

# 高性能・変動精度・高信頼性数値解析 手法とその応用

Numerical Methods with High-Performance, Trans  
Precision and High Reliability

jh200037-NAH

研究代表者：中島研吾（東京大学情報基盤センター）

第12回JHPCNシンポジウム（オンライン）  
2020年7月9日

# ポストムーア時代に向けた低・変動精度演算による新計算原理の必要性

- 低・変動精度演算
  - 計算量・計算時間・消費電力・メモリ容量・I/O削減
- 科学技術計算の多くは倍精度(64bit)で実施されているが, 条件の良い問題であれば, 単精度(32bit)・半精度(16bit), 混合精度でも可能
- 我々は不必要に高精度な計算を実施して貴重な時間やエネルギーを無駄にしている・・・かも知れない
- 昨今, 低精度演算の活用は盛ん
  - 混合精度演算の研究, アプリへの適用は既に行われている
  - Approximate Computing: 省電力のための研究
    - 元々は画像処理・機械学習
    - FPGAを含むハードウェア, システムソフトウェア, コンパイラ
  - ポストムーア時代へ向けた重要な技術の一つ

# 高性能・変動精度・高信頼性数値解析手法と その応用: 背景と目的



JHPCN (学際大規模情報基盤共同利用・共同研究拠点)

2018/2019/2020年度共同研究課題

代表: 中島研吾 (東京大学情報基盤センター・理研R-CCS)

副代表: 横田理央 (東京工業大学)・市村強 (東京大学地震研究所)

日米独 3ヶ国  
16機関 (産学官), 30+人

- 通信最適化 (Serial: メモリ階層, Parallel: 分散) による **高性能アルゴリズム**
- **低・変動精度** 演算の採用による省電力・省エネルギーの実現 (半精度・単精度・倍精度・4倍精度)
- 精度保証による **高信頼性アルゴリズムの実現**
  - 数値計算による近似解 (数値解) は様々な計算誤差を含み, 計算結果の信頼性の観点から, 数値解の正しさを数学的に保証する必要がある (精度保証)。
  - 低精度・変動精度使用時, 悪条件問題には重要, 実問題で現れる大規模疎行列・H行列への応用例は非常に少ない



WASEDA University  
早稲田大学



北海道大学  
HOKKAIDO UNIVERSITY



芝浦工業大学  
SHIBURA INSTITUTE OF TECHNOLOGY  
Established 1927



東京大学  
THE UNIVERSITY OF TOKYO



Tokyo Woman's Christian University  
東京女子大学



東京工業大学  
Tokyo Institute of Technology



名古屋大学



九州大学  
KYUSHU UNIVERSITY



国立研究開発法人  
国立環境研究所  
National Institute for Environmental Studies



RIST



RIKEN  
R-CCS



FAU  
FRIEDRICH-ALEXANDER  
UNIVERSITÄT  
ERLANGEN-NÜRNBERG



DLR  
Deutsches Zentrum  
für Luft- und Raumfahrt  
German Aerospace Center



BERKELEY LAB



intel



NVIDIA

中島研吾(東大)(代表)全体統括・高性能アルゴリズム  
市村 強(東大)(副代表)高性能アルゴリズム・地震シミュレーション  
横田理央(東工大)(副代表)高性能アルゴリズム・精度保証  
藤田航平(東大)高性能アルゴリズム・地震シミュレーション  
星野哲也(東大)高性能アルゴリズム・性能最適化  
坂本龍一(東大)電力測定・自動チューニング(AT)  
有間英志(東大)電力測定・AT  
古村孝志(東大)地震シミュレーション  
近藤正章(東大)精度保証・電力測定・AT  
奥田洋司(東大)工学シミュレーション・精度保証  
森田直樹(東大)工学シミュレーション・精度保証  
荻田武史(東女大)精度保証・AT  
田中一成(早大)精度保証・AT  
尾崎克久(芝工大)精度保証・AT  
河合直聡(東大)高性能アルゴリズム・量子科学シミュレーション  
Sameer Deshmukh(東工大)高性能アルゴリズム・精度保証  
大友広幸(東工大)高性能アルゴリズム・精度保証  
Peter Spalhoff(東工大)高性能アルゴリズム・精度保証  
Qianxiang Ma(東工大)高性能アルゴリズム・精度保証  
Muhammad Ridwan Apriansyah(東工大)高性能アルゴリズム・精度保証  
Gerhard Wellein(FAU Erlangen-Nürnberg, Germany)高性能アルゴリズム・量子科学シミュレーション・SELL-C- $\sigma$   
Achim Basermann(DLR, Germany)高性能アルゴリズム・量子科学シミュレーション  
Osni Marques(Lawrence Berkeley National Laboratory, USA)高性能アルゴリズム・AT

