたがって、物理と数理の接点としてのデータベースや擬似(ミニ)アプリの必要性が認識されてきたため、疎行列データベース(最大 100 万次元まで)[23]・固有値問題むけ擬似(ミニ)アプリ[24]を公開した

### (2) 当初計画の達成状況について

(1)であげたテーマ(A)-(G)のうち当初計画は(C)-(F) であるが全て達成され、発展テーマとして(A)(B)(G)となった。特に(C)(アルゴリズムの実装と最適化)は当初予想以上に達成され、テーマ(A)(B)へと発展した。

結果として、昨年度の 10 倍となる 1 億原子(100 ナノスケール, 世界最大)計算が「京」コンピュータで達成され、全コア計算に至るまでの強スケーリング性が示せたことは、当初目的を遥かに超える成果と考えている。

(注 6) Takeo Hoshi, Yohei Akiyama, Tatsunori Tanaka and Takahisa Ohno, "Ten-million-atom electronic structure calculations on the K computer with a massively parallel order-N theory", J. Phys. Soc. Jpn. 82, 023710, 4pp (2013)

(注 7)通常の(密行列型固有値問題)計算では、計算コストが $O(N^3)$ でありメモリコストは $O(N^2)$ であるため、非常に大きな系($N\to\infty$)では計算コストがボトルネックになることが多い。しかし本研究のオーダー$N$ 法では、計算コスト・メモリコストともに $O(N)$ のため、どちらがボトルネックになるかはケースバイケースである。

(注 8) 電池関連物質研究[14]は、山崎久嗣(トヨタ自動車)を共著者とする、産学連携研究である。

## 4. 今後の展望

アモルファス状共役高分子系など 100 ナノスケール実デバイスを目標とした応用研究を行うと同時に、理論拡張(計算できる物理量を増やす)ことを目指したい。理論拡張としては特に、ナノエレクトロニクスの基礎である量子電気伝導計算(非平衡グリーン関数理論)(図 7)の超大規模超並列計算に注力したい(注 9)。

(注 9) 量子電気伝導(非平衡グリーン関数)計算コードは自作しているが、超大規模・超並列計算には対応していない:Hiroshi Shinaoka, Takeo Hoshi and Takeo Fujiwara, "Ill-Contact Effects of d-Orbital Channels in Nanometer-Scale Conductor", J. Phys. Soc. Jpn. 77, 114712, (2008)

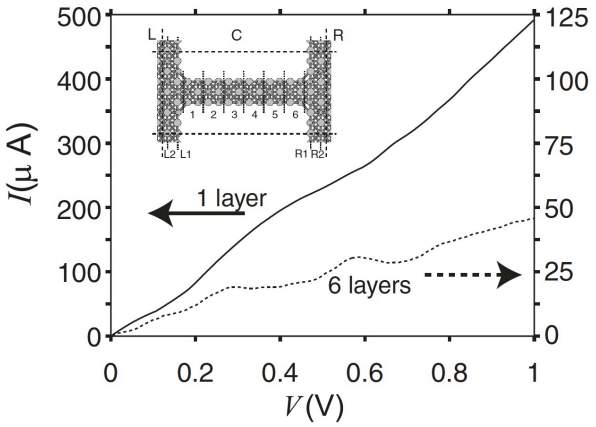



図 7: 量子電気伝導計算の例(注 9)。金属ナノワイヤ系における、電流-電圧(I-V)曲線。

## 5. 研究成果リスト

### (1) 学術論文

[1] S. Nishino and T. Fujiwara, "Parametrization scheme with accuracy and transferability for tight-binding electronic structure calculations with extended Huckel approximation and molecular dynamics simulations", J. Mol. Model. 19, 2363 (2013).

[2] A. Imakura, T. Sogabe, and S.-L. Zhang, "An efficient variant of the restarted shifted GMRES for solving shifted linear systems", J. Math. Res. Appl., 33, 127 (2013).

[3] J. Jia, T. Sogabe, and M.E.A. El-Mikkawy," Inversion of k-tridiagonal matrices with Toeplitz structure", Comput. Math. Appl., 65 (2013).

[4] T. Miyata, S. Honda, R. Naito, S.-L. Zhang, "A numerical approach to surface Green's functions via generalized eigenvalue problems", Japan Journal of Industrial and Applied Mathematics, 30, pp 653-660 (2013).

### (2) 国際会議プロシーディングス

[5] T. Hoshi, K. Yamazaki, Y. Akiyama, " Novel linear algebraic theory and one-hundred-million-atom electronic structure calculation on the K computer ", JPS Conf. Proc. 1, 016004, 4pp. (2014).

[6] Y. Ohtani, T. Fujiwara, S. Nishino, T. Suzuki, S. Yamamoto, and Y. Zempo, "Automatic Determination of Tight Binding Parameters in Bulk Systems", MRS Proceedings 1523, mrsf12-1523-qq06-08 (2013)

[7] Dongjin Lee, Takafumi Miyata, Tomohiro Sogabe, Takeo Hoshi, Shao-Liang Zhang, " An interior eigenvalue problem from electronic structure calculations ", Japan J. Indust. Appl. Math. 30, 625 (2013)

### (3) 国際会議発表

[8] (招待講演) Takeo Hoshi, Takafumi Miyata, Tomohiro Sogabe, *Shao-Liang Zhang*, ' Krylov subspace theories and ultra-large-scale electronic state calculations on the K computer', 18-20., June, 2013, The 9th East Asia SIAM Conference The 2nd Conference on Industrial and Applied Mathematics (EASIAM-CIAM 2013) , Institut Teknologi Bandung, Bandung, West Java, Indonesia.

