

T2K と ES2 による密行列固有値ソルバの性能評価

片桐 孝洋

東京大学情報基盤センター 特任准教授

1. はじめに

ペタフロップスコンピュータを複数台所有できる時代となった。並列数 (コア数) は 10 万コアにも達する。ペタフロップスの演算性能をもつ計算機環境で世界最高性能を目指すアプリケーションは、10 万の超並列性を 1 つのアプリケーション内で達成しなくてはならない。

従来開発されてきた分散メモリ並列計算機用の密行列用数値計算ライブラリは、高効率並列実行のために超大規模行列での実行を前提としている。2010 年 6 月現在、スーパーコンピュータのランキングサイト TOP500 において 1 位のオークリッジ国立研究所 Jaguar では、密行列の連立一次方程式の次元数が約 550 万次元にも達している。従来の数値計算ライブラリは、このような大規模行列で高効率になることを前提に設計され、特化したアルゴリズムと実装方式が実現されている。

ところが密行列ソルバは計算量が $O(n^3)$ で増加するため、超大規模行列の実行は並列化されているとはいえ「事実上の」時間的制約がある。たとえば、スーパーコンピュータ運営上の制約で、1 ユーザが全系占有できる時間は高々数時間に限定されるであろう。そこで論文[1]では、小規模行列に対してコア数を増加しても台数効果が見込める超並列固有値ソルバの開発を目的とした。通信時間削減と単体性能向上の観点から従来方式の改良を行った。

本原稿の目的は以下の通りである。従来から開発してきた自動チューニング (AT) 機能付き数値計算ライブラリ ABCLib_DRSSSED[2] は階層キャッシュを有するマルチコアを CPU としてノードにもつスーパーコンピュータでは、ブロック化がされていないため大規模サイズで性能の低下を生じていた。しかし、ABCLib_DRSSSED はベクトル計算機向きの演算実装となっている。このため、現存するベクトル型スーパーコンピュータ NEC SX-9 の性能評価が興味深い。そこで、独立行政法人海洋開発研究機構の地球シミュレータセンターに設置されている NEC SX-9/E によるスーパーコンピュータ (ES2) と、2008 年 6 月から東京大学情報基盤センターで稼働している T2K オープンスパコン (HITACHI HA8000 クラスタシステム) との性能比較を行うことを目的にした。

本原稿の構成は以下のとおりである。2 節で AT 機能付き固有値ソルバ ABCLib_DRSSSED の概略を説明する。3 節では、ES2 を利用した ABCLib_DRSSSED の性能評価を行う。最後に、本稿で得られた知見を述べる。

2. AT 機能付き固有値ソルバ ABCLib_DRSSSED

(1) 概要と背景

ABCLib_DRSSSED[2] は、対称実数固有値問題の任意の個数の固有値・固有ベクトルを計算できる機能をもつ並列数値計算ライブラリである。現在公開されている ABCLib_DRSSSED 1.04 は、

MPI-1 と Fortran90 で実装されている。ライブラリが提供する最適化方針と性能パラメタの定義範囲 (実装方式) 内で性能が自動的に最適化される機能を有し、これを自動チューニング (AT) 機能とよぶ。性能モデリングとして、任意次数の多項式近似による実行時間予測機能、および、サンプリング点の指定機能が実装されている。

ライブラリ上のアルゴリズムとして、対称固有値ソルバに利用される密対称行列用 Householder 三重対角化、三重対角対称行列用の固有値計算のための二分法、三重対角対称行列用の固有ベクトル計算のための逆反復法、および密対称行列の固有ベクトル計算のための Householder 逆変換の各ルーチンが利用できる。さらに、密行列用 QR 分解ルーチン (ただし、行列 R を保存せず、直交化の目的で直交行列 Q のみ求める機能) も提供している。

