

# GPUでの交換相互作用

平面波基底第一原理計算プログラム

xTAPPへの実装

鳥取大学 工学部 吉本 芳英

# 交換相互作用

- 電子間相互作用の近似の順位
  1. 電子密度間のクーロン相互作用
  2. 1次の簡約化密度行列による交換相互作用
  3. 電子相関
- 局所密度近似では2.以降が近似

# 密度汎関数近似の発展

- 局所密度近似
  - 経済的であるが、絶対精度は万能でない
- 拡張：交換相互作用の計算を含む近似手法  
「hybrid型」
  - PBE0, B3LYP, HSE, LC など
- 交換相互作用の計算コストは大きい

# 交換相互作用の計算

$$E_x = -\frac{1}{2} \sum_{i,j,\sigma} \int d\mathbf{r} d\mathbf{r}' \frac{\psi_{i,\sigma}^*(\mathbf{r}) \psi_{j,\sigma}^*(\mathbf{r}') \psi_{j,\sigma}(\mathbf{r}) \psi_{i,\sigma}(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|}$$

$$V_x[\psi_{i,\sigma}](\mathbf{r}) = - \sum_j \int d\mathbf{r}' \frac{\psi_{j,\sigma}^*(\mathbf{r}') \psi_{i,\sigma}(\mathbf{r}') \psi_{j,\sigma}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|}$$

静電ポテンシャルの形

$$E_x = \frac{1}{2} \sum_{i,\sigma} \int d\mathbf{r} \psi_{i,\sigma}^*(\mathbf{r}) V_x[\psi_{i,\sigma}](\mathbf{r})$$

# 交換相互作用の計算

- $i, j$ のペア
  - $\psi_{j,\sigma}^*(\mathbf{r})\psi_{i,\sigma}(\mathbf{r})$  から「ポテンシャル」
  - FFTを使う： $O(N_A \log N_A)$
- $i$ 一つにつき、「ポテンシャル」は $N$ 個
- $i$ は $N$ 個