岩下武史(北大)高性能アルゴリズム  
深谷 猛(北大)高性能アルゴリズム  
池原 紘太(北大)高性能アルゴリズム  
塙 敏博(東大)性能最適化・電力測定  
伊田明弘(東大)高性能アルゴリズム  
片桐孝洋(名大)高性能アルゴリズム・精度保証・AT  
大島聡史(名大)高性能アルゴリズム・性能最適化  
山梨祥平(名大)高性能アルゴリズム・精度保証・AT  
八代 尚(環境研)大気シミュレーション  
井上弘士(九大)精度保証・電力測定・自動チューニング  
荒川 隆(RIST)大気シミュレーション  
成瀬 彰(NVIDIA)精度保証・AT  
堀越将司(インテル)精度保証・AT  
大島聡史(名大)高性能アルゴリズム・性能最適化

# 目的(1/2)

- エクサスケールシステムにおける高性能数値アルゴリズム実現には、メモリ・ネットワークの階層の深化に対応した通信最適化 (Serial, Parallel), 省電力・省エネルギー (以下「省電力」) に向けた検討が必要である。
- Approximate Computing (S. Mittal, A Survey of Techniques for Approximate Computing, ACM CSUR 48-4, 2016) は、低精度・混合精度演算の積極的活用により計算時間短縮, 消費電力削減を図る試みであり、従来は画像認識等の計算精度の要求されない分野を対象としていたが、昨今は数値計算において半精度から四倍精度まで演算精度を動的に変動させる変動精度 (Transprecision) も含めた研究が進められている。
- 数値計算による近似解 (数値解) は様々な計算誤差を含み、計算結果の信頼性の観点から、数値解の正しさを数学的に保証する必要があり、低精度・混合 / 変動精度使用時、悪条件問題には重要であるが、実問題で現れる大規模疎行列・H行列への応用例はほとんどない。
- 本研究では、JHPCNシステム群の中で消費電力当たり計算性能 (GFLOPS/W値) の高いシステムを主たるターゲットとして、以下を実施する:

## 目的(2/2)

- ① 疎行列演算, H行列演算, ステンシル演算等の代表的数値アルゴリズム, 各アプリケーション(地震, 大気科学, 量子科学, 構造力学)について, Serial・Parallel通信最適化に着目した高性能最適化手法を各システムに実装し, 低精度・混合/変動精度演算について検討し, 消費電力を測定する:アルゴリズム開発・消費電力測定
- ② 疎行列演算やH行列演算を対象として, 特に悪条件問題における実用的な精度保証法を確立する。更に①の各アルゴリズム, アプリケーションについて所望の結果精度達成という条件下で, 計算時間や消費電力を最小化する最適演算精度を自動チューニング技術により動的に制御する手法を確立する:精度保証・自動チューニング
- ③ 本研究によって開発された高性能・変動精度・高信頼性数値解法を, 自動チューニング機構を有するアプリケーション開発・実行環境ppOpen-HPC(JST-CREST 2011-2018, <https://github.com/Post-Peta-Crest/ppOpenHPC>)及び(計算+データ+学習)融合を実現する革新的ソフトウェア基盤h3-Open-BDEC(科研費基盤S 2019-2023, <http://nkl.cc.u-tokyo.ac.jp/h3-Open-BDEC/>)に実装し, 東大Oakforest-PACS(OFP), 東大Reedbush(RBH, RBL), 東大Oakbridge-CX(OBCX), 東工大Tsubame-3(TSB3)で公開する。将来的には「富岳」も含む他のセンターのスパコンへの導入も視野に入れる。

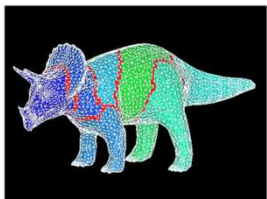
# ppOpen-HPC

自動チューニング機構を有する  
アプリケーション開発フレームワーク

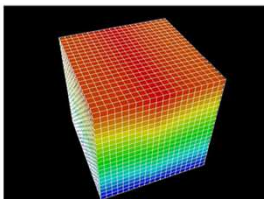


- JST/CREST + DFG/SPPEXA (2011-2018)
- 7機関, 50名以上が参加
- **オープンソースソフトウェア**

- ✓ <http://ppopenhpc.cc.u-tokyo.ac.jp/>
- ✓ <https://github.com/Post-Peta-Crest/ppOpenHPC>
- ✓ 英語ドキュメント, MITライセンス