ABCLib_DRSSSED は、ベクトル化 (もしくは、疑似ベクトル化) 機能を有するプロセッサをもつノードにおいて、超並列環境 (2000 年頃において 1000 並列以上) で高速に動作する方式が採用されている。したがって、現在主流となっている階層キャッシュを有するプロセッサからなる超並列計算機環境においては、行列サイズが大きくなるとキャッシュミスを生じて性能低下が起こる。ただし台数効果では、十分な性能向上が期待できる。

また論文 [1] では、固有ベクトル計算に先進的解法である MRRR 法を実装することで、従来では計算時間の 90% 以上を占めていた固有ベクトル計算部分を、T2K オープンスーパーコンピュータにおいて、固有ベクトル計算の占める割合が 1% 未満となる方式を実現した。したがって、キャッシュをもつマルチコア計算機 (スカラマシン) では MRRR 法を採用すると、ほとんどの計算時間が Householder 三重対角化と Householder 逆変換の時間となる。

(2) T2K オープンスーパーコンピュータ (東大) における性能

ABCLib_DRSSSED の Householder 三重対角化 (TRD)、Householder 逆変換 (HIT)、および MGS 法による直交化 (QR 分解) (MGSAO) の変換各処理の GFLOPS 値を 図 1 にそれを示す。

図 1 では、TRD と HIT において行列サイズ 2,000 を超えてから演算性能が $1/4 \sim 1/5$ に低下する。一方、MGSAO では、このような性能低下がない。MGSAO はブロック化アルゴリズムを採用していることから、ブロック化の効果がきわめて大きいことを意味している。また、行列サイズ 1,000 の時のワーキングスペースは、TRD と HIT では概算で $4 * N$ (ここで、 N は行列サイズ) となるため $4 * 1000 * 8 \approx 32K$ である。したがって L1 キャッシュにデータが乗る上限となることが予想される。このような振る舞いは、キャッシュを有する CPU では当然である。

一方、キャッシュに乗るサイズでもピーク性能 (147.2 GFLOPS) に対する効率は 12% ~ 17% であり十分でない。この理由は ABCLib_DRSSSED 1.04 の実装においては、(1) TRD と HIT はブロック化が採用されていないこと ; および、(2) MGSAO ではブロック化が採用されているがループアンローリングやタイリングなどの自動チューニング項目が最内側ループに対して行われていない ; という、AT ライブラリ自体の最適化戦略から生じたものと推測される。ABCLib_DRSSSED 開発時に想定した計算機では、最内側ループ長が長くとも性能劣化が起こりにくいアーキテクチャを想定していたためである。

