# Post Peta CRESTで開発した *升行列ライブラリ升ACApK* (ライブラリと使用法について)

#### 伊田 明弘 (東京大学)

PCCCワークショップin大阪2020 2020年 2月 21日 富士通株式会社 関西システムラボラトリ(大阪)

- **1** *HACApK*ライブラリの概要
- 2 低ランク構造行列法
- 3 (格子) 升行列法の並列化
- 4 HACApKライブラリの使用方法
- 5 今後の取り組み

#### 概要:ppOpen-HPC

#### JST-CREST

- ・ "ポストペタスケール高性能計算に資するシステムソフトウェア技術の創出"領域
- "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"
   研究代表者: 中島 研吾(東京大学・情報基盤センター)

Download site: http://ppopenhpc.cc.u-tokyo.ac.jp/



# FDM, FEM, BEMで用いられる計算格子







Pictures are transcribed from

FDM: de Havilland Aircraft Company, VINAS Co. Ltd FEM: Prof. Nakazima (Tokyo Univ.) BEM: μ-TEC Co. Ltd

# FDM, FEM, BEM による離散化



### <u> 升ACApK ライブラリ</u>

- ■境界要素法(BEM)シミュレーション向け高速計算ライブラリ ■境界要素法のボトルネック
  - •g[u] = f, where  $g[u](x) := \int_{\Omega} \kappa(x, y) u(y) dy$

• Degenerate kernel:  $\kappa(x, y) \cong \sum_{\nu=1}^{r} \kappa_{1}^{\nu}(x) \kappa_{2}^{\nu}(y)$ , for far remote x, y

e.g. 
$$\kappa(x,y) \propto |x-y|^{-1}$$

#### ■大規模計算に向けて

- ► 密行列に対し近似手法の導入
   ・低ランク構造行列法: O(N<sup>2</sup>) ⇒ O(NlogN)

   光行列,格子光行列,BLR行列
- ▶ 分散メモリ並列計算
  - Hybrid MPI+X programming model

# *升*ACApK ライブラリ

■対応プラットフォーム

•CPU clusters •Multi- GPUs(整備中)

### ■ プログラミング言語

- Fortran90+OpenMP
- •CUDA, CUDA Fortran, OpenACC (GPU対応)
- ■必要な外部ライブラリ
  - •MPI
- 公開形態
  - Open source 
     MIT license

### ■ ダウンロード サイト

https://github.com/Post-Peta-Crest/ppOpenHPC/tree/MATH/HACApK

http://ppopenhpc.cc.u-tokyo.ac.jp(旧Version)

## <del>升ACApK</del>の適用例

#### Static electric field analysis

- Potential operator:  $\mathcal{V}[u](x) := \int_{\Gamma} \frac{1}{4\pi \|x y\|} u(y) dy, \quad x \in \Gamma$
- Induced surface charge is calculated in half-infinite domain.





Analysis result

### HACApKの適用例(CDWM)

■Calculate LLG equation with terms of spin torque effect

$$\frac{\partial M}{\partial t} = -\gamma M \times H - \frac{\alpha}{M_S} M \times \frac{\partial M}{\partial t} - (\mathbf{u} \cdot \nabla) M - \frac{\beta}{M_S} M \times [(\mathbf{u} \cdot \nabla) M]$$
  
**Effective field**  

$$H = \frac{1}{4\pi\mu_0} \int_V \nabla \cdot M \frac{\mathbf{r}}{r^3} dV + \frac{2K_u}{M_s^2} (M \cdot \mathbf{k}) \mathbf{k} + \frac{2A}{M_s^2} \Delta M + H_{app}$$
  
**Demagnetizing field**  $H^D$   
Other field components  $H^O$ 

#### Discretization method

- Time : Explicit Runge–Kutta  $M|_t$ ,  $H|_t \implies M|_{t+\Delta t}$
- Space : Magnetic material is divided into N cubic elements



### HACApKの適用例(CDWM)

Current-induced Domain Wall Motion (CDWM)CDWM is induced by a spin-polarized current in magnetic nanowires.



### <del>升ACApK</del>の適用例

# [Ida : *IEEE Trans. on Mag. 2017*] Micromagnetic simulation of spin torque oscillator



Oscillating frequency spectrum is calculated in magnetic materials.



•H-matrices accuracy of  $\varepsilon = 10^{-5}$  is required to reproduce oscillation processes.

<del>ℋACApK</del>の適用例



### <del>升ACApK</del>の適用例

#### [Mifune : Superconductor Science and Technology 2019]

#### Superconductor Simulation accelerated by AMG

-Governing equation: Faraday's law and Biot–Savart's

$$\nabla \times \left(\frac{1}{\sigma_{\rm s}[T]} \nabla \times \mathbf{n}T\right) \cdot \mathbf{n} + \frac{\partial}{\partial t} \left(\frac{\mu_0 t_{\rm s}}{4\pi} \int_{S'} \frac{(\nabla \times \mathbf{n}'T') \times \mathbf{r} \cdot \mathbf{n}}{r^3} \, \mathrm{d}S'\right) = 0$$

Galerkin and backward finite difference methods

$$A[\boldsymbol{T}_{[m-1]}] \cdot \boldsymbol{T}_{[m]} + B \cdot (\boldsymbol{T}_{[m]} - \boldsymbol{T}_{[m-1]}) = 0.$$