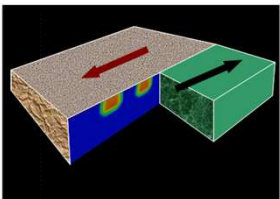
FEM  
Finite Element Method



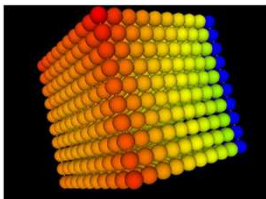
FDM  
Finite Difference Method



FVM  
Finite Volume Method



BEM  
Boundary Element Method



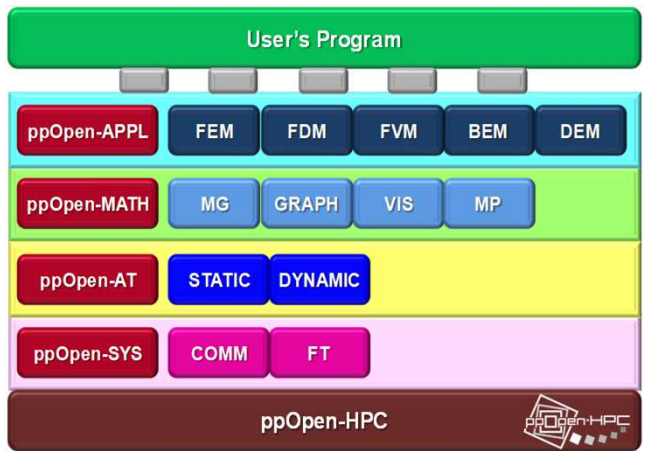
DEM  
Discrete Element Method

Framework  
Appl. Dev.

Math  
Libraries

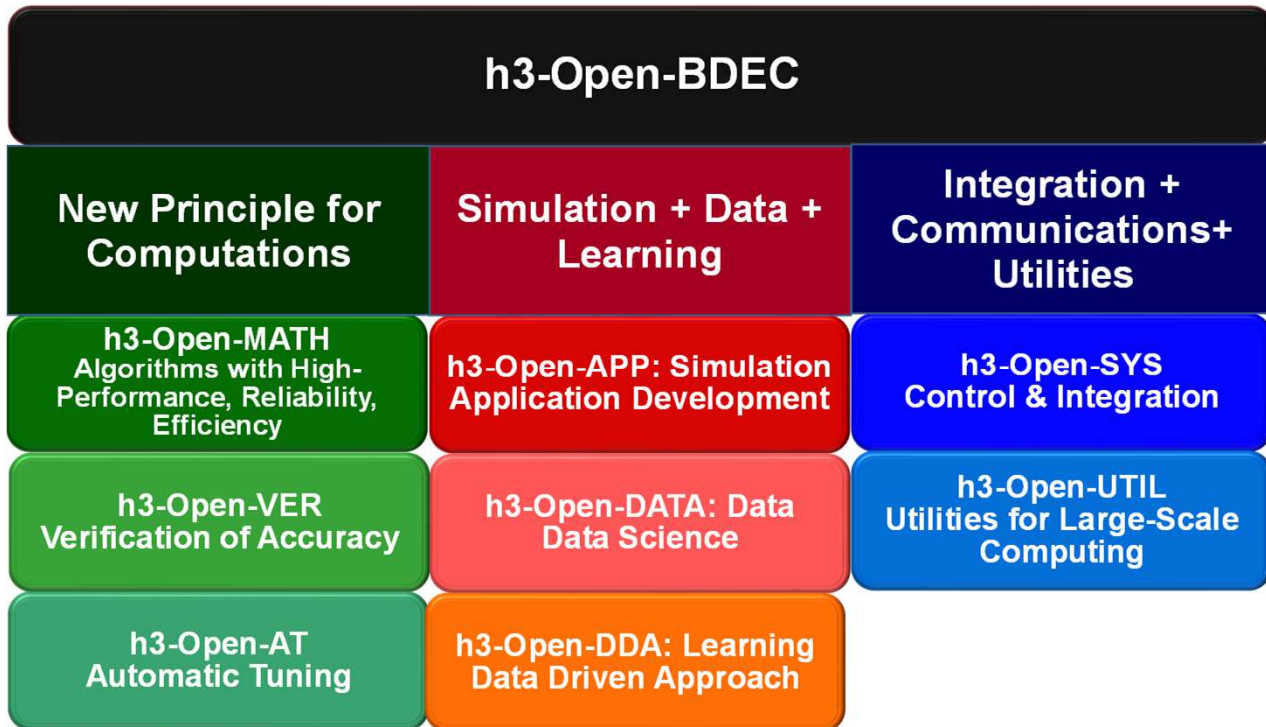
Automatic  
Tuning (AT)

System  
Software



# h3-Open-BDEC

Extension of ppOpen-HPC for Integration of (Simulation + Data + Learning) in the Exa/Post Moore Era, Initial Target: **BDEC@U.Tokyo**



**h3=**  
Hierarchical,  
Hybrid,  
Heterogeneous

**Big Data & Extreme Computing**



# JHPCN公募型共同研究として実施する必要性

- JHPCNは多様な計算機環境を備え、実用的なシステムとして最もGFLOPS/W値の高いOFF, RBH, RBL, OBCX, Tsubame-3等の大規模システムを有し、本研究の目指す高性能・変動精度・高信頼性数値解法の研究には最適である。
- RBL, OBCXでは「ノード固定」における設定カスタマイズにより、個別ノードの消費電力測定が可能である。
- JHPCNは様々な分野の専門家を擁し、本研究のような学際的研究を推進する体制を容易に構築でき、北大, 東大, 東工大, 名大, 九大各センターから様々な分野の研究者が参加している。
- JHPCN各センターはオープンソースソフトウェア活用に積極的であり、本研究の成果を公開, 各センターのスパコンにデプロイし, 講習会等の普及活動を協力して行い, 利用者拡大, ソフトウェアの更なる改良が可能となる。