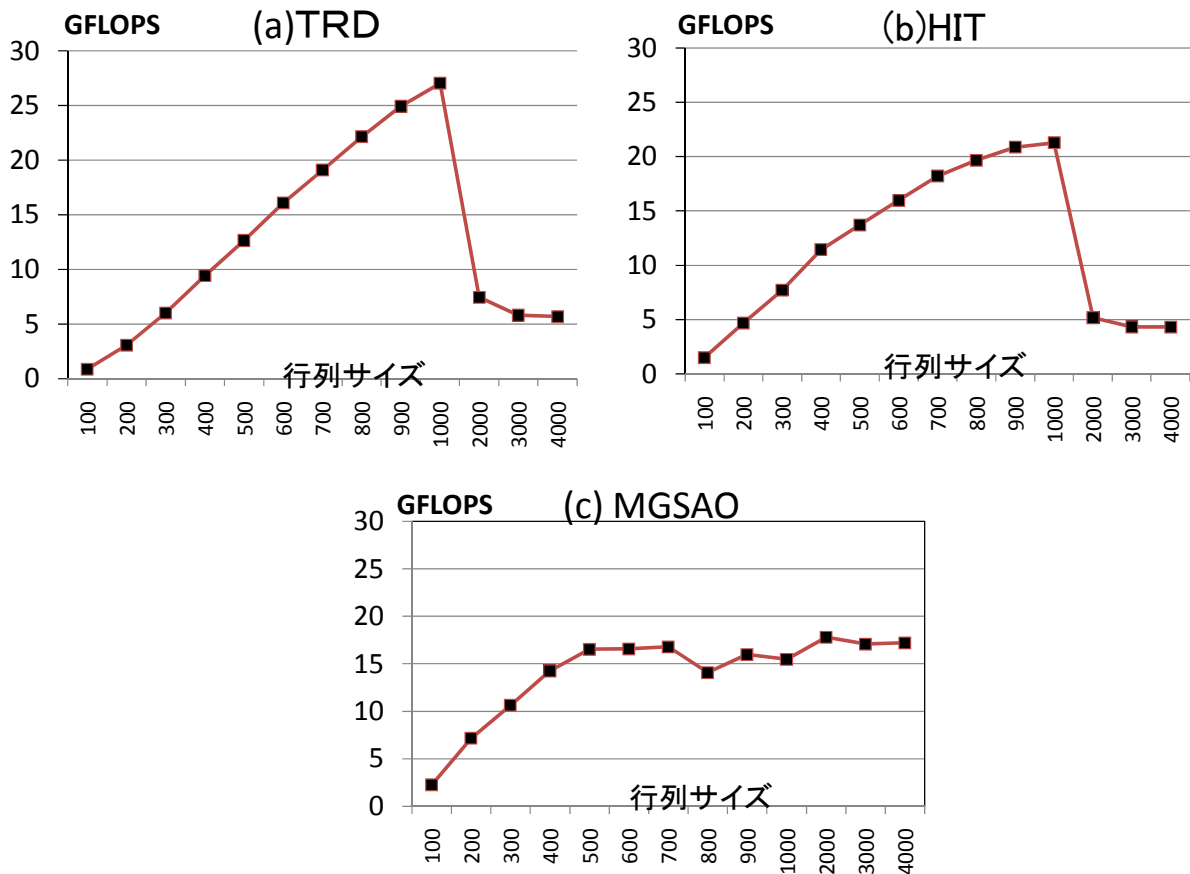


図 1 T2K オープンスーパーコンピュータ（東大）での各処理の GFLOPS 値
 （1 ノード(16 コア)、ピュア MPI 実行）
 (a)TRD : Householder 三重対角化、(b)HIT: Householder 逆変換、
 (c)MGS AO: MGS 法による直交化 (QR 分解)

(3) ベンチマークとしての特徴

ベンチマークの観点からの ABCLib_DRSSSED の各処理について以下にまとめる。

- 対称実数固有値ソルバ
 - Householder 三重対角化部（以降、TRD）
 - ◇ BLAS 2 演算性能評価
 - 行列-ベクトル積演算性能
 - 行列更新演算性能
 - ◇ マルチキャスト通信性能評価
 - プロセッサのグリッド ($p \times q$) における、同時放送処理 p 個（従事プロセッサ数 q 個）の性能評価
 - Householder 逆変換部（以降、HIT）
 - ◇ BLAS 1 演算性能評価
 - 必要な固有ベクトル数(k)に応じて BLAS 1- k 更新となる行列更新演算性能。全固有ベクトルが必要なときは、BLAS 2 演算となる。
 - ◇ コレクティブ通信性能評価
 - 集団通信性能評価 (Gather 演算)。

- 直交化 (QR 分解)
 - 修正 Gram-Schmidt 演算部 (以降、MGSAO)
 - ◇ BLAS3 演算性能評価
 - 行列更新演算性能評価。ただし、最外ループの刻み幅はブロック幅となる。
 - ◇ 1 対 1 通信性能評価
 - 通信粒度がブロック幅の逆数で変化する。つまり通信回数が、上記 BLAS3 のブロック幅の逆数となる実装における評価。

対称実数固有値ソルバではブロック化が実装されていないので、階層メモリを有するプロセッサで、かつキャッシュミスヒット時の性能劣化が大きなプロセッサでは、大規模問題実行時に性能劣化が引き起こることが予想される。一方、QR 分解ではブロック化アルゴリズムの採用により、大規模問題実行時の性能劣化を食い止めることが期待できる。ブロッキングある／なしでの性能差も、ABCLib_DRSSSED におけるプロセッサ性能の評価指標の 1 つである。

