Mixed mode in CASINO

Lucian Anton

NAG, UK

Dario Alfè,

UCL, UK

Cray User Group

Edinburgh, May 24-27, 2010

Experts in numerical algorithms
and HPC services

Outline

- Introduction
- OpenMP sectors
  - One particle orbitals
  - Jastrow
  - Ewald
  - Update $\overline{D}$
  - e-e relative distances
- Performance results and discussion

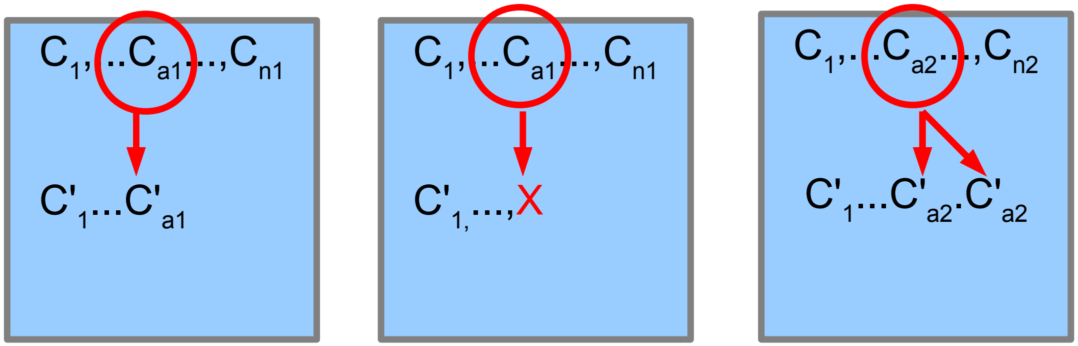
Introduction I

- CASINO is Quantum Monte Carlo code used to compute electronic structure properties of large molecules or solids.
- It scales up to 40,000 cores.
- For models with large number of electrons one configuration computation needs to be accelerated (N^2 scaling).

CASINO is developed by TCM group, Cambridge University.

http://www.tcm.phy.cam.ac.uk/~mdt26/casino2.htm

Introduction II



$$t = N_{step} \times \frac{N_{pop}}{P} \times \frac{t_{config}}{N_{thread}} \qquad t_{config} \sim N_e^2$$

Numerical kernels: OPO

$$i\hbar \frac{\partial \Psi(R,t)}{\partial t} = -\frac{\hbar^2}{2\mathrm{m}} \sum_{i=1}^{N_e} \nabla_i^2 \Psi(R,t) + V(R,R_I)\Psi(R,t)$$

$$D_\uparrow = \begin{vmatrix} \phi_1(r_1) & \phi_1(r_2) & \cdots & \phi_1(r_{N_\uparrow}) \\ \vdots & \cdots & \cdots & \vdots \\ \phi_{N_\uparrow}(r_1) & \phi_{N_\uparrow}(r_2) & \dots & \phi_{N_\uparrow}(r_{N_\uparrow}) \end{vmatrix}$$

One particle orbital (OPO)

$$\Psi = e^{J(\alpha,R,R_I)} D_\uparrow D_\downarrow$$

$$J(R) = \sum_{\substack{i<j, \\ \sigma_i,\sigma_j}} u_{\sigma_i,\sigma_j}(\alpha, |r_i - r_j|)$$

Numerical kernels: JASTROW

$$i\hbar \frac{\partial \Psi(R,t)}{\partial t} = -\frac{\hbar^2}{2\text{m}} \sum_{i=1}^{N_e} \nabla_i^2 \Psi(R,t) + V(R,R_I)\Psi(R,t)$$

$$D_\uparrow = \begin{vmatrix} \phi_1(r_1) & \phi_1(r_2) & \cdots & \phi_1(r_{N_\uparrow}) \\ \vdots & \cdots & \cdots & \vdots \\ \phi_{N_\uparrow}(r_1) & \phi_{N_\uparrow}(r_2) & \ldots & \phi_{N_\uparrow}(r_{N_\uparrow}) \end{vmatrix}$$

Jastrow (JAS)

$$\Psi = e^{J(\alpha, R, R_I)} D_\uparrow D_\downarrow$$

$$J(R) = \sum_{\substack{i<j, \\ \sigma_i, \sigma_j}} u_{\sigma_i, \sigma_j}(\alpha, |r_i - r_j|)$$

Numerical kernels: Ewald summation

$$i\hbar\frac{\partial \Psi(R,t)}{\partial t} = -\frac{\hbar^2}{2\mathrm{m}}\sum_{i=1}^{N_e}\nabla_i^2\Psi(R,t) + V(R,R_I)\Psi(R,t)$$

$$D_\uparrow = \begin{vmatrix} \phi_1(r_1) & \phi_1(r_2) & \cdots & \phi_1(r_{N_\uparrow}) \\ \vdots & \cdots & \cdots & \vdots \\ \phi_{N_\uparrow}(r_1) & \phi_{N_\uparrow}(r_2) & \ldots & \phi_{N_\uparrow}(r_{N_\uparrow}) \end{vmatrix}$$