# 研究の意義

- 本研究は、最先端のスパコン向けに開発された高性能数値アルゴリズムに対して、低精度・混合／変動精度演算を適用し、精度保証、そのための自動チューニング手法を開発する試みとしては初めてのものであり、様々なアプリケーションへの適用により、低精度・混合／変動精度演算の科学技術シミュレーションへの有効性を検証できる。
- 開発したアルゴリズム、アプリケーションの消費電力の直接測定によって、各計算の特性と低精度・混合／変動精度演算の有効性を消費電力の観点から検討可能となる。

# 本研究の沿革(1/3)

- 本研究は、計算科学・計算機科学・数値アルゴリズム分野の研究者の協力のもと、様々な数値アルゴリズムの最新アーキテクチャに向けた最適化、低精度・変動精度導入による高速化、消費電力節約、および精度保証に基づく最適精度選択のための自動チューニング選択手法の確立と実アプリケーションでの検証を目指したものである。
- 本研究では以下の各項目について研究開発を実施する
  - ステンシル計算(①構造格子, ②半構造格子)
  - 疎行列演算(③一般行列, ④悪条件問題, ⑤Adaptive CG(局所前処理による))
  - ⑥H行列, ⑦精度保証, ⑧消費電力測定, ⑨自動チューニング手法, の各項目についての研究開発を実施する。
- 本研究は2018年度から、3年計画として実施している。

# 研究実施 項目の概要

実施項目	担当者	目標精度	概要
①ステンシル計算: 構造格子	深谷・岩下・古村・成瀬	半・単	<ul style="list-style-type: none"> <li>三次元地震波動伝搬解析コード「Seism3D」(古村他)</li> <li>深谷・岩下の時空間タイリング手法による最適化</li> <li>変動精度演算を適用した大規模シミュレーションによる性能検証, 精度評価</li> </ul>
②ステンシル計算: 半構造格子	深谷・岩下・八代・荒川・成瀬	半・単	<ul style="list-style-type: none"> <li>三次元全球大気解析コード「NICAM」(八代他)</li> <li>時空間タイリング手法, 変動精度演算を適用した大規模シミュレーションによる性能検証, 精度評価</li> </ul>
③疎行列演算: ppOpen-MATH	岩下・中島・星野・塙・大島・奥田・森田・Wellein	半・単・倍	<ul style="list-style-type: none"> <li>ppOpen-HPC疎行列反復法ソルバー群(Pipeline法, Dynamic Loop Scheduling, hCGA法)(中島他)</li> <li>SELL-C-<math>\sigma</math>, 変動精度演算機能を適用</li> <li>不均質場固体力学・地下水流れ, FrontISTR(奥田)[2]等による性能検証, 精度評価</li> </ul>
④疎行列演算: pK-Open-SOL(悪条件)	河合・伊田・中島・星野・塙・大島・Wellein・Basermann	半・単・倍 ・4倍	<ul style="list-style-type: none"> <li>量子科学シミュレーションの固有値計算の過程で得られる悪条件問題を対象とする前処理付き反復法ソルバー</li> <li>悪条件問題向けのブロック化・対角シフト機能の他, 世界初の分散並列オーダーリング手法を実装(河合他)</li> <li>SELL-C-<math>\sigma</math>, 変動精度演算機能を適用</li> <li>量子科学分野の大規模シミュレーションによる性能検証, 精度評価, 悪条件問題における単精度実数演算適用の可能性について併せて検討</li> </ul>
⑤疎行列演算: Adaptive CG	市村・藤田・堀越・成瀬	半・単・倍	<ul style="list-style-type: none"> <li>三次元地震解析コード「GHYDRA」(市村・藤田)で使用されている手法, 混合精度演算も導入済</li> <li>半精度も含めた変動精度演算, 大規模シミュレーションによる性能検証, 精度評価</li> </ul>
⑥H行列演算: HACApK	伊田・岩下・横田・星野・塙・大島・河合・Basermann	単・倍 4倍	<ul style="list-style-type: none"> <li>ppOpen-HPCの一環として開発(伊田・岩下)</li> <li>変動精度演算の実装, 悪条件問題向けの前処理手法</li> <li>電磁気学, 地震発生サイクル等の大規模シミュレーションによる性能検証, 精度評価</li> <li>悪条件問題における単精度実数演算適用の可能性について併せて検討</li> </ul>
⑦精度保証手法	荻田・片桐・尾崎・奥田・横田・井上・近藤・森田・堀越・成瀬		③～⑥を対象として, 精度保証を適用可能な行列クラスを抽出し, 悪条件問題における実用的な精度保証法を確立するとともに, 数値解の精度を改善するための効率的な反復改良法(Iterative Refinement)を精度保証法と融合
⑧消費電力測定	近藤・井上・坂本・有間・塙		①～⑥を対象として, 近藤, 井上等によるJST-CREST「ポストベタスケールシステムのための電力マネージメントフレームワークの開発」の成果, 知見を適用し, RAPL(Running Average Power Limit)等によりReedbush-L, 東大情基セのクラスター(IBM P9+NVIDIA V100)等を使用した電力測定を実施し, 変動精度演算による消費電力削減効果を実証
⑨自動チューニング手法	片桐・荻田・横田・井上・近藤・有間		上記①～⑧の知見を元に, 所望の精度で計算時間や消費電力を最小化する最適演算精度を, アプリケーション・係数行列の性質, 問題サイズ, ハードウェア環境等に基づき自動チューニング技術によって動的に制御する手法を確立, ppOpen-AT(ppOpen-HPCの自動チューニング機構)に実装