(4) 性能パラメタの説明

AT 性能パラメタの観点では、ABCLib_DRSSSED の性能パラメタは以下に分類される [2]。

- 対称実数固有値ソルバ 性能パラメタ
 - Householder 三重対角化用
 - ◇ 通信実装方式切り替え *ictr*
 - ◇ 行列-ベクトル積アンローリング段数制御 *imv*
 - ◇ 行列更新演算アンローリング段数制御 *iud*
 - Householder 逆変換用
 - ◇ 通信実装方式切り替え *ichit*
 - ◇ 行列更新演算アンローリング段数制御 *ihit*
- QR 分解 性能パラメタ
 - ブロック幅調整 *ibl*
 - 枢軸ブロック演算用
 - ◇ 最外ループアンローリング段数制御 *iop*
 - ◇ 第 2 ループアンローリング段数制御 *isp*
 - それ以外の演算用
 - ◇ 最外ループアンローリング段数制御 *ioo*
 - ◇ 第 2 ループアンローリング段数制御 *iso*

(5) 性能パラメタの定義域：現在の実装における制約

ABCLib_DRSSSED 1.04 では、自動チューニングパラメタについて、以下の実装がなされている。

- 対称実数固有値ソルバ 性能パラメタ定義域
 - Householder 三重対角化用
 - ◇ 通信実装方式切り替え
 - $ictr = \{1 \text{ 対 } 1 \text{ 通信関数を用いた実装, 集団通信関数を用いた実装}\}$
 - ◇ 行列-ベクトル積アンローリング段数制御
 - $imv = \{1 \text{ 段}, 2 \text{ 段}, \dots, 16 \text{ 段}\} : \text{最外ループのみ}$

- ◇ 行列更新演算アンローリング段数制御
 $iud = \{1 \text{ 段}, 2 \text{ 段}, \dots, 16 \text{ 段}\} : \text{最外ループのみ}$
- Householder 逆変換用
 - ◇ 通信実装方式切り替え
 $ichit = \{\text{放送関数を用いた実装}, 1 \text{ 対 } 1 \text{ ブロッキング通信関数を用いた実装}, 1 \text{ 対 } 1 \text{ ノンブロッキング関数を用いた実装}\}$
 - ◇ 行列更新演算アンローリング段数制御
 $ihit = \{1 \text{ 段}, 2 \text{ 段}, \dots, 16 \text{ 段}\} : \text{最外ループのみ}$
- QR 分解 性能パラメタ定義域
 - ブロック幅調整
 $ibl = \{1, 2, 3, 4, 8, 16\}$
 - 枢軸ブロック演算用
 - ◇ 最外ループアンローリング段数制御
 $iop = \{1 \text{ 段}, 2 \text{ 段}, 3 \text{ 段}, 4 \text{ 段}\}$
 - ◇ 第2ループアンローリング段数制御
 $isp = \{1 \text{ 段}, 2 \text{ 段}, 3 \text{ 段}, 4 \text{ 段}, 8 \text{ 段}, 16 \text{ 段}\}$
 - それ以外の演算用
 - ◇ 最外ループアンローリング段数制御
 $ioo = \{1 \text{ 段}, 2 \text{ 段}, 3 \text{ 段}, 4 \text{ 段}\}$
 - ◇ 第2ループアンローリング段数制御
 $iso = \{1 \text{ 段}, 2 \text{ 段}, 3 \text{ 段}, 4 \text{ 段}, 8 \text{ 段}, 16 \text{ 段}\}$

3. 性能評価

(1) 評価環境

地球シミュレータシステム 2 (ES2) のノードは NEC SX-9/E で、8 コア (8CPU) を搭載しており、メモリは 128GB である。理論最大演算性能は、ノードあたり 819.2GFLOPS である。CPU の動作周波数は 3.2MHz で、1 クロック当たり 4 浮動小数点演算が可能であり、ベクトルパイプラインが 8 つある。したがって、1CPU 当たり $3.2\text{GHz} \times 4\text{FLOPS} \times 8 = 102.4\text{GFLOPS}$ の性能となる。