[9] (招待講演) Takeo Hoshi, Takafumi Miyata, Tomohiro Sogabe, Shao-Liang Zhang, "Novel linear algebraic theory and one-hundred-million-atom quantum material simulations on the K computer", 14-17., Oct., 2013, The First International Workshop on Computational Science and Engineering , National Taiwan University Taipei, Taiwan

[10] Dongjin Lee, Takafumi Miyata, Takeo Hoshi, and Shao-Liang Zhang, "An Interior Eigenvalue Problem and Its Numerical Solution", The 9th East Asia SIAM Conference The 2nd Conference on Industrial and Applied Mathematics (EASIAM-CIAM 2013) , Institut Teknologi Bandung, Bandung, West Java, Indonesia

[11] Takeo Hoshi, Keita Yamazaki, Yohei Akiyama, 'Novel linear algebraic theory and one-hundred-million-atom electronic structure calculation on the K computer', 14-19., Jul, 2013, The 12th Asia Pacific Physics Conference (APPC12), Makuhari Messe, Chiba.

### (4) 国内会議発表

[12] (招待講演) 星健夫,”京での超大規模量子物質シミュレーションとクリロフ部分空間”,日本応用数理学会三部会連携「応用数理セミナー」 , 東京大学, 12 月 27 日 (2013).

[13] (招待講演)星健夫, 曽我部知広, 宮田考史, 李東珍,張紹良,井町宏人,山元進,西野信也,秋山洋平,山﨑渓太, “数学・計算物質科学・HPC 分野の融合としての超大規模電子状態計算”, 東京大学物性研スーパーコンピュータ共同利用・CMSI 合同研究会 , 東京大学, 12 月 10-13 日 (2013)

[14] 西野信也,藤原毅夫,山崎久嗣, “固体電解質 $Li_{4-x}Ge_{1-x}P_xS_4$ におけるリチウム拡散機構”, 日本物理学会 , 3 月 27-30 日, 東海大学 (2014)

[15] 星健夫, 秋山洋平, 山崎渓太, 横山誠也,” アモルファス状共役高分子系の超大規模電子状態計算”, 日本物理学会 , 3 月 27-30 日, 東海大学 (2014)

[16] 星健夫,”京における 100 ナノスケール電子状態計算と大行列数理解法”, セミナーレクチャー, 神戸大学, 11 月 11 日(2013).

[17] 山崎渓太, 秋山洋平, 星健夫,″Python を用いた電子状態計算むけ波動関数可視化ツールの開発 ″,日本物理学会年会, 徳島大学, 9 月 25-28 日(2013).

[18] 星健夫, 秋山洋平, “京コンピュータにおける超大規模電子状態計算”, 日本物理学会年会, 徳島大学, 9 月 25-28 日(2013).

[19] 星健夫, “京コンピュータにおける超大規模量子物質シミュレーション”, 日本応用数理学会 2013 年度年会,オーガナイズドセッション「ペタスケール環境における固有値解法」, アクロス福岡, 福岡, 9 月 25-28 日 (2013)

[20] 山崎渓太, 秋山洋平, 星健夫, “Python を用いた電子状態計算むけ波動関数可視化ツールの開発”,応用物理学会 中国四国支部学術講演会, 香川大学, 香川, 7 月 27 日(2013).

[21] 秋山洋平, 星健夫, “超大規模電子状態計算におけるナノ複合カーボンの可視化解析”, 応用物理学会 中国四国支部学術講演会, 香川大学, 香川, 7 月 27 日(2013).

[22] 星健夫, 秋山洋平, “超大規模超並列電子状態計算を中核とした物理・数理・HPC の融合研究”,学際大規模情報基盤共同利用・共同研究拠点第 5 回シンポジウム, THE GRAND HALL (品川) , 東京, 7 月 11-12 日 (2013).

[23] 山崎渓太, 秋山洋平, 星健夫, “Python を用いた波動関数可視化ツールの開発”, 科研費新学術領域「コンピューティクスによる物質デザイン」研究会、東京大学, 7 月 8-9 日 (2013).

[24] 星健夫, “京コンピュータにおける 1 億原子電子状態計算”, 科研費新学術領域「コンピューティクスによる物質デザイン」研究会、東京大学, 7 月 8-9 日 (2013).

**(5) その他 (特許, プレス発表, 著書等)**

[25] 行列データベース公開(実問題における大規模疎行列データ) ; ELSES matrix library; http://www.elses.jp/matrix/

[26] 固有値問題計算むけ擬似 (ミニ) アプリ公開 ; Eigen Test; http://www.damp.tottori-u.ac.jp/~hoshi/eigen_test/

[27] 第 18 回 (2013 年度) 応用物理学会中国四国支部学術講演会発表奨励賞受賞 ; 発表[21]に対して。