#### Previous study

- Apply  $\mathcal{H}$ ACApK to integral
- Solve by Newton-Raphson using BiCG-STAB

#### New approach

- Set *A* as preconditioning matrix in the BiCG-STAB
- Apply AMG to the preconditioner

HACApKの適用例



Time [year]

#### ℋACApKの適用例(地震発生物理学:動的破壊シミュレーション)

2016年熊本地震の例:

- 熊本側の本震と、大分側の動的誘発地震の発生を動力学的に再現
- 現実的な断層形状を考慮
- 動的誘発地震のメカニズムを解明: 本震が放射した強い応力波が50km以上を伝播



HACApKの適用例(地震発生物理学:動的破壊シミュレーション)

オリジナルBEM

積分核: K<sub>ijm</sub>

- 高速領域分割法(FDPM): Ando(2016, GJI) 積分核の領域ごとの時間依存性と積分領域形状の違いを利用
  - *in* Domain F →時空間依存だが高々面積分  $(K_{i j m})$ • 積分核:  $\begin{cases} K_{ij}^{IA} G_m + K_{ij}^{IB} & in \text{ Domain I} \\ K_{ij}^{S} & in \text{ Domain S} \end{cases}$  →時間項を分離した体積積分 →時間項のない面積分  $\Sigma_j \Sigma_m K_{ijm} D_{jn-m}$  $K_{i\,j}^{IA} G_m = \begin{bmatrix} k_1 \\ \vdots \\ k_j \end{bmatrix}_i \begin{bmatrix} g_1 & \cdots & g_m \end{bmatrix}$ 積分領域の分割 Space 効率化:密行列→ベクトルの積 Domain-F FDPMによる計算コストの削減 ~AS/CI+At •  $O(M^2N) \rightarrow O(M^2)$ メモリ Domain-S S-wave fro Domain-
    - $O(MN^3) \rightarrow O(MN^2) \sim O(M^2)$  演算数 (M: 要素数. N:時間ステップ数)

まだ減らしたい→H行列法適用

~AS/CT+At

三角錐:

動弾性の影響範囲

16

### HACApKと密行列のメモリ使用量

 $\blacksquare$   $\mathcal{H}$ -matrices with ACA reduce memory usage.



- **1** *HACApK*ライブラリの概要
- 2 低ランク構造行列法
- 3 (格子) 光行列法の並列化
- 4 HACApKライブラリの使用方法
- 5 今後の取り組み

#### 低ランク行列近似(LRA)とベクトルの直積

■ベクトルの直積は「ランク1の行列」

$$a \otimes b = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{pmatrix} \otimes (b_1 \quad b_2 \quad b_3) = \begin{pmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \\ a_4b_1 & a_4b_2 & a_4b_3 \\ a_5b_1 & a_5b_2 & a_5b_3 \\ a_6b_1 & a_6b_2 & a_6b_3 \end{pmatrix}$$

■任意の行列はベクトルの直積(ランク1行列)で近似できる(精度は度外視)

$$\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \\ c_{41} & c_{42} & c_{43} \\ c_{51} & c_{53} & c_{53} \\ c_{61} & c_{62} & c_{63} \end{pmatrix} \simeq \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{pmatrix} \otimes (b_1 \quad b_2 \quad b_3)$$

■基底の異なるランク1行列をk個足すと「ランクk行列」になる

■任意の行列はランクk行列で近似できる

・kが十分小さければ低ランク近似



低ランク構造行列法

■積分方程式法に現れる密行列に対する近似手法



### ACAを用いた低ランク行列近似

### Adaptive Cross Approximation



■選択するベクトルの本数により使用メモリ量と近似精度を制御可能.

### Procedure of ACA with full pivoting (ACA\_full)

Pivoting and updating matrix need computational cost  $O(N^2)$ .



### Procedure of ACA with partial pivoting (ACA\_part)



#### 低ランク構造行列法の近似原理

■ツリー法, FMM, パネルクラスタリング法などと近似原理は同じ.



### 低ランク構造行列の生成プロセス

### ■ HACApK では青字の方法を採用している

- **Step 1** Clustering & Construction of a cluster tree
  - Method based on geometric information
  - Method based on nested dissection
  - AMG-like method
- Step 2 Construction of a partition■ There're only blocks(frames of sub-matrices).
- Step 3 Fill in sub-matrices
  - Low-rank approximation
    - Cross approximation i) ACA ii) ACA+
    - Interpolation
    - -Hybrid method of interpolation and ACA







### $\mathcal{H}ACApKにおける行列分割の方法(従来<math>\mathcal{H}行列の場合)$

#### ■幾何情報に基づくクラスタリングを利用

Admissible condition: min{diam( $\Omega_s^h$ ), diam( $\Omega_t^h$ )}  $\leq \eta dist(\Omega_s^h, \Omega_t^h)$ 



#### 従来升行列法を用いた行列分割の例





### 従来光行列法を用いた行列分割の例



#### 低ランク構造行列法と近縁手法の種類

#### ■行列分割構造と要素保存法



・青字: HACApKで取り扱える手法

### 既存の低ランク構造行列法の比較

■分割構造の違いにより、一長一短



- Complicated structure
   効率的な演算・通信
   パターンの構築が困難
- -High compressibility

Memory : **O**(**NlogN**)

• Strong admissible cond.