ノード間結合網は 2 段ファットツリーでフルバイセクション結合である。ノード当たりのバンド幅は $8\text{GB/sec} * 8\text{plane} = 64\text{GB/sec}$ である。2 段目スルーput は $8\text{G/sec} * 2 \text{ line} * 10\text{pair} * 8\text{SW} * 8\text{plane} = 10\text{TB/sec}$ である。

コンパイラは、NEC SX Fortran90 Rev.404 2010/03/01 で、コンパイラオプションは、**ピュア MPI** (MPI のみによる実行) で、-chopt である。実験期間は、2010 年 7 月 9 日～7 月 20 日である。

以降の性能評価では、問題サイズ (100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000) において性能パラメタを全て自動チューニングして最高速となるパラメタが自動設定されている。また、N=10,000 実行については、8,000 の最適パラメタを設定し実行されている。

ABCLib_DRSSD 1.04 におけるデフォルトパラメタは、以下のとおりである。

- $ictr = \{1 \text{ 対 } 1 \text{ 通信関数を用いた実装}\}$

- $imv = \{8 \text{ 段}\}$
- $iud = \{6 \text{ 段}\}$
- $ichit = \{\text{放送関数を用いた実装}\}$
- $ihit = \{1 \text{ 段}\}$
- $ibl = \{4\}$
- $iop = \{4 \text{ 段}\}; isp = \{8 \text{ 段}\};$
- $ioo = \{4 \text{ 段}\}; iso = \{8 \text{ 段}\};$

以上のデフォルトパラメタに対する AT による速度向上について評価する。アンローリング段数を 1 段に固定し、コンパイラ最適化による効果のみと比べたものではない¹。コンパイラ最適化のみより、本デフォルトパラメタの方が経験的に高速となる。

(2) 自動チューニングの効果とベクトル機での性能

地球シミュレータ ES2 の 1 ノード (8 コア) による ABCLib_DRSSSED 1.04 の実行性能を図 2 にのせる。プロセッサ・グリッドは 2x4 である。