# 研究計画：3年計画（2018～2020年度）

- 1年目（FY.2018）

- Intel Xeon Phi(KNL)・NVIDIA P100向け最適化, 実装及び精度評価
- 1年目は電力測定も含めて「一通りやってみた」という感じ

- 2年目（FY.2019）

- 精度保証手法の確立

- 3年目（FY.2020）

- 自動チューニング手法確立
- ppOpen-HPC等へ実装・公開

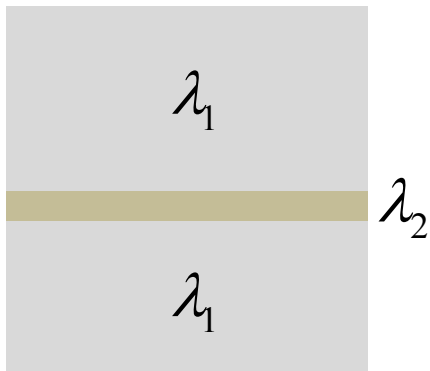
## 本研究の沿革(2/3)

- 2018年度は、上記①～⑩の項目について、(1)性能最適化、(2)多様な計算精度、(3)精度保証、(4)実アプリケーションを対象とした消費電力測定、を中心として予備的な検討を実施し、学会等で発表した他、ISC18、SC18、ISC19、SC19等の国際会議の展示ブースでも紹介した。
- 2018年度の段階で低精度・混合／変動精度演算をアルゴリズム、最適化、精度保証、消費電力まで含めて扱った研究事例はほとんどなかったが、SC19(2019年11月)では、低精度・混合／変動精度演算を数値計算に取り入れた研究事例が多数見られた。

# 三次元不均質ポアソン方程式 ICCG法 [KN 2018]

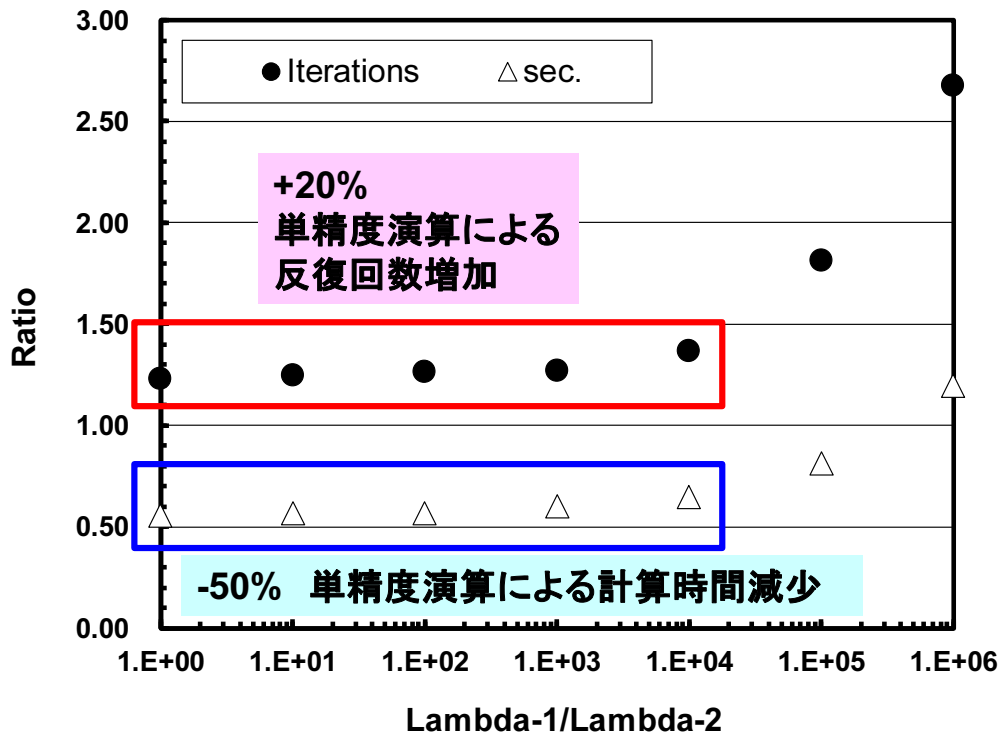
条件数( $\lambda_1/\lambda_2$ )と反復回数・計算時間の関係