Ewald Summation (EWA)

$$\Psi = e^{J(\alpha,R,R_I)} D_\uparrow D_\downarrow$$

$$J(R) = \sum_{\substack{i<j,\\ \sigma_i,\sigma_j}} u_{\sigma_i,\sigma_j}(\alpha, |r_i - r_j|)$$

Numerical kernels: Update D

$$i\hbar\frac{\partial \Psi(R,t)}{\partial t} = -\frac{\hbar^2}{2\mathrm{m}}\sum_{i=1}^{N_e}\nabla_i^2\Psi(R,t) + V(R,R_I)\Psi(R,t)$$

$$D_\uparrow = \begin{vmatrix} \phi_1(r_1) & \phi_1(r_2) & \cdots & \phi_1(r_{N_\uparrow}) \\ \vdots & \cdots & \cdots & \vdots \\ \phi_{N_\uparrow}(r_1) & \phi_{N_\uparrow}(r_2) & \dots & \phi_{N_\uparrow}(r_{N_\uparrow}) \end{vmatrix}$$

$$\Psi = e^{J(\alpha,R,R_I)} D_\uparrow D_\downarrow$$

$$J(R) = \sum_{\substack{i<j,\\ \sigma_i,\sigma_j}} u_{\sigma_i,\sigma_j}(\alpha, |r_i - r_j|)$$

Update $\bar{D}$ ($\bar{D}$)

Numerical kernels: R_{ee}

$$i\hbar\frac{\partial\Psi(R,t)}{\partial t}=-\frac{\hbar^2}{2\mathrm{m}}\sum_{i=1}^{N_e}\nabla_i^2\Psi(R,t)+V(R,R_I)\Psi(R,t)$$

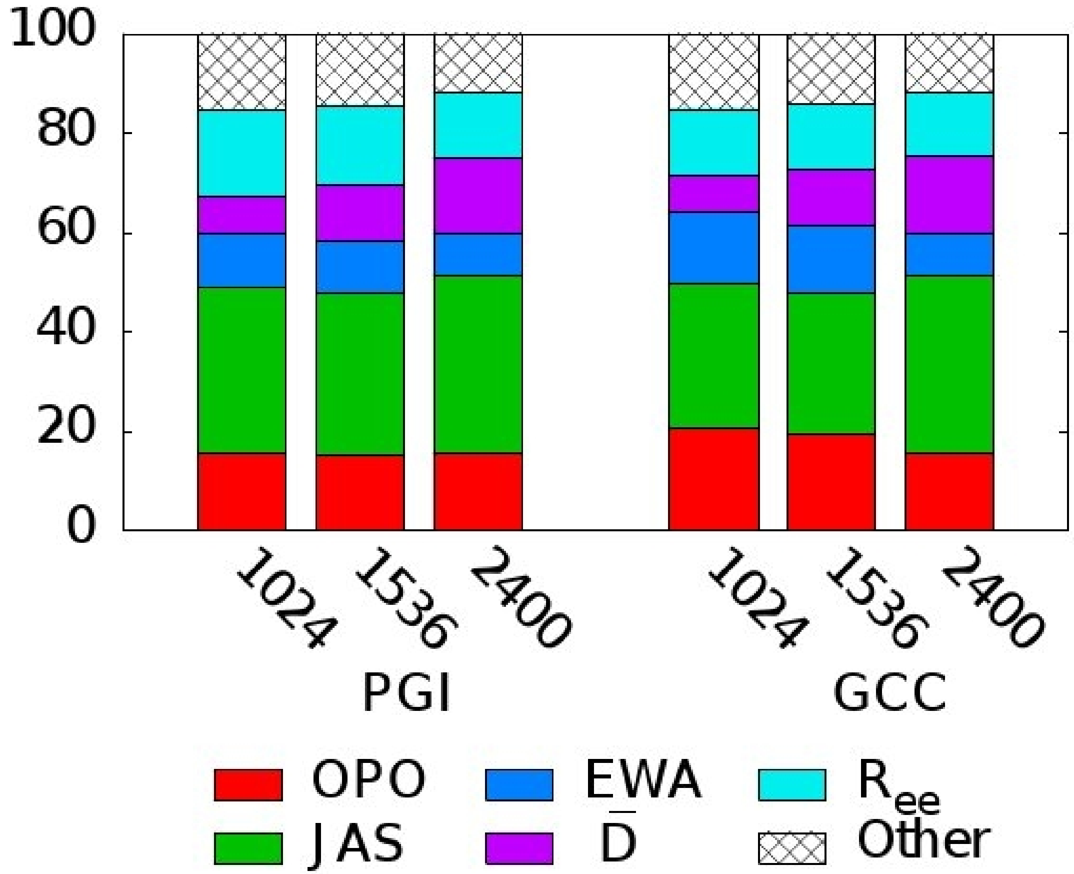
$$D_\uparrow=\begin{vmatrix}\phi_1(r_1) & \phi_1(r_2) & \cdots & \phi_1(r_{N_\uparrow})\\ \vdots & \cdots & \cdots & \vdots\\ \phi_{N_\uparrow}(r_1) & \phi_{N_\uparrow}(r_2) & \dots & \phi_{N_\uparrow}(r_{N_\uparrow})\end{vmatrix}$$