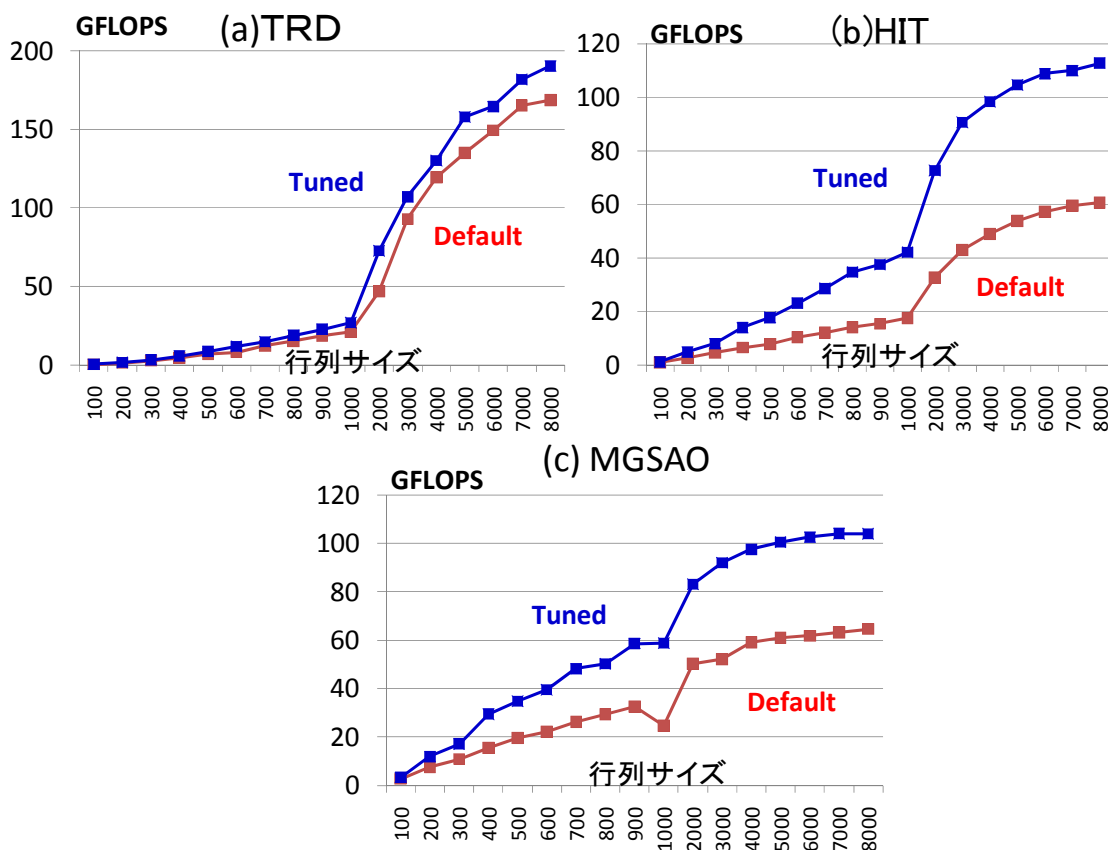


図 2 ES2 での各処理の GFLOPS 値 (1 ノード(8 コア)、ピュア MPI 実行)。

Default は ABCLib_DRSSSED のデフォルトパラメタによる実行性能。

Tuned は ABCLib_DRSSSED の自動チューニング後による実行。

図 2 から、自動チューニングの効果は処理により異なるが、TRD で 1.1x-1.5x、HIT で 1.3x-2.4x、MGSAO で 1.3x-2.3x 程度である。この効果は T2K より高いが、ABCLib_DRSSSED のカ

¹ これをベースラインとして AT の効果を評価する論文も多い。このベースラインを用いた評価に比べ、本性能評価の基準では経験的に AT の効果が低めに算出される。

一ネルがベクトル計算機向きになっていることから驚くべきことではない。

また GFLOPS 性能であるが、ベクトル計算機の特徴から問題サイズが大きくなるほど GFLOPS 値が高くなる。特に、問題サイズ 1,000 以下は GFLOPS 値の伸びが低い。2,000 以上でより効果的に GFLOPS 値が向上していく。これは、ベクトル長と性能向上（立ち上げオーバーヘッドの隠ぺい効果）の関連である。さらに、TRD と HIT は $N=8,000$ で飽和性能でないと思われる。さらに大規模サイズで GFLOPS 値が伸びていくと思われる。

性能を考察する。

小規模サイズ 1,000 のとき性能は、TRD で 26.8GFLOPS、HIT で 42.1GFLOPS、MGSAO で 58.8GFLOPS である。T2K の同サイズの性能と比べると、それぞれ、約 25GFLOPS、約 20GFLOPS、約 15GFLOPS なので、TRD についてはほぼ同等、HIT については約 2 倍、MGSAO については約 4 倍である。

中規模サイズ 8,000 のとき性能は、TRD で 190GFLOPS、HIT で 112GFLOPS、MGSAO で 104GFLOPS である。T2K の $N=4,000$ の性能と比べると、T2K は飽和性能であるので、それぞれ、約 5GFLOPS、約 5GFLOPS、約 15GFLOPS となる。TRD については約 38 倍、HIT については約 22 倍、MGSAO については約 7 倍の速度向上があるといつてよい。

以上のように、キャッシュサイズに収まるような小規模サイズではベクトル機の恩恵が少なく、最大でも約 4 倍の速度向上となる。ところが、キャッシュサイズに収まらないアクセスについて、ブロック化されていないプログラムで性能向上の効果が最大で 38 倍と大きく、ベクトル機の恩恵を享受できる。