- Simple structure like block-divided matrix 密行列用アルゴリズムを 流用可能
- Low compressibility
   Memory : *O*(*N*<sup>1.5</sup>)
- •Strong admissible cond.



- Relatively simple structure
- Very high compressibility for low dimensional problem
- Worse asymptotic complexity for high dimensional problem

Weak admissible cond.(All off-diagonal : Low-rank)

### 格子升行列法 [A.Ida : IPDPS'18]

#### ■低ランク構造行列法の新手法

- -Hybrid of  $\mathcal{H}$ -matrices and BLR-matrices
- -Advantages of lattice  $\mathcal{H}$ -matrices
  - High compressibility of  $\mathcal H\text{-matrices}$
  - Convenience of matrix arithmetic with BLR-matrices



- 格子升行列のメモリ使用量
- ■格子サイズの決め方に依存
  - *l*:格子サイズ
  - • $l = N \square$  従来 $\mathcal{H}$ 行列
  - *l* = 1 □ 溶行列



- $l = \alpha N$  for a constant  $\alpha (\frac{1}{N} \le \alpha \le 1)$ : 格子数固定 二〉要素数:  $f(N) = 2\gamma N \log N + N(2\gamma \log \alpha + 2/\alpha - \gamma)$
- ・格子サイズ l はMPIプロセス数に応じて決められるべき

### 低ランク構造行列の行列分割方法

•既存升行列









▪BLR行列

ב×ב

$\begin{array}{c} \beth_1^2 \\ \times \beth_1^2 \end{array}$	$\Sigma_1^2 \times \Sigma_2^2$	$\Box_1^2 \times \Box_3^2$	$\begin{array}{c} \beth_1^2 \\ \times  \beth_4^2 \end{array}$
$\Sigma_2^2 \times \Sigma_1^2$	$\square_2^2 \times \square_2^2$	$\Sigma_{2}^{2} \times \Sigma_{3}^{2}$	$\square_2^2 \times \square_4^2$
$\Sigma_3^2 \times \Sigma_1^2$	$\Sigma_3^2 \times \Sigma_2^2$	$\Sigma^2_3 \times \Sigma^2_3$	$\begin{array}{c} \Sigma_3^2 \  imes \Sigma_4^2 \end{array}$
$\square^2_4 \times \square^2_1$	$\Sigma_4^2 \times \Sigma_2^2$		$\begin{matrix} \beth_4^2 \\ \times  \beth_4^2 \end{matrix}$

$\exists_1^2 \\ \times \exists_1^2$	$\Sigma_1^2 \times \Sigma_2^2$	$\square_1^2$ × $\square_3^2$	$\square_1^2 \times \square_4^2$
$\square_2^2 \times \square_1^2$	$\Sigma_2^2 \times \Sigma_2^2$	$\begin{array}{c} \Sigma_2^2 \\ \times \ \Sigma_3^2 \end{array}$	$\square_2^2 \times \square_4^2$
$\square_3^2 \times \square_1^2$	$\square_3^2 \times \square_2^2$	$\begin{array}{c} \Sigma_3^2 \\ \times \ \Sigma_3^2 \end{array}$	$\begin{array}{c} \Sigma_3^2 \  imes \Sigma_4^2 \end{array}$
$\square_4^2 \times \square_1^2$	$\square_4^2 \times \square_2^2$	$\square_4^2 \times \square_3^2$	$\square_4^2 \times \square_4^2$





	$\begin{array}{c} \beth_1^2 \\ \times \beth_1^2 \\ \square_2^2 \\ \times \beth_1^2 \end{array}$	$\begin{array}{c} \square_1^2 \\ \times \square_2^2 \\ \square_2^2 \\ \times \square_2^2 \end{array}$	$\begin{array}{c} \beth_1^2 \\ \times \beth_3^2 \\ \square_2^2 \\ \times \beth_3^2 \end{array}$	$\begin{array}{c} \beth_1^2 \\ \times  \beth_4^2 \\ \square_2^2 \\ \times  \beth_4^2 \end{array}$	
ב×ב	$2^2_3 \times 2^2_1$	$\begin{array}{c} 2\\ 2_3^2\\ \times 2_2^2\end{array}$	$\Sigma_3^2$ $\times \Sigma_3^2$	$\square_3^2 \times \square_4^2$	
				$\begin{matrix} \beth_4^2 \\ \times  \beth_4^2 \end{matrix}$	

•	$\begin{array}{c} \beth_1^2 \\ \times \beth_1^2 \end{array}$	$\square_1^2 \times \square_2^2$	$\Box_1^2 \times \Box_3^2$	$\begin{array}{c} \beth_1^2 \\ \times  \beth_4^2 \end{array}$	
	$\begin{array}{c} \beth_2^2 \\ \times  \beth_1^2 \end{array}$	$2^2_2 \times 2^2_2$	$\begin{array}{c} \Sigma_2^2 \\ \times \ \Sigma_3^2 \end{array}$	$\begin{array}{c} \beth_2^2 \\ \times \square_4^2 \end{array}$	
		$\square_3^2 \times \square_2^2$	$\begin{array}{c} \Sigma_3^2 \\ \times \ \Sigma_3^2 \end{array}$	$\begin{matrix} \beth_3^2 \\ \times  \beth_4^2 \end{matrix}$	
	$\square_4^2 \times \square_1^2$	$\square_4^2 \times \square_2^2$	$egin{array}{c} \Sigma_4^2 \  imes \Sigma_3^2 \end{array}$	$\square_4^2 \times \square_4^2$	

