大規模分子軌道計算における並列部分対角化 プロジェクト名: FMO-MO 法による大規模分子軌道計算

梅田宏明²,渡邊寿雄⁵,池上努⁵

1 筑波大学大学院 システム情報工学研究科
 2 筑波大学 計算科学研究センター
 3 産業技術総合研究所 計算科学研究部門
 4 東京工業大学 学術国際情報センター
 5 産業技術総合研究所 情報技術研究部門

2009年5月14日

はじめに

- 大規模なタンパク質の分子軌道計算
 - FMO-MO 法による大規模分子の Fock 行列生成(高い並列性)
 より大規模分子を扱うための計算量とメモリー使用量の削減
 - フロンティア軌道付近の固有ベクトル計算(部分対角化)
 超並列計算環境向き対角化法の開発
- T2K 環境への移行
 - T2K 以前の計算環境
 - クラスタ:AIST Super Cluster P32 subsystem
 - 1ノード: two mono-Opteron 2.0GHz

大規模化に向けての課題

計算対象例:上皮成長因子受容体 (EGFR) 二量体

17,246 原子, 6-31G 基底関数, 96,234 基底

行列は約 10 万次元,非零要素数は約 5 億. semi-sparse で要素配置が広範囲に広がった行列.

非零要素数は 1000 万次元の QCD 行列の場合と同程度. 全要素数の 5%程度で密行列として扱うほどではない.



Fock matrix

T2K での並列部分対角化のための課題:

- **①** パラメータ依存性が高く、ノード間のロードバランスが悪い
- 線形方程式が解きにくく前処理のコストやメモリー使用量が大きい
- 政行列処理において計算効率が悪い

対象とする並列固有値解法

周回積分を用いた固有値解法

複素平面上の周回積分を用いて目的とする部分空間を抽出する.

一般化固有値問題 $Ax = \lambda Bx$ の固有対 $: (\lambda_i, x_i)$, $1 \leq i \leq n$

m 個の固有値 $\lambda_1, \ldots, \lambda_m$ が Jordan 曲線 Γ 内にあるとする.



零でないベクトル
$$v\in\mathbb{C}^n$$
に対して $s_k=rac{1}{2\pi\mathrm{i}}\int_{\Gamma}z^k(zB-A)^{-1}Bv\mathrm{d}z$ とする、このとき、 $\mathrm{span}\{s_0,\ldots,s_{m-1}\}=\mathrm{span}\{x_1,\ldots,x_m\}$

T. Sakurai and H. Tadano, Hokkaido Math. J. 36 (2007).

ブロック化による安定化

1本の入力ベクトル v から 1本の出力を得る. Single Input Single Output (SISO)

$$s_k = rac{1}{2\pi\mathrm{i}}\int_{\Gamma} z^k (zB-A)^{-1}Bv\mathrm{d}z$$

線形独立な L 本の入力ベクトル v_1, \ldots, v_L を用いる. Multiple Input Multiple Output (MIMO)

$$s_k^{(l)} = rac{1}{2\pi\mathrm{i}}\int_{\Gamma} z^k (zB-A)^{-1}Bv_l\mathrm{d} z, \hspace{1em} l=1,2,\ldots,L$$

低次のモーメントベクトルだけを使うことで方法の安定性を向上させる.

Block Krylov 部分空間

L本の右辺ベクトルに対して線形方程式を解く:

$$(zB-A)y_l = Bv_l, \quad 1 \le l \le L$$

Krylov 部分空間 $(r_0 \in \mathbb{C}^n)$: $r_0, Ar_0, A^2r_0, A^3r_0, A^4r_0, \cdots$



Block Krylov 部分空間 $(R_0 \in \mathbb{C}^{n \times L})$: $R_0, AR_0, A^2R_0, \cdots$



Block COCG 法: Block Krylov 部分空間において解を探索する複素対称行列向き解法

ブロック化による効果

Block SS 法:

固有値解法部分において,複数の低次モーメントベクトルを用いて必要な部 分空間を生成

- 低次のモーメントベクトルを用いることで安定性が向上
- 精度が向上し、より広範囲の固有値を一度に求めることが可能
- Block COCG 法:

線形方程式の解法部分において,複数 seed によって Krylov 部分空間を生成

- •1反復あたりの探索空間を広げることで反復回数を低減
- 問題ごとの反復回数のばらつきを抑えてロードバランスを改善
- ●より軽い前処理で収束するようにし、計算時間とメモリー使用量を削減
- 複数右辺ベクトルで計算することで行列データの再利用性が向上
- 対称行列の行列ベクトル積や前処理における前進後退代入を複数右辺ベクトル で並列化

数値実験

Epidarmal Growth Factor Receptor (EGFR) dimmer

- Test problem:
 - EGFR (Epidermal Growth Factor Receptor) dimmer (17,246 Atoms)
 - Basis function: 6-31G (96,234 basis functions)
 - Matrix size: 96,234 × 96,234
 - Number of nonzero elements : 456,807,648
- Test environment:
 - AIST Super Cluster
 - T2K-Tsukuba system



AIST Super Cluster の結果

HOMO-LUMO 付近に 8 区間を設定し, 各区間ごとに計算. 1 区間あたり 16 ノード (32CPU) を割り当て. 94 組の固有対が得られた.

Wall-clock time with 128 nodes (256CPU)



(各ノードは two mono-Opteron 2.0GHz, 6 GB of memory)

T2K-Tsukuba システムの結果

T2K-Tsukuba 上で MIMO による高速化と Block COCG 法によるメモリー削 減の効果の検証.

- 1 領域あたり 4 ノード (16 ソケット)を割り当てた
- 1 ソケットあたり 1 つの方程式を解いた
- ソケット内 (4 コア) は OpenMP による並列化
- 前処理行列の生成は Intel Math Kernel Library の PARDISO を用いた

計算時間の改善

最初の BCAST を除いた計算時間の比較

AIST Super Cluster P-subsystem (Opteron Dual CPU 2.0GHz) 128 ノード (256CPU) で約 580 秒



T2K-Tsukuba system 32 ノード (512 コア) で約 235 秒

反復回数の変化

ブロック数 L を変えたときの Block COCG 法の残差履歴の比較



-: L = 1, -: L = 2, -: L = 4, -: L = 8, -: L = 16

反復解法の計算時間

前処理生成のパラメータによる Block COCG 法の計算時間 (秒)



計算時間の比較

前処理パラメータ $\delta=5.0 imes10^{-4}$ のときの Block COCG 法の計算時間の比較 (メモリーが十分にあるとき)



(FMO-MO 法による大規模分子軌道計算)

大規模分子軌道計算における並列部分対角化

計算時間の比較

前処理パラメータ $\delta=1.6 imes10^{-3}$ のときの Block COCG 法の計算時間の比較 (メモリー使用量を節約したとき)



ブロック化による精度の向上

ブロック数 L を変えたときの Block SS 法の固有対の精度



おわりに

マルチコアでの効率改善:

- MIMO によるデータ再利用性の向上
- Block COCG 法の提案
- 前処理におけるパラメータ依存性とメモリー使用量の低減
- Block SS 法による安定性の向上

今後の課題:

- マルチコアでの効率改善ノウハウの蓄積
 HPC 分野との連携
- MIMO による効率改善の応用
 - アプリケーション分野との連携
- 大規模分子の FMO 向けのモデリング