一方、 $N=8,000$ でのピーク性能に対する効率は 23.1%、TRD では 13.6%、MGSAO では 12.6% である。これらは飽和性能ではないのでさらに向上する可能性がある。TRD のような三重対角化処理は両側変換とよばれ、行列サイズが処理中に変化する行列積処理や、反復ごとに行列サイズが縮小していくのは TRD と同じであるが、片側変換となる LU 分解や QR 分解に比べ、実行性能が出ない。このような両側変換処理であることを考慮すると、TRD の性能は悪くはない性能といえる。一方、TRD や MGSAO は不十分であるが、通信最適化やストリップマイニングなどベクトル向き実装が考慮されていないので、さらに演算効率を上げられる。また、自動チューニング機能による人手によらない最適化による速度向上を考慮すると、開発コストの観点から状況により妥当な性能といえるかもしれない。

(3) 小規模行列での実行における処理の内訳

図 3 に並列処理においては小規模行列となる $N=10,000$ の実行に対し、固有値ソルバにおける各処理の内訳を 1 ノード (4 コア (2x2 構成))、1 ノード (8 コア (2x4 構成))、および、4 ノード (2 コア (4x8 構成)) のものについて載せる。なお、テスト行列は Frank 行列である。

図 3 から、MRRR 法における固有ベクトル計算の部分が ES2 にとっては大部分を占め、約 74% となる。T2K では占める割合が約 1% であったが、これは、MRRR 法の主計算である二分法がベクトル化できる実装ではなく、スカラ計算機向きであることによると思われる。この解決には、ベクトル機向きの多分法や多固有値多分法[3]を実装するか、分割統治法などの行列積ベースの方法に変更する必要がある。一方、Householder 三重対角化と Householder 逆変換の占める割合は高々 20% である。本性能評価は、全固有値、全固有ベクトルを計算する場合で、固有ベクトル計算部分の演算時間が最悪となる場合であるが、ABCLib_DRSSD の ES2 上での実行では、固有ベクトル計算部分の高速化が必須であるといえる。

実行時間[秒]