	$\exists_1^2 \\ \times \exists_1^2$	$\Sigma_1^2 \times \Sigma_2^2$	$\square_1^2$ × $\square_3^2$	$\square_1^2 \times \square_4^2$
	$\square_2^2 \times \square_1^2$	$\square_2^2 \times \square_2^2$	$\Sigma_2^2 \times \Sigma_3^2$	$\square_2^2 \times \square_4^2$
	$\Sigma_3^2 \times \Sigma_1^2$	$\square_3^2 \times \square_2^2$	<sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>00</sup> <sup>01</sup> <sup>00</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup> <sup>01</sup>	$\begin{matrix} \beth_3^2 \\ \times  \beth_4^2 \end{matrix}$
	$\Sigma_4^2 \times \Sigma_1^2$	$\Sigma_4^2 \times \Sigma_2^2$	3     5       × 23     × 24       >23     × 24       >6     23       × 23     × 24       >6     × 24	$ \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \end{array} \\ \end{array} \\ \end{array} \\ \end{array} \end{array} $ $ \begin{array}{c} \begin{array}{c} \end{array} \\ \end{array} $ $ \begin{array}{c} \end{array} \\ $ $ \begin{array}{c} \end{array} \\ \end{array} $ $ \begin{array}{c} \end{array} \\ \end{array} $ $ \begin{array}{c} \end{array} \\ $ $ \begin{array}{c} \end{array} \\ $ $ \end{array} $ $ \end{array} $ $ \begin{array}{c} \end{array} \\ $ $ \end{array} $ $ \end{array} $ $ \begin{array}{c} \end{array} \\ $ $ \end{array} $

33

- **1** *HACApK*ライブラリの概要
- 2 低ランク構造行列法
- 3 (格子) 升行列法の並列化
- **4** *H*ACApKライブラリの使用方法
- 5 今後の取り組み

境界要素法解析で必要なHACApKの2機能

演算は部分行列ごとに実施

- 低ランク構造行列生成
  - ・低ランク近似(LRA)は各ブロックで個別独立に実施
- 低ランク構造行列-ベクトル積
  - ・部分行列-ベクトル積の集まり



#### HACApKにおける低ランク構造行列法の並列化

- ■低ランク構造行列の生成において
  - ・step 3 (最も時間がかかる部分)のみ並列化. (step1,2も並列化した版は公開準備中)
  - ・MPIの通信は不要.

step1 クラスタッリー 作成 step2 H行列構造 作成

全MPIプロセスで冗長計算

step3 部分行列の計算(ACA)

並列計算

- ■低ランク構造行列-ベクトル積計算において
- ・全MPIプロセスがベクトル全体を持つ(ベクトル分散版は公開準備中) ・MPIプロセス間の通信が必要.

### ■上記の両方の並列計算において

- ・計算コアへのデータ割り当て方式は共通にする.
- ・演算は部分行列を単位として行う.
- ・計算コアへ割り当てられたデータは部分行列の集合となる.

### HACApKにおける従来H行列の並列計算手法

基本的に光行列を行方向にスライスするが、部分行列は切断しない

■ MPIプロセス割当て戦略

① 各ブロックの高さR(M<sub>k</sub>)を可能な限り最小化
 ⇒ 通信データサイズの最小化

② MPIプロセス間の計算負荷不均衡を最小化



- 従来 $\mathcal{H}$ 行列-ベクトル積  $y \coloneqq \widetilde{A}x$ 
  - ・各MPIプロセスで部分行列-ベクトル積を行った後に隣接間通信



### 格子升行列に対するMPIプロセス割当て戦略

#### ■「プロセスグリッド」に基づく2次元サイクリック割当て

・各MPIプロセスへの割当て:同色で塗られた格子

-複雑な構造は逐次orスレッド並列では許容範囲 -密行列向けの既存アルゴリズムを流用可能

■ 例: 3×2 プロセスグリッド (6 MPI-processes)



Assignment global view



Lattice  $\mathcal{H}$ -matrix





		88

□ : Process(2,1)



格子 $\mathcal{H}$ 行列-ベクトル積,  $y \coloneqq \widetilde{A}x$ 

アルゴリズム 1 : "LatticeHMVM\_fullVector" ■各MPIプロセスで部分行列-ベクトル積を実行し部分ベクトルを得る ■全プロセスで全ベクトル y を得るために,

 Obtain a part of the result vector by MPI\_All\_reduce in each row in process grid.



② Exchange parts of vector by MPI\_iSEND and iRECV in each column.