$$\Psi=e^{J(\alpha,R,R_I)}D_\uparrow D_\downarrow$$

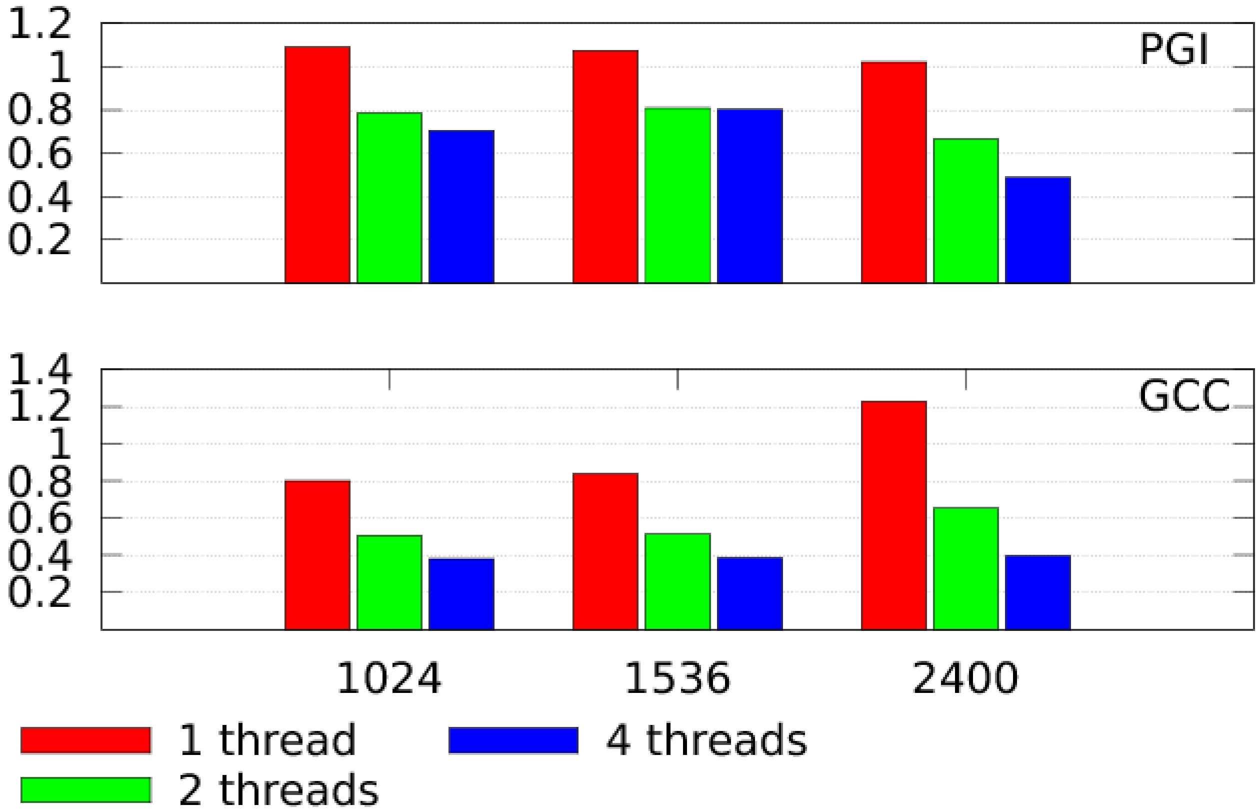
$$J(R)=\sum_{\substack{i<j,\\ \sigma_i,\sigma_j}}u_{\sigma_i,\sigma_j}(\alpha,|r_i-r_j|)$$

e-e distances (R_{ee})