全部で $O(N_A^3 \log N_A)$

計算量が大きい

# GPUの活用

- CPUとの比較
  - 演算能力 大
  - メモリバンド幅 大
  - メモリ量 小
- メモリバンド幅：FFTを高速計算
- デバイスへの転送が必要

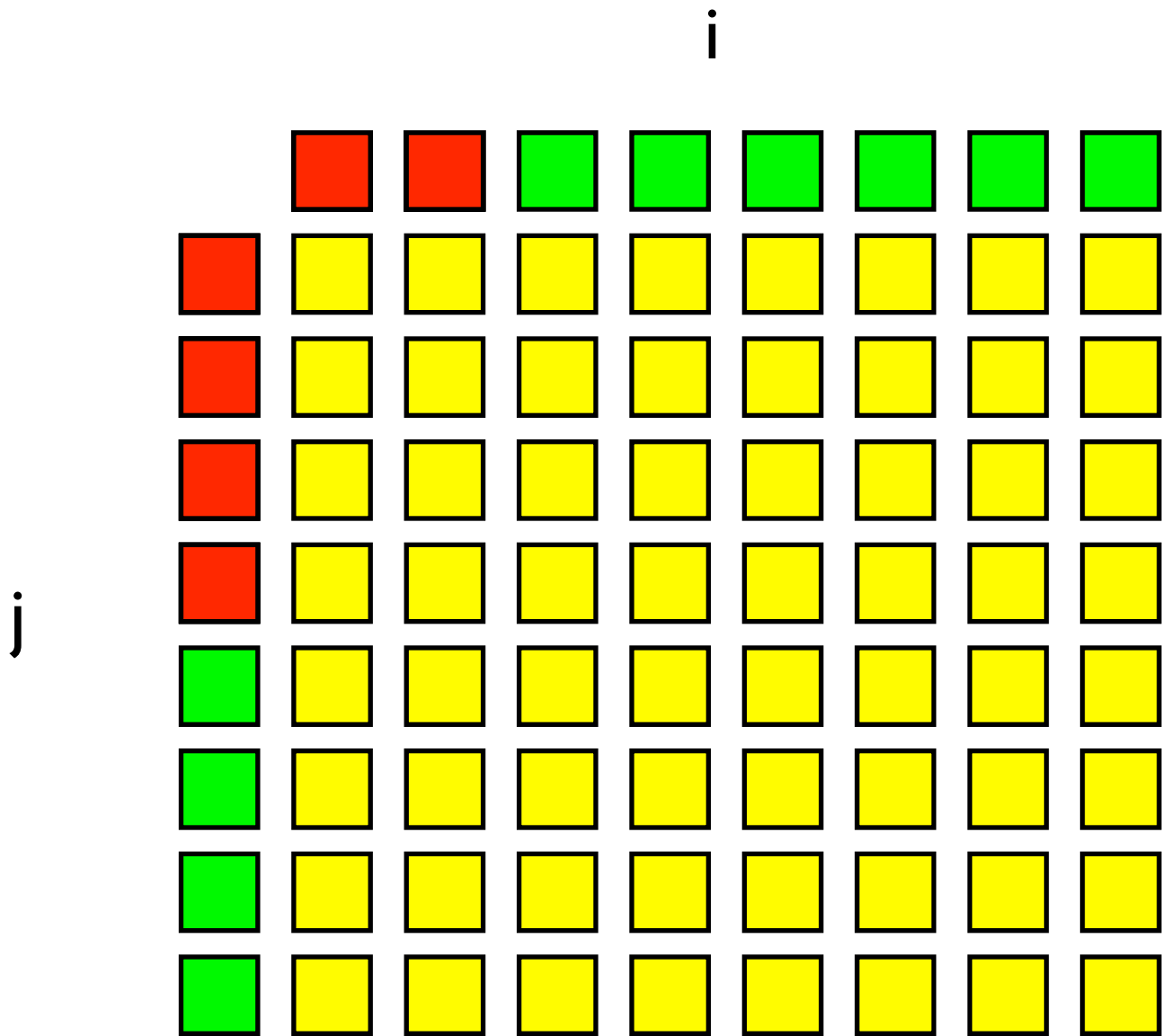
# 問題点

- CPU・GPU間のデータ転送はかなり遅い
  - ~ 5 GB/s、 行き帰りが必要
  - CPUのメモリバンド幅は>20 GB/s
- 一回のFFTだと転送時間の元が取れない
  - Xeon X5690 で $128^3$ のFFTに8.5 ms
  - 5 GB/sで $128^3$ の転送1回に6.7 ms

# 計算のブロック化

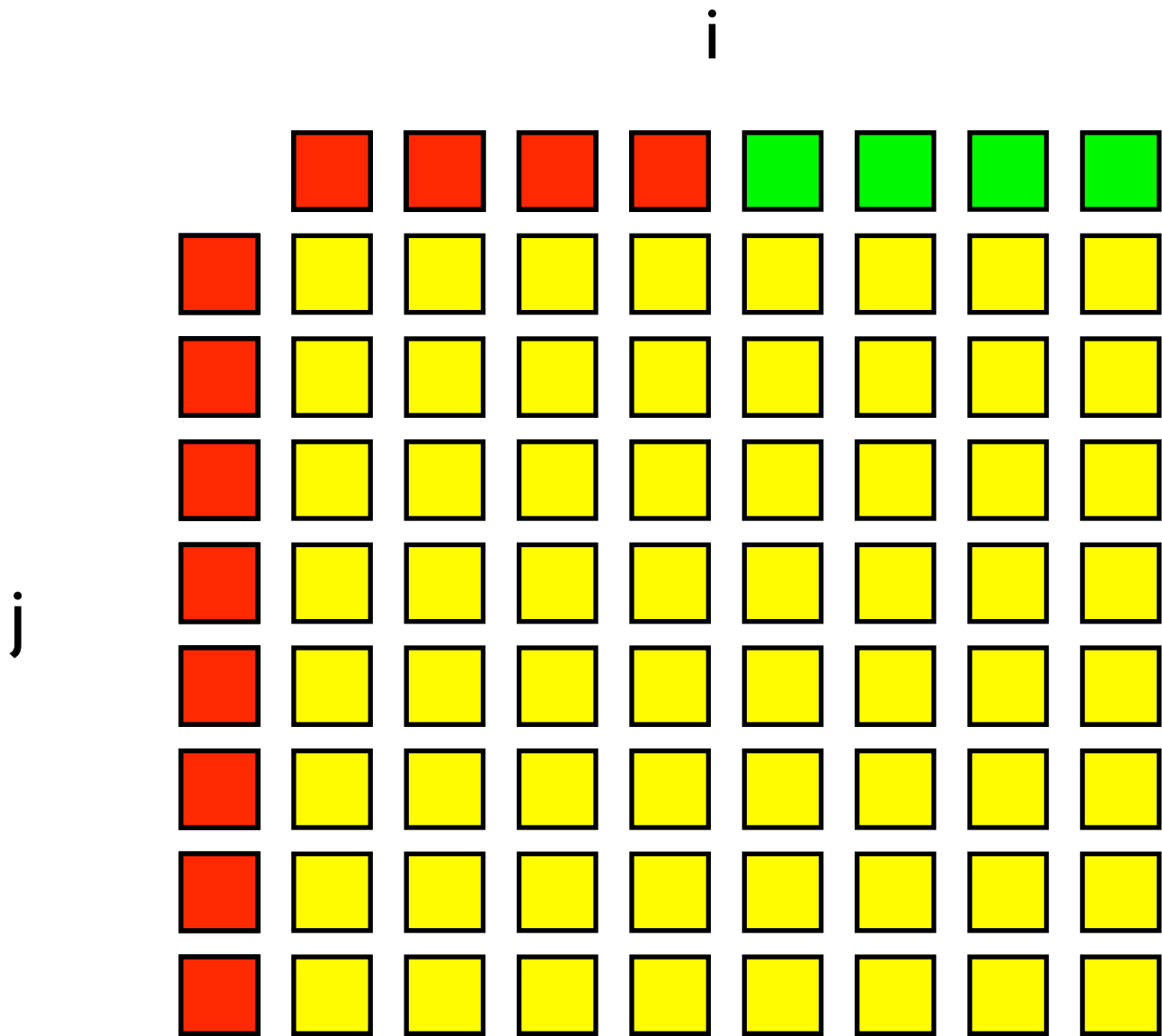
- $i, j$ のペアについて計算が必要
  - 入力データは $N$ 個、演算は $N^2$ 個、結果データは $N$ 個
- $N_{\text{blk}}$ 個  $\times$   $N_{\text{blk}}$ 個ずつブロック化して計算