### 格子 $\mathcal{H}$ 行列-ベクトル積, $y \coloneqq \widetilde{A}x(ベクトル分散版)$

アルゴリズム 2 : "LatticeHMVM \_partVector" ■正方形プロセスグリッド限定:  $N_{pr} = N_{pl} = \sqrt{N_p}$ •各MPIプロセスが全ベクトルを持つ必要がない

■次の Ãxを計算するため必要な部分ベクトルのみを通信

① MPI\_reduce in each row to diagonal processes

② MPI\_Broadcast in each column from diagonal processes to other processes





### 数值実験: Current induced Domain Wall Motion(CDWM)

#### ■ (格子) ℋ-行列の近似精度

• Tolerance  $\varepsilon = 1.0e-4$  for  $||A - \tilde{A}||_F \le \varepsilon$ .

#### ■計算メッシュ

Thin films of magnetic material



- A: Dense matrix,  $\tilde{A}$ :  $\mathcal{H}$ -matrix
- Single film model
   N =20,910
  - N = 31,860
  - *N* =116,100
- Double film model
   N =232,200
- Triple film model
   N =348,300
- ■計算機: Fujitsu PRIMERGY CX2550 (北海道大学)
  - Processor : Intel Xeon<sup>™</sup> Gold 6148 (40cores/node)
  - •Memory : 384GB/node, Network : 12.5 GB/s, Fat Tree.

格子光-行列の計算オーダー

- 格子サイズ l の決め方で  $O(N \log N) \sim O(N^2)$ 
  - ・ $l = N \square$  従来  $\mathcal{H}$ -行列 ・ $l = 1 \square$  密行列
- MPI数で格子サイズを決定  $(l = N/\sqrt{9N_p})$ :  $O(N \log N)$ •Memory usage of Lattice  $\mathcal{H}$ -matrices
  - -Calculation time of Lattice  $\mathcal{H}$ MVM when using 400 MPI processes



### MPI プロセス間の計算負荷均衡



・格子升行列:効率低下の度合いが穏やか



• " Single film model" N =20,910



#### (格子) ℋ行列-ベクトル積の並列計算性能

- 以下のMPIプロセス数まで計算速度向上を観測
  - ·従来升行列:約100 MPI プロセスまで
  - ・格子升行列(全ベクトル保持):約1,600 MPIプロセスまで
  - ・格子升行列(ベクトル分散): 少なくとも 3,600 MPIプロセスまで
- MPIプロセス数 $N_p$ > 100では格子 $\mathcal{H}$ 行列が従来 $\mathcal{H}$ 行列より高速 ■ MPIプロセス数 $N_p$ > 600では格子 $\mathcal{H}$ 行列(ベクトル分散)が最速



- **1** *HACApK*ライブラリの概要
- 2 低ランク構造行列法
- 3 (格子) 光行列法の並列化
- 4 HACApKライブラリの使用方法
- 5 今後の取り組み

### **HACApKで仕様を公開している関数**

■関数一覧と機能		赤字:必ず使用される関数	
	関数名	機能、コメント	
	HACApK_init	環境設定;最初に呼ばれる必要あり	
	HACApK_generate	低ランク構造行列の生成	
	HACApK_solve	線形方程式の求解	
	HACApK_adot_pmt_lfmtx_p	行列・ベクトル積計算	
	HACApK_free_leafmtxp	低ランク構造行列のメモリ開放	
	HACApK_finalize	後始末	

・境界要素法シミュレーションに最低限必要な機能を供給

- ■静的な問題を解く場合に使用される関数:HACApK\_solve 例:静電場解析
- ■動的な問題を解く場合に使用される関数:HACApK\_adot\_pmt\_lfmtx\_p 例:マイクロマグネティクス計算、超電導解析、地震サイクル解析、動的破壊解析

### HACApK利用コードの構成例



#### ユーザ定義関数 HACApK element ij

#### ■2つの要素番号と各要素上の情報を受け取り、行列要素値を返す

```
real*8 function HACApK_entry_ij(i,j,st_bemv)
integer :: i,j
type(st_HACApK_calc_entry) :: st_bemv
        :
    return HACApK_entry_ij
end function
```



#### ■境界要素積分に必要な情報を格納する構造体を宣言しておく

```
type(st_HACApK_calc_entry)
integer :: nd,lp61
real*8,pointer :: ao(:)

integer :: int_ex
real*8 :: dbl_ex
integer,pointer :: int_ex1(:),int_ex2(:,:)
real*8,pointer :: dbl_ex1(:),dbl_ex2(:,:)
end function
```



Pseudo code 2.4.1: example of the basic usage of the HACApK.

### HACApKへの入力情報

### ■行列サイズ:nd

lrtrn=HACApK\_init(nd,st\_ctl,st\_bemv)

allocate(coord(nd, 3), rhs(nd), sol(nd))

### ■ *升行列への*要求精度:ztol

lrtrn=HACApK\_generate(st\_leafmtxp,st\_bemv,st\_ctl,coord,ztol)
lrtrn=HACApK\_solve(st\_leafmtxp,st\_bemv,st\_ctl,rhs,sol,ztol)

#### ■線形方程式の右辺ベクトル: rhs

lrtrn=HACApK\_solve(st\_leafmtxp,st\_bemv,st\_ctl,rhs,sol,ztol)

#### ■行列ベクトル積の元ベクトル:x

lrtrn=HACAkK\_adot\_pmt\_lfmtx\_p(st\_leafmtxp,st\_bemv,st\_ctl,ax,x)