単精度／倍精度の比: 少ないほど良い, Intel Xeon Broadwell-EP 1ノード



$$\nabla \cdot (\lambda \nabla \phi) + f = 0$$

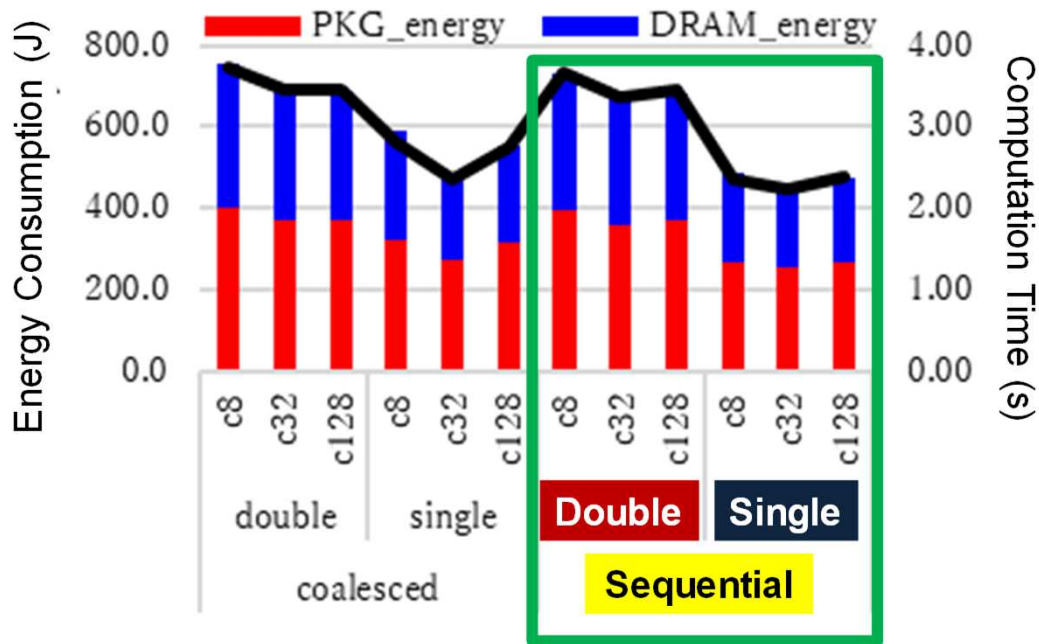
Reedbush:  
Intel Xeon Broadwell-EP  
Single Node:  
18 cores x 2 soc's



# 3D不均質ポアソン方程式 ICCG法:消費エネルギー測定(J)

$\lambda_1 = \lambda_2$ : Reedbush: Intel Xeon Broadwell-EP

- $128^3$  DOF
- Coalesced/Sequential
- 単精度・倍精度
- 色数: 8, 32, 128
- 計算時間に比例した消費エネルギー(J)
- Watt値は単精度になると増加する場合もあり
  - 計算密度増加