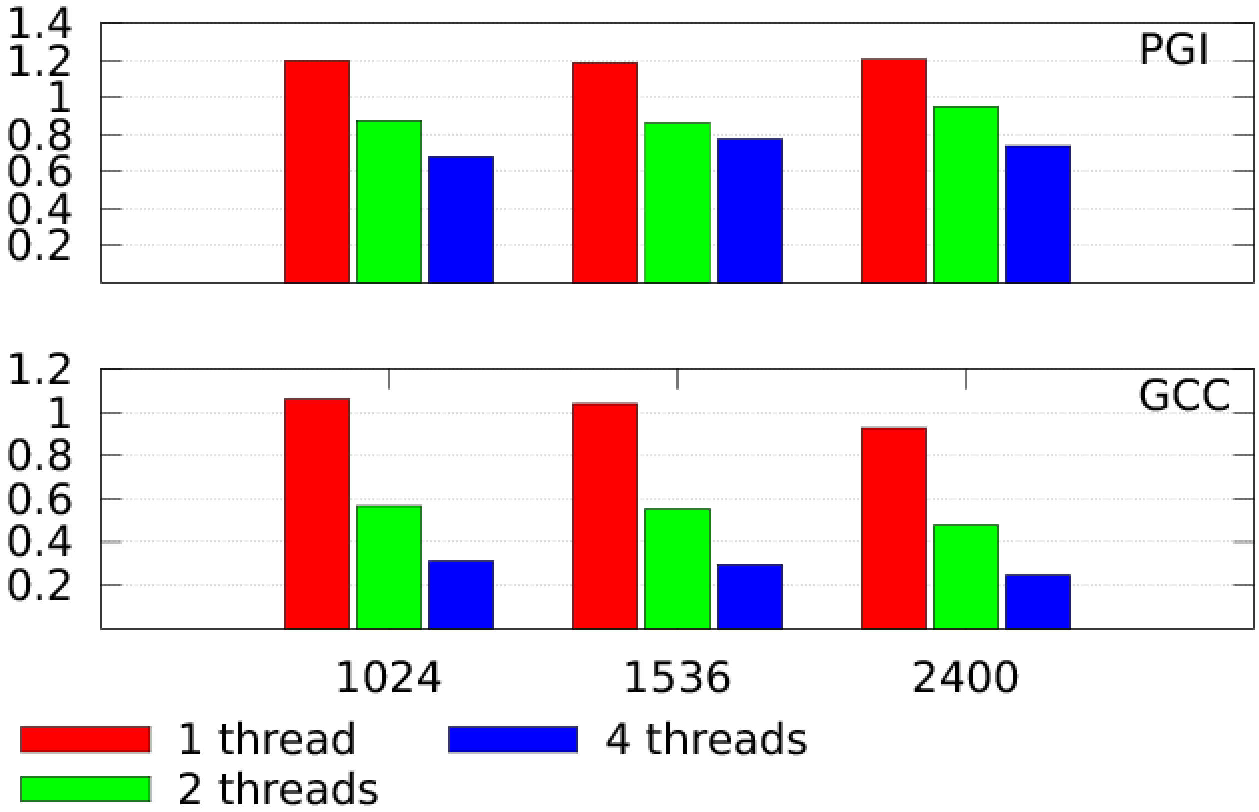
Time consumed by numerical kernels



OPO

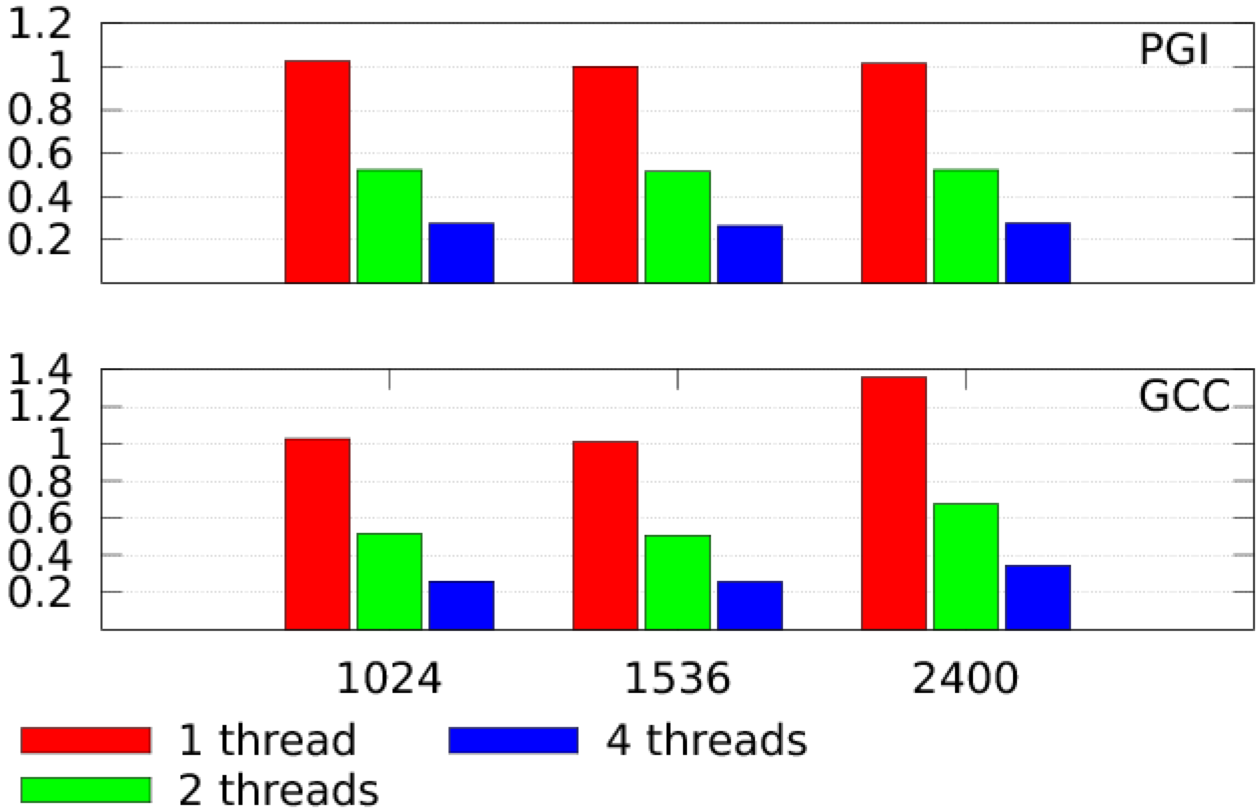

1.2
1
0.8
0.6
0.4
0.2
PGI
1.4
1.2
1
0.8
0.6
0.4
0.2
GCC
1024
1536