**HACApKへの入力情報**:座標情報coord(:,:)

■クラスタリングに必要な座標情報

・行列インデックスと関連付けられている三次元デカルト座標
 ・要素が関連付けられている場合には、要素重心の座標



#### HACApKのログ

#### ■ 制御パラメタst\_ctl%param(1)の値が1以上で出力



### **HACApKの一歩進んだ使い方**

#### ■制御パラメタ: st\_ctl%param(1:100)

番号	初期値	説明
1	1	ログ出力(0:エラーのみ、1:標準出力、2:デバグ用出力)
10	1	低ランク構造行列の種類(1:通常 <b>光</b> 行列, 10:格子 <b>光</b> 行列, 20:BLR行列)
11	0	近似行列の精度チェック(0:チェックしない, 1:チェックする)
21	15	最小クラスタサイズ
51	2.0	クラスタ間の距離判定パラメータ
61	1	部分行列での低ランク近似手法選択(1:ACA法, 2:ACA+法)
63	1000	部分行列の最大ランク

#### ■制御パラメタの設定方法

・初期化関数HACApK\_initと近似行列生成関数HACApK\_generateの間で設定

lrtrn=HACApK\_init(nd,st\_ctl,st\_bemv)
allocate(coord(nd,3),rhs(nd),sol(nd))
ztol=1.0e-5; st\_ctl%param(10)=10; st\_ctl%param(11)=1;
lrtrn=HACApK\_generate(st\_leafmtxp,st\_bemv,st\_ctl,coord,ztol)

#### 制御パラメタ: st\_ctl%paramの解説



- **1** *HACApK*ライブラリの概要
- 2 低ランク構造行列法
- 3 (格子) 光行列法の並列化
- 4 HACApKライブラリの使用方法
- 5 今後の取り組み

#### 低ランク構造行列分解への取り組み

- ■取り組み中の行列分解
  - ►LU分解 ►QR分解
- ■想定する利点
  - ▶密行列計算より軽い
    - ・メモリ使用量: $O(NlogN) < O(N^2)$ · 注答号 : $O(Nlog^2 N) \ll O(N^3)$
    - •演算量 :  $O(N\log^2 N) \ll O(N^3)$
  - ▶疎行列の不完全分解より高精度
  - ▶ ランク数の調整により精度が制御可能?

#### ■想定する使い道

- ▶低精度:クリロフ部分空間法の前処理、固有値の近似値計算
- ▶高精度:密行列による完全分解の代替

#### 格子 光行列に基づく数値線形代数の構築

- ■データ科学計算および科学技術計算の大規模化・高速化
  - ► $O(N^2)$ 以上の密行列演算を $O(N\log^p N)$ 以下に低減
  - ►密行列線形代数ライブラリ(BLAS, LAPACK, ScaLAPACK)の置き換え
- ■格子升行列計算アルゴリズム群の開発:
  - ▶行列演算種類の拡充
    - ・開発済み:格子H行列生成,行列-ベクトル積,LU分解
    - 開発中:QR 分解(BLRでは開発済)
    - •開発予定:行列-行列積,固有值計算,特異值計算
  - ►格子光行列法の深化
    - ・複素数化および長方行列への対応(現状は実数正方行列のみ対応可能)
       ・テンソルへの拡張
  - ▶データ科学分野の問題への適用
    - ・最尤推測法に用いられる各種確率分布に基づく共分散行列への適用
    - ・3次元ユークリッド空間と高次元データ空間の違いを克服
    - ・ベイズ推論および深層学習へと適用範囲拡大

### 科学技術計算での密行列演算回避



#### 階層的低ランク行列近似の適用範囲

 1 線形積分作用素から得られる密行列の効率的な近似行列表現を得る
 2 上記①で得られた階層型行列に対し、 行列計算(加法, 乗法, LU分解, QR分解など)を行う
 3 線形微分作用素から得られる疎行列に対し、近似逆行列を計算する
 4 線形微分作用素から得られる疎行列に対し、近似LU分解を計算する

③の直感的理解(Bebendorf(2008)より)・微分の逆作用素は積分のはず

$$-\Delta u = f \text{ in } \Omega \coloneqq (a, b), u(a) = u(b) = 0$$

$$\Box \quad 2 x = n = n$$

$$A = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & -1 & 2 & -1 & & \\ & & & -1 & 2 & -1 & \\ & & & & -1 & 2 & -1 \\ & & & & & -1 & 2 \end{bmatrix}$$

線形積分作用素の 離散化行列に類似 ⇒階層型行列

$$A^{-1} = \frac{1}{9} \begin{bmatrix} 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 7 & 14 & 12 & 10 & 8 & 6 & 4 & 2 \\ 6 & 12 & 18 & 15 & 12 & 9 & 6 & 3 \\ 5 & 10 & 15 & 20 & 16 & 12 & 8 & 4 \\ 4 & 8 & 12 & 16 & 20 & 15 & 10 & 5 \\ 3 & 6 & 9 & 12 & 15 & 18 & 12 & 6 \\ 2 & 4 & 6 & 8 & 10 & 12 & 14 & 7 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix}$$