N=10,000

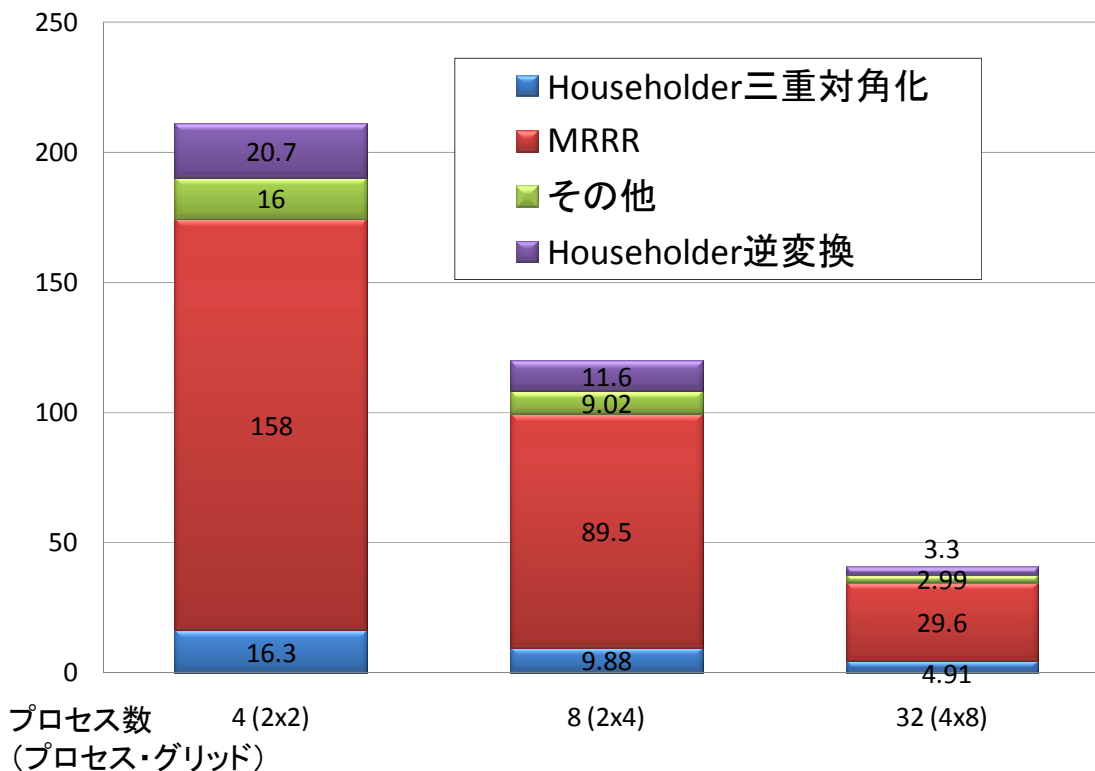


図 3 小規模問題サイズ (N=10,000) による固有値ソルバにおける各処理の内訳。

ES2 の 1 ノード (4 コア (2x2 構成))、1 ノード (8 コア (2x4 構成))、
および、4 ノード (2 コア (4x8 構成))。

演算精度については、4 ノード (32 プロセス) 実行時において、真の固有値からの最大誤差は $7.417\text{D}-9$ 、固有ベクトル直交性 (フロベニウスノルム $\|X^T X - I\|_F$) の最大値は $1.559\text{D}-10$ 、および固有方程式の最大残差に関する 2 ノルム ($\|A x_i - \lambda_i x_i\|_2$) の最大値は $4.597\text{D}-8$ であった。

3. おわりに

本稿では、自動チューニング機能付き固有値ソルバ ABCLib_DRSSSED をベクトル計算機である地球シミュレータ ES2 で評価し、同規模の実行性能を T2K オープンスパコン (東大) の結果と比較した。

性能評価の結果、ES2 での実行性能はベクトル計算を考慮した ABCLib_DRSSSED の実装により比較的良好で、T2K の時に生じた大規模サイズでの性能劣化を生じない。また、Householder 三重対角化においては、1 ノードあたりのピーク性能比 23.1%を達成した。また、中規模サイズで T2K の 1 ノード当たりの性能向上比は、最高で約 38 倍程度になった。

一方、利点ばかりでなく、ベクトル機特有の欠点も明らかとなった。1,000 次元以下の超小規模実行では T2K との性能差が無いか、あっても高々 4 倍程度の速度向上となる。ノード当たりの価格性能比からすると、コスト的に有効とは思えない状況もある。また、T2K で 1%以下であった MRRR 法による固有ベクトル計算部分が全体処理の 70%を占める。固有ベクトル計算部分の高速化が課題となる。

今後の課題として、分割統治法など固有ベクトル計算部分の解法を変更し、再度性能評価す

ることがあげられる。また、固有ベクトル計算の解法の選択自体を自動チューニング機能として入れ込むことも、スカラ機とベクトル機との計算機構成の違いを吸収するために重要な自動チューニング機能となる。演算のさらなる高速化としては、ストリップマイニングなど、ベクトル計算機向きの実装方式を自動チューニング機能として入れ込み、さらなる高性能化を実現する方式を開発することがあげられる。

謝辞

本研究の一部は、平成 22 年度 地球シミュレーター一般公募プロジェクト、先進・創出分野、「問題適合型高精度計算ライブラリの開発」（代表：長谷川秀彦（筑波大学 大学院図書館情報メディア研究科））の支援による。

参 考 文 献

- [1] 片桐孝洋：ペタフロップス環境における小規模行列用対称密行列固有値ソルバに向けて一逆変換の改良，情報処理学会論文誌：コンピューティングシステム（ACS），Vol. 3 No. 2, pp. 1-8（2010 年 6 月）
- [2] Takahiro Katagiri, Kenji Kise, Hiroki Honda, and Toshitsugu Yuba: ABCLib_DRSSD: A Parallel Eigensolver with an Auto-tuning Facility, Parallel Computing, Vol. 32, Issue 3, pp. 231-250（2006）
- [3] 片桐孝洋：対称三重対角固有値ソルバにおける多固有値多分法の開発，東京大学情報基盤センター スーパーコンピューティングニュース，VOL. 9, No. 5（2007 年 9 月）