2400
1 thread
4 threads
2 threads

Jastrow

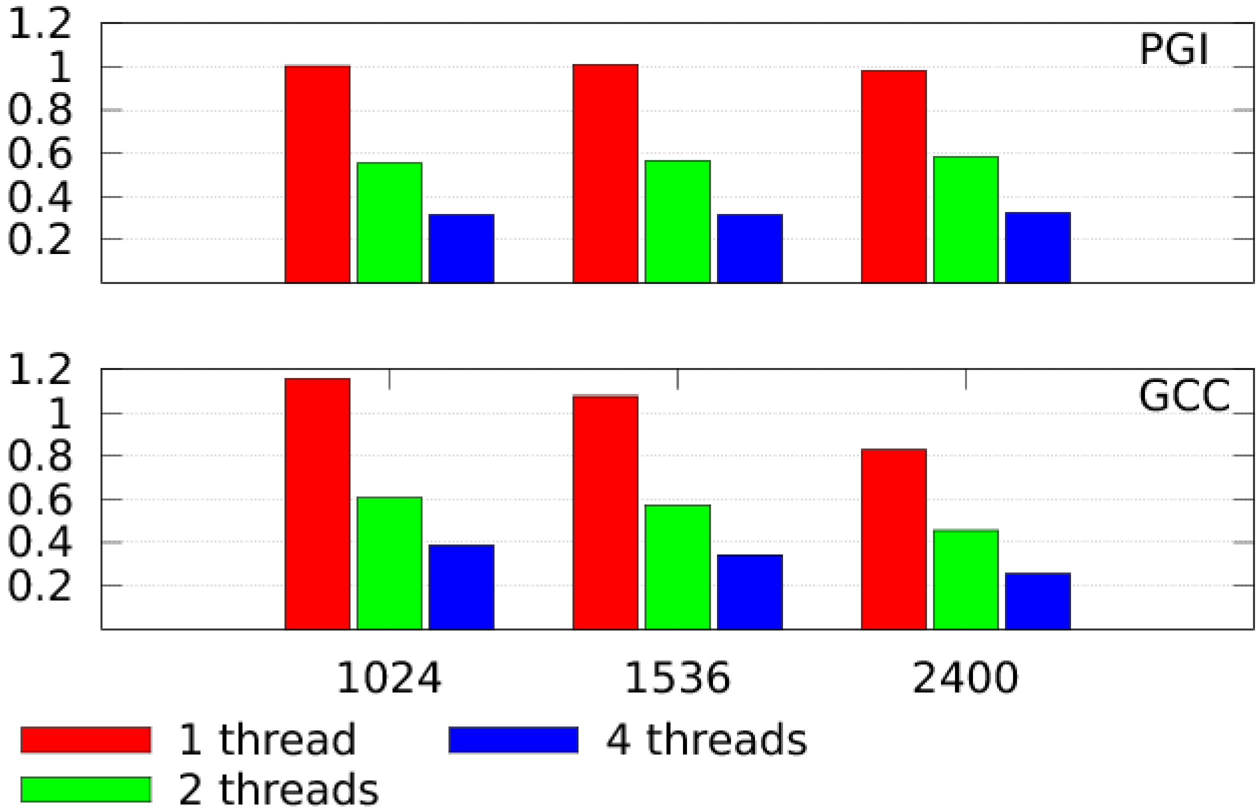


1.4
1.2
1
0.8
0.6
0.4
0.2
PGI
1.2
1
0.8
0.6
0.4
0.2
GCC
1024
1536
2400
1 thread
4 threads
2 threads