#### データ科学計算への格子光行列適用



メモリ:  $O(N \log N)$ 

- Dimensionality Reduction
  - ▶主成分分析のホットスポット:共分散行列の固有値・固有ベクトル計算
    - ・密行列 :メモリ $O(N^2)$ ,演算量 $O(N^3)$ ・格子 $\mathcal{H}$ 行列:メモリ $O(N\log N)$ ,演算量 $O(N\log^2 N)$

■ディープニューラルネットワークのハイパーパラメータ予測

- ►ヘッセ行列とヤコビ行列の関係式から最適な学習率やバッチサイズを見積り
- ・McCandlishらの研究では対角近似 ⇒ 格子升行列適用により予測精度向上

#### 格子升行列の最新アーキテクチャへの高性能実装法

■ハードウェアの性能を引き出せるアルゴリズム開発

■最新計算機アーキテクチャ向け格子H行列計算コード開発:

- ►次世代CPU 向けコード開発
  - ・スパコン「京」の次世代機「富岳」で採用されるCPU「A64FX」
  - ・メニーコア型CPU,ベクトル命令セット採用,小容量高速メモリ搭載
- ►GPU向けコード開発
  - ・マルチコアCPU 向けが既開発の演算(格子H 行列生成, 行列-ベクトル積, LU 分解) ・テンソルコアの効率的な利用法
- ► FPGA を用いた格子 *H*行列分割生成および行列分解
  - ・格子ℋ行列分割構造の作成ルーチン(条件分岐が多く、CPU・GPUに不向き)
  - ・格子ℋ行列分解(深い演算パイプラインと細粒度での並列処理が可能なFPGA向き)
- ▶混合精度演算による計算の効率化
  - ・部分行列のランク数に応じて数値の保存精度を決定
  - ・誤差が蓄積する縮約演算の量に応じて演算精度を決定
- ►動的手法を用いた高速化(ヘテロジニアスなシステムへの対応)
  - ・タスク並列言語(Cilk Plus, Tascell)を用いた動的負荷分散
  - ・ランタイムライブラリ(StarPU, PaRSECなど)の活用

### Acceleration of $\mathcal{H}$ -mat. generation using KNL

#### Test model of electrostatic field analysis

- Perfect conducting sphere  $P[u](x) := \int_{\Omega} \frac{1}{4\pi ||x-y||} u(y) dy, x \in \Omega$
- ► Dielectric sphere
- $D[u](x) := \int_{\Omega} \frac{\langle x y, n(y) \rangle}{4\pi \|x y\|^3} u(y) dy, x \in \Omega$ including branch divergence



Ground

[Hoshino : CLUSTER 2018]

User-defined functions depend on these integral equations

Perfect

**BDW** 

Perfect

KNL

SIMD design



►KNL : Intel Xeon Phi 7250, 68 core

**Evaluation Environments** 

- ► Compiler : Intel compiler 18.0.1
  - -gopenmp -O3 -ipo -align array64byte -xAVX2 (BDW) -xMIC-AVX512 (KNL)

#### Performance comparison

- ► BDW : approximately **2x** speedup
- ►KNL : over **4x** speedup
  - In the case of dense matrix generation, new design achieved at most 6.6x speedup



Dierectric

BDW

Original

#### **Coefficient H-matrix generation** on BDW and KNL

Dierectric

KNL

### Acceleration of $\mathcal{H}$ -mat.-vec. using GPU, KNL

[Hoshino : ICCS 2018]



### HACApK関連の査読付き論文リスト

- 1. Rise Ooi, Takeshi Fukaya, Takeshi Iwashita, <u>Akihiro Ida</u> and Rio Yokota, "Accelerating Multithreaded Linear Solver with Mixed Precision Hierarchical Matrix Computation and Data Structure", International Conference on High Performance Computing in Asia-Pacific Redion (HPC Asia 2020), pp.92-101.
- Yu Pei, George Bosilca, Ichitaro Yamazaki, <u>Akihiro Ida</u> and Jack Dongarra, "Evaluation of Programming Models to Address Load Imbalance on Distributed Multi-Core CPUs: A Case Study with Block Low-Rank Factorization", In: 2019 IEEE/ACM Parallel Applications Workshop, Alternatives To MPI (PAW-ATM). IEEE, 2019. (In press)
- 3. <u>Akihiro Ida</u>, Tadashi Ataka and Atsushi Furuya, "Lattice H-matrices for Massively Parallel Micromagnetic Simulations of Current-induced Domain Wall Motion", IEEE Transactions on Magnetics, (In press)
- Zhengyang Bai, Tasuku Hiraishi, Hiroshi Nakashima, <u>Akihiro Ida</u> and Masahiro Yasugi. "Parallelization of Matrix Partitioning in Construction of Hierarchical Matrices using Task Parallel Languages", Journal of Information Processing, Vol. 27 (2019) pp. 840-851.
- 5. <u>Akihiro Ida</u>, Hiroshi Nakashima, Tasuku Hiraishi, Ichitaro Yamazaki, Rio Yokota and Takeshi Iwashita, "QR Factorization of Block Low-Rank Matrices with Weak Admissibility Condition", Journal of Information Processing, Vol. 27 (2019) pp. 831-839.
- 6. Satoshi Ohshima, Ichitaro Yamazaki, <u>Akihiro Ida</u> and Rio Yokota, "Optimization of Numerous Small Dense-Matrix–Vector Multiplications in H-matrix Arithmetic on GPU", In: Auto-Tuning for Multicore and GPU (ATMG) In conjunction with the IEEE MCSoC-19, Singapore (Oct 1-4, 2019)
- 7. Ichitaro Yamazaki, <u>Akihiro Ida</u>, Rio Yokota and Jack Dongarra, "Distributed Memory Lattice H-matrix Factorization", The International Journal of High Performance Computing Applications (IHPCA) (2019) 1094342019861139.
- 8. Takeshi Mifune, Naoki Tominaga, Yusuke Sogabe, Yudai Mizobata, Masahiro Yasunaga, Akihiro Ida, Takeshi Iwashita and Naoyuki Amemiya, "Large-scale electromagnetic field analyses of coils wound with coated conductors using a current-vector-potential formula-tion with a thin-strip approximation", Superconductor Science and Technology(SUST).
- 9. Tetsuya Hoshino, Akihiro Ida, Toshihiro Hanawa and Kengo Nakajima, "Load-balancing-aware Parallel Algorithms of H-matrices with Adaptive Cross Approximation for GPUs", In Proceedings of IEEE International Conference on Cluster Computing (CLUSTER 2018)