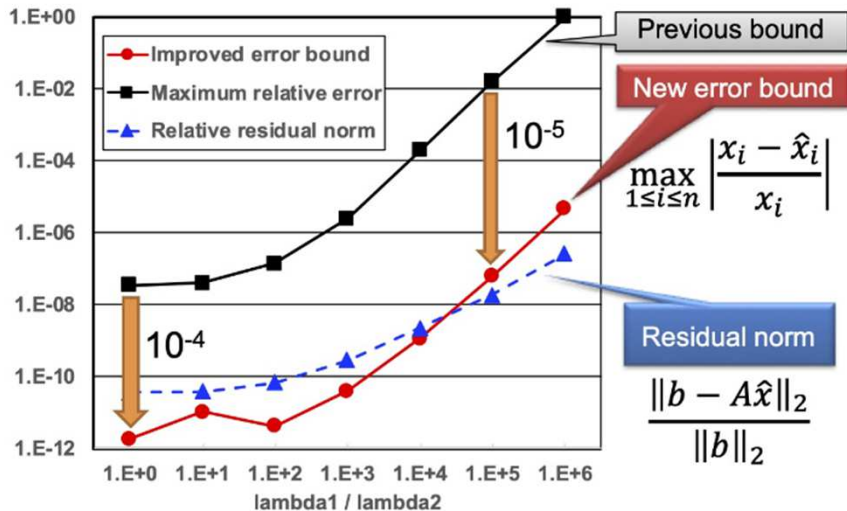




# 本研究の沿革(3/3):2019年度までに得られた知見

- 従来, 倍精度演算が適用されていた科学技術計算に単精度・半精度・混合演算を適用し, 反復改良法 (Iterative Refinement) 併用により, 同等の精度の計算結果がより短時間で得られる場合がある
  - 一般に, 計算時間短縮に比例して消費エネルギー (Joule) は減少する
- 悪条件問題では, 低精度演算では正解が得られない場合がある。
  - 特に半精度演算は変数の範囲が限定されるため注意が必要であり, 精度の要求されない反復法前処理等に適用するべきである。
- 従来のM疎行列向け精度保証手法 (T. Ogita他, 2001) は, 相対誤差上限の見積もりが厳しめであったため, より現実的な手法を開発し, 悪条件問題への有効性が示された。
- 演算精度の影響は, 問題規模・疎行列格納手法, アーキテクチャにより多様
- 局所的に演算精度を変更する手法の開発に着手し, 有効性が示された。
- 2019年度成果については「jh190042-NAH」へ

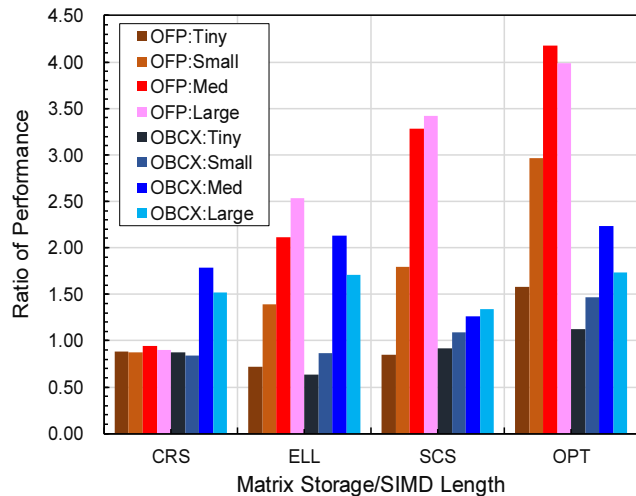
# 実問題向け精度保証手法の開発(荻田, 尾崎, 中島)



- 係数行列AがM行列性を持つ場合, 高速な精度保証法[Ogita et al., Computing, 2001]を適用可能
- 不均質場におけるポアソン方程式を解くICCG法への適用事例では, 条件数が大きくなると誤差限界が過大評価になっている可能性
- 2019年度は[Ogita et al., Computing, 2002]に基づく手法を開発し数値実験を実施
- 図に示すように, 提案方式によって誤差限界の過大評価が大幅に抑制でき, 近似解と精度保証の計算時間は同程度である

# 問題規模・行列格納手法と演算精度の関係

(中島, 坂本, 近藤他).



倍精度演算(FP64)と比較すると単精度演算(FP32)では計算時間が減少するが、アーキテクチャ、最適化の程度によってその影響は様々

- 2019年度は、ポアソン方程式を解くICCG法に対して、OFP(Intel Xeon Phi), OBCX(Intel CLX)1ノードにおいて様々な疎行列格納形式、問題規模に対して計算を実施した。
- 図は、CRS・倍精度の計算効率を1.00とした場合の単精度における計算効率

– CRS, ELL: Ellpack-Itpack, SCS: SELL-C- $\sigma$ , OPT: SCSを最適化

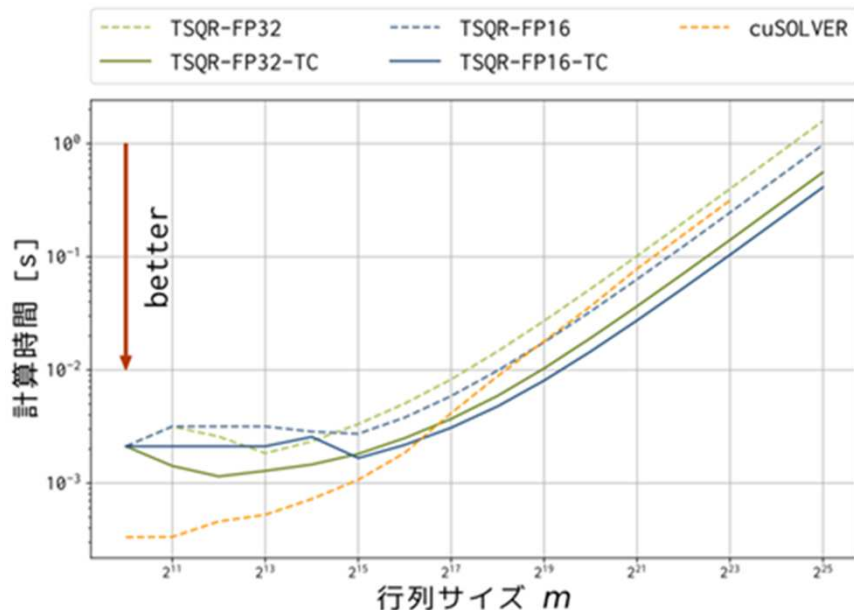
– OFPではベクトル化が不十分な場合には単精度による計算時間短縮効果は得られない。