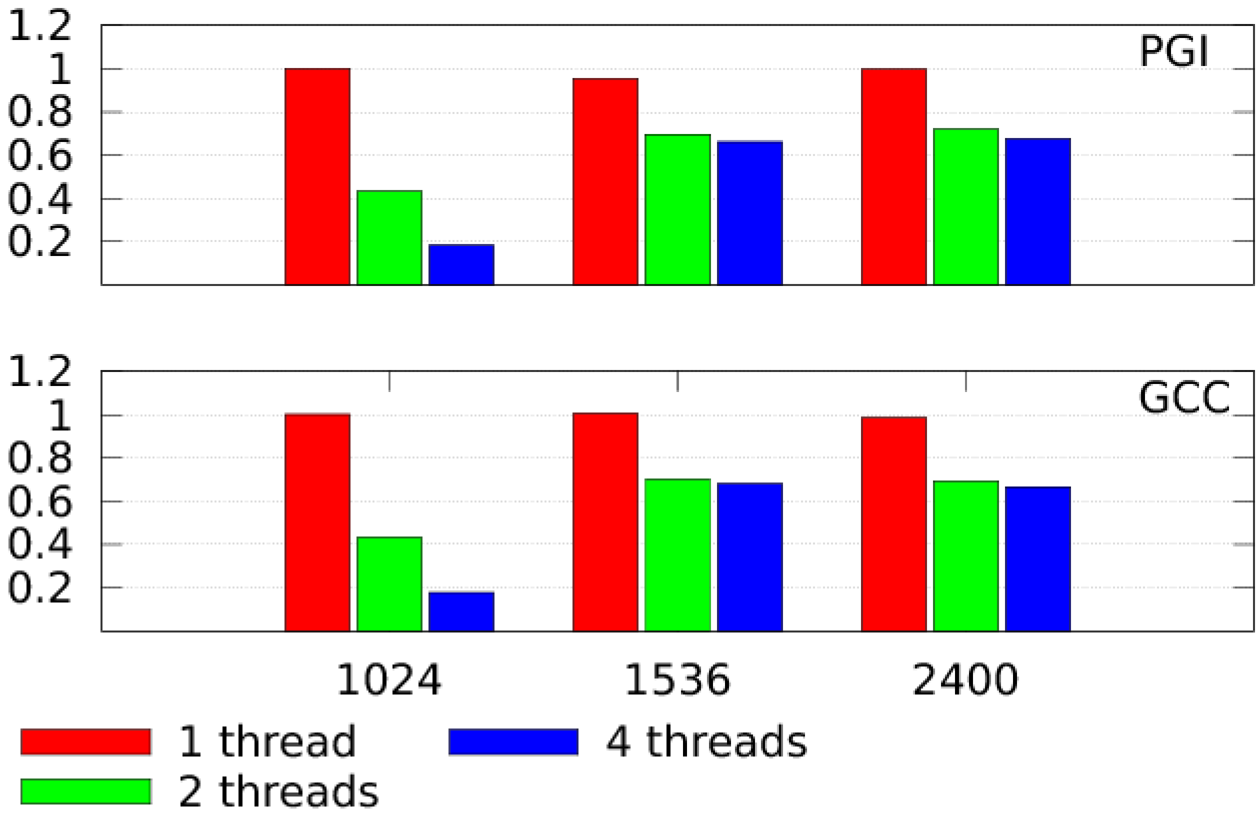
Ewald


1.2
1
0.8
0.6
0.4
0.2
PGI
1.4
1.2
1
0.8
0.6
0.4
0.2
GCC
1024
1536
2400
1 thread
4 threads
2 threads