### HACApK関連の査読付き論文リスト

- 10. Tetsuya Hoshino, Akihiro Ida, Toshihiro Hanawa and Kengo Nakajima,"Design of Paral-lel BEM Analyses Framework for SIMD Processors", In: Shi Y. et al. (eds) Computational Science ? ICCS 2018. ICCS 2018. Lecture Notes in Computer Science, vol 10860. Spring-er, Cham
- 11. Akihiro Ida, "Lattice H-Matrices on Distributed-Memory Systems", 32nd IEEE Interna-tional Parallel & Distributed Processing Symposium (IPDPS 2018), Vancouver, Canada (May 21-25, 2018), pp.389-898.
- Ichitaro Yamazaki, Ahmad Abdelfattah, Akihiro Ida, Satoshi Ohshima, Stanimire Tomov, Rio Yokota, and Jack Dongarra "Analyzing Performance of BiCGStab with Hierarchical Matrix on GPU clusters", 32nd IEEE International Parallel & Distributed Processing Sym-posium (IPDPS 2018), Vancouver, Canada (May 21-25, 2018), pp930-939.
- 13. Naoki Tominaga, Takeshi Mifune, Akihiro Ida, Yusuke Sogabe, Takeshi Iwashita, Naoyuki Amemiya, "Application of hierarchical matrices to large-scale electromagnetic field analyses of coils wound with coated conductors", IEEE Transactions on Applied Su-perconductivity, Volume 28, Issue 3 (April 2018), pp.1-5.
- 14. Satoshi Ohshima, Ichitaro Yamazaki, <u>Akihiro Ida</u> and Rio Yokota, "Optimization of Hierarchical matrix computation on GPU", In: Asian Conference on Supercomputing Frontiers. Springer, Cham, 2018. p. 274-292.
- 15. <u>Akihiro Ida</u>, Hiroshi Nakashima and Masatoshi Kawai, "Parallel Hierarchical Matrices with Block Low-rank Representation on Distributed Memory Computer Systems", International Conference on High Performance Computing in Asia-Pacific Redion (HPC Asia 2018), Tokyo, Japan (Jan. 28-31, 2018), pp.232-240.
- 16. <u>Akihiro Ida</u>, Tadashi Ataka, Takeshi Mifune, Yasuhito Takahashi, Takeshi Iwashita and Tadashi Furuya, "Application of Improved H-matrices in Micromagnetic Simulations of Spin Torque Oscillator", IEEE Transactions on Magnetics, Vol. 54, Issue 3
- 17. Takeshi Iwashita, <u>Akihiro Ida</u>, Takeshi Mifune and Yasuhito Takahashi, "Software Framework for Parallel BEM Analyses with H-matrices Using MPI and OpenMP", Procedia Computer Science 108C (2017) pp.2200–2209.
- <u>Akihiro Ida</u>, Takeshi Iwashita, Takeshi Mifune and Yasuhito Takahashi, "Variable Preconditioning of Krylov Subspace Methods for Hierarchical Matrices with Adaptive Cross Approximation," IEEE Transactions on Magnetics, Vol. 52, Issue 3

### HACApK関連の査読付き論文リスト

- 19. Takeshi Iwashita, Naokazu Takemura, <u>Akihiro Ida</u> and Hiroshi Nakashima, "A new fill-in strategy for IC factorization preconditioning considering SIMD instructions", The 13-th IEEE International Symposium on Parallel and Distributed Processing with Applications (IEEE ISPA-15), Helsinki Finland (2015)
- 20. <u>Akihiro Ida</u>, Takeshi Iwashita, Makiko Ohtani and Kazuro Hirahara, "Improvement of Hierarchical Matrices with Adaptive Cross Approximation for Large-scale Simulation", Journal of Information Processing, Vol. 23, No. 3 (2015) pp. 366-372.
- 21. <u>Akihiro Ida</u>, Takeshi Iwashita, Takeshi Mifune and Yasuhito Takahashi, "Parallel Hierarchical Matrices with Adaptive Cross Approximation on Symmetric Multiprocessing Clusters", Journal of Information Processing, Vol. 22, No. 4 (2014) pp. 642-650.