– OFP, OBCX共に、問題規模、データ移動量が大では単精度による時間短縮効果が顕著

– 最適精度選択自動チューニング手法策定には、これらの項目を考慮する必要がある

# H行列向け手法の高速圧縮 手法の開発(横田)

- H行列計算の中で最も工夫を要するのは、密行列を低ランク行列に圧縮する部分である。
  - 圧縮には一般的にrandomized SVDが用いられ、その内部カーネルは複数のQR分解から構成される。
- 2018年度は多くの小さなQR分解をGPU上で並列に行うbatched QR分解を開発したが、2019年度はこれをTSQR(Tall & Skinny)に拡張し、任意の大きさの行列のrandomized SVDを行うためのQR分解の枠組みを開発した。
- NVIDIA V100上でFP32を用いた場合の計算時間はcuSOLVER による QR分解と比較して最大2.17倍高速となった



# 2020年度研究計画(1/4)

- 2019年度末時点で、ほぼ当初の予定通り目標を達成したが、2020年度は各アルゴリズムの、演算精度、最適化、アーキテクチャの他、問題規模も考慮して消費電力・エネルギーへの体系的な影響評価を継続して実施する他、低精度・混合／変動精度演算に関する研究開発を、これまでの研究成果を元に継続して実施し、自動チューニング手法確立を目指す。
  - 精度保証手法については、疎行列演算に加えて、H行列、偏微分方程式解法向け手法の研究開発も実施する。
  - 各センターの保有するNVIDIA V100, 東大情報基盤センターに2019年度末迄に導入される富士通FX700(A64FX(富岳))クラスタ等を使用して、新アーキテクチャ向け検討も実施する。
  - 特に富士通FX700クラスタを使用した、半精度演算(FP16)のフィージビリティースタディを様々な手法、アプリケーションについて重点的に実施する。

# 2020年度研究計画(2/4)

## • ステンシル計算(①, ②):

- 差分法コード, 地震波動伝播コードSeism3D(①), 全球大気計算コードNICAM(②)を対象として, 計算時間・消費エネルギー最小化のための演算精度選択手法の研究開発を継続して実施する。

## • 疎行列演算(③, ④, ⑤):

- 一般行列, 悪条件行列, Adaptive CGとともに, 反復改良法適用による低精度・混合／変動精度演算の安定化, 反復法前処理に対する低精度演算適用の検討, 様々な手法(Pipeline法, Dynamic Loop Scheduling, hCGA法, SELL-C- $\sigma$ ), 計算条件, 問題規模等の各演算精度による影響について, 安定性・実行時間の評価, 各CPU, GPUでの消費電力測定を継続して実施する。
- 計算時間・消費エネルギーを最小にするための演算精度選択手法の研究開発を継続して実施する(⑦, ⑨と関連)。
- 2019年度から新たに研究項目として追加した, 大規模分散並列問題で局所的・動的に演算精度を変動, 動的に負荷分散を適用する手法の研究開発を継続して実施する。

# 2020年度研究計画(3/4)

## • H行列(⑥):

- 低精度・混合／変動精度演算に関する検討を継続して実施し、悪条件問題向けの前処理手法の開発、電磁気学、地震発生サイクル等シミュレーションによる性能検証、精度評価、電力測定を実施する。
- また、悪条件問題を中心とした精度保証手法について検討を実施する(⑦, ⑨参照)。

## • 電力測定(⑧), 自動チューニング(⑨):

- ①～⑦の各項目と連携して、アルゴリズム、実装手法、問題規模、諸パラメータ、演算精度、アーキテクチャと消費電力、計算時間、計算結果の関係を体系化し、所望の計算精度を得るための最適パラメータ設定を自動的に決定する手法の検討を実施し、ppOpen-HPC, h3-Open-BDECへの実装を行う。

# 2020年度研究計画(4/4)

## • 精度保証(⑦):

- 疎行列演算(反復解法)については, 2019年度研究成果を元に, 悪条件問題に対して反復改良法と組み合わせた精度保証手法の研究開発を実施する。
  - ほぼ完成
- H行列については, 悪条件問題を中心とした精度保証の研究開発を実施する。
  - 階層的低ランク近似法により, 密行列における行列計算の演算量を低減することができる。
  - 低精度演算を用いる場合にはH行列のランクを小さくすることができ, 大幅に少ない演算量で同等の計算精度を実現できる。
  - H行列のランクが特異値の減衰率に依存することに着目し, 様々なアプリケーションで用いられる行列について特異値の観点から, 低精度演算と低ランク近似の適用範囲を明らかにし, 密行列直接解法における効率的な精度保証付き計算法を開発する。
- 三次元不均質場における楕円型偏微分方程式の解の精度保証付き数値計算を実現
  - 有限要素法等の離散化する際に生じる誤差も考慮し, 2019年度迄に開発した疎行列計算に対する精度保証法と組み合わせることにより, 対象問題の離散化誤差を含めた包括的誤差評価を高速に達成する。