R_{ee}

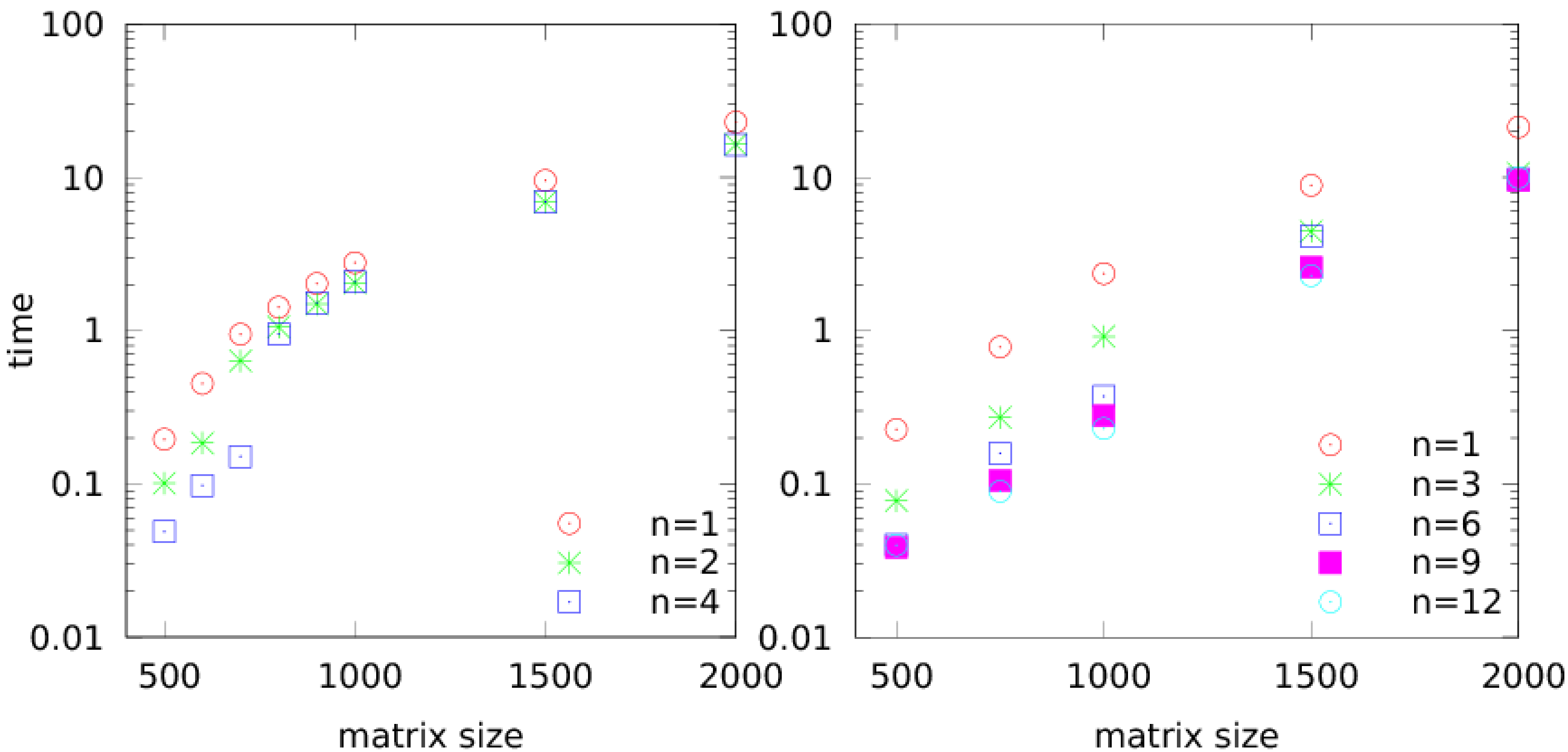

1.2
1
0.8
0.6
0.4
0.2
PGI
GCC
1024
1536
2400
1 thread
2 threads
4 threads

$\bar{D}$ matrix



1.2
1
0.8
0.6
0.4
0.2
PGI
GCC
1024
1536
2400
1 thread
2 threads
4 threads

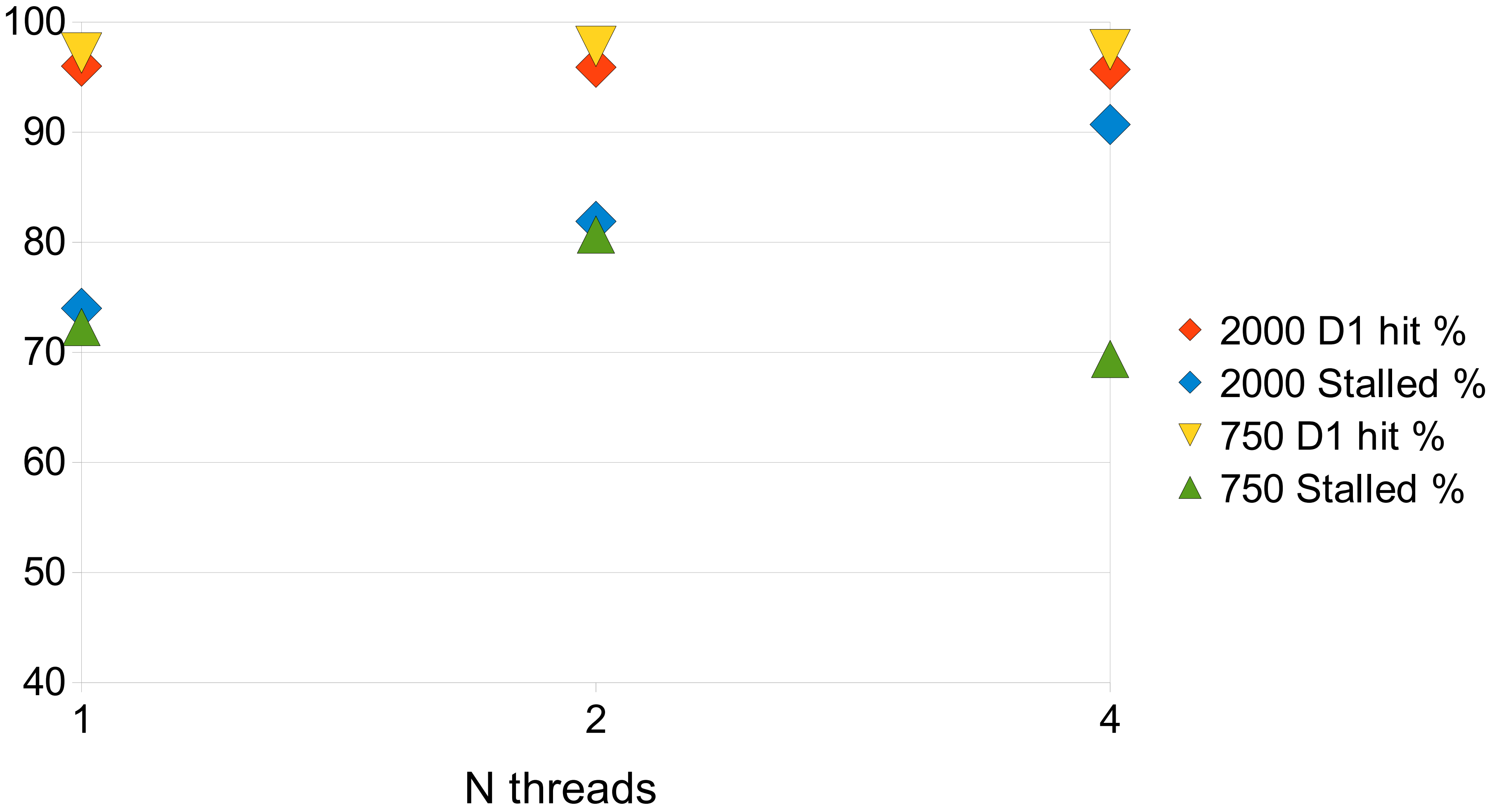
Scaling loss in matrix $\overline{\mathrm{D}}$ computation