演算1  
 転送2  
 比：1/2

$i, j$  のペアを一つずつ順番に処理



演算4  
 転送4  
 比：1

i, j のペアを4つずつグループごとで処理

# 演算 / データ転送

- $N_{\text{blk}} \times N_{\text{blk}}$  でブロック化
- 一ブロックあたり
  - 演算量は  $N_{\text{blk}}^2$  倍
  - データ転送量は  $N_{\text{blk}}$  倍
  - 演算 / データ転送は  $N_{\text{blk}}$  倍
- データ転送を目立たなくできる

# 使用GPU

- AMD Radeon HD 6950 2GB
  - 倍精度浮点演算512 GFLOPS
  - × CUDA → OpenCL
  - メモリ帯域幅144GB/s
  - ~2.6 万円
- NVIDIA Tesla C2000シリーズと同等の性能でずっと安い
  - Tesla C2070: 515 GFLOPS、144GB/s、~21 万円
  - Xeon X5690: 83 GFLOPS、32GB/s、~14 万円

# ベンチマーク条件

- CPU: Xeon X5690 (6core, 3.46GHz)
- GPU: AMD Radeon HD 6950 2GB
- Infiniband QDR
- 8 CPU vs 8 GPUで比較
- Silicon 216原子, Gamma 点
- プログラムパッケージxTAPPを基礎
- ブロックサイズ20

# ベンチマーク結果

## I SCF時間

平面波のカット オフ波数 [a.u.]	3.6	4.0	4.8	5.0	5.4	6.4
FFTメッシュ	$72^3$	$80^3$	$96^3$	$100^3$	$108^3$	$128^3$
Xeon X5690 [s]	378	549	994	1188	1583	2255
Radeon HD 6950 [s]	169	297	342	500	534	749
加速率	2.23	1.84	2.91	2.37	2.96	3.01

# ベンチマーク結果

- 最大 3 倍の高速化
- 十分に大きなデータについては  
バンド幅律速
- $160/32 = 5$ 倍が理想性能比
- 基数 5 が苦手？

# AMD vs NVIDIA

ブロック演算部のみでの比較

	Radeon HD 6950	Tesla C2070
コアの経過時間 [s]	5.43	5.01

- 128<sup>3</sup> FFT メッシュ
  - ブロックサイズ 20
  - AMDAPPML 2.5 vs CUFFT 4.0
- 条件



# ブロック演算部の中身

経過時間 [s]	データ 転送	FFT	その他の GPU演算	その他
Radeon HD 6950	0.53	3.09	1.05	0.76
Tesla C2070	0.55	3.05	1.22	0.19

- $128^3$  FFT メッシュ

条件

- ブロックサイズ 20
- AMDAPPML 2.5 vs CUFFT 4.0

# まとめ

- GPUを用いて交換相互作用の計算を高速化
- Xeon X5690 vs Radeon HD 6950で最大3倍
  - AMD Radeon HD 6950は2.6万円
- ブロック化でCPU・GPU間のデータ転送に必要な時間を小さくした