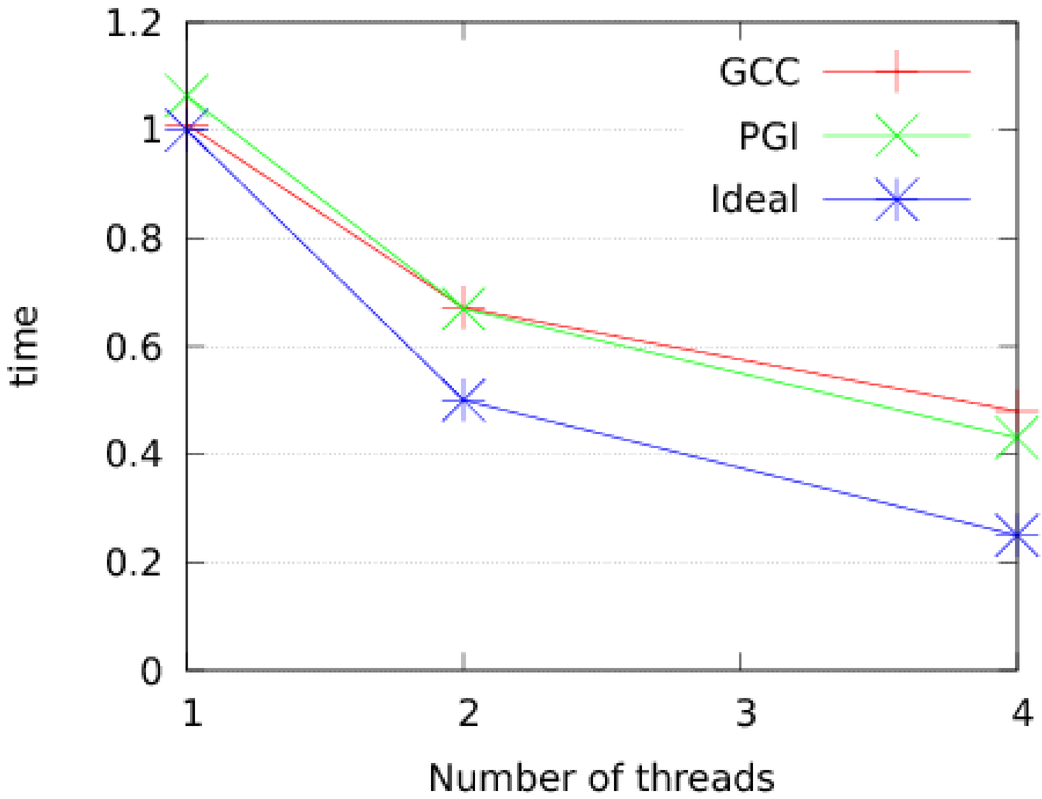
$\sqrt{(4L_2+L_3)/8} \approx 724$

$\sqrt{(6L_2+L_3)/8} \approx 1086$

PAPI counters



Aggregate performance N_e=1024



Conclusions

- Mixed mode in CASINO achieves speed up factors in the range of 1.5/4 for 2/4 OpenMP threads for models with 1000 electrons on XT4.
- Our analysis predicts that the same speed up factors extends up to 2000 electrons models on XT6 due to larger cache memory.
- We have found a memory bound kernel ($\overline{D}$) that need further study.
- What is cause of the large overheads?

HECToR

HIGH END COMPUTING TERASCALE RESOURCE

A Research Councils UK High End Computing Service