œ

Theory NC PAW LOBPCG SCF

mplementation

Hypothesis

Principles

Results

Benchmarks

Optimisations

NC

PAW

PIOWIC

Conclusion

Conclusion

Perspectives

BandFFT parallelization of ABINIT

François Bottin & Gilles Zérah

CEA/DIF, BP 12 Bruyères-le-Châtel Département de Physique Théorique et Appliquée Service de Physique de la Matière Condensée

3rd International ABINIT Developer Workshop Liège (Belgium), 29-31 January 2007

œ

Theory
NC
PAW
LOBPCG
SCF
Implementation
Hypothesis
Principles
Results
Benchmarks
Optimisations
NC
PAW
Howto
Conclusion
Conclusion
Perspectives

• Large supercell *ab initio* calculations are very time consuming.

Parallelization: one of the main computational tasks



Parallelization: one of the main computational tasks

Theory
NC
PAW
LOBPCG
SCF
Implemer
Hypothesis

Large supercell *ab initio* calculations are very time consuming.Increase of the powerful and number of processors of

supercomputers

Site	Computer	Proc.	Tflops
DOE/NNSA/LLNL (US)	BlueGene	131072	281
NNSA/Sandia National Laboratories (US)	Red Storm	26544	101
IBM T. J. Watson Research Center (US)	BGW	40960	91
DOE/NNSA/LLNL (US)	ASC Purple	12208	76
Bacelona Supercomputing Center (Spain)	Mare Nostrum	10240	63
NNSA/Sandia National Laboratories (US)	Thunderbird	9024	53
CEA/DAM (France)	TERA-10	9968	53
NASA/Ames Research Center/NAS (US)	Columbia	10160	52
GSIC Center (Japan)	TSUBAME	11088	47
Oak Ridge National Laboratories (US)	Jaguar	10424	43



Parallelization: one of the main computational tasks

Theory
NC
PAW
LOBPCG
SCF
Implemer
Hypothesis

- Large supercell *ab initio* calculations are very time consuming.
- Increase of the powerful and number of processors of supercomputers

Site	Computer	Proc.	Tflops
DOE/NNSA/LLNL (US)	BlueGene	131072	281
NNSA/Sandia National Laboratories (US)	Red Storm	26544	101
IBM T. J. Watson Research Center (US)	BGW	40960	91
DOE/NNSA/LLNL (US)	ASC Purple	12208	76
Bacelona Supercomputing Center (Spain)	Mare Nostrum	10240	63
NNSA/Sandia National Laboratories (US)	Thunderbird	9024	53
CEA/DAM (France)	TERA-10	9968	53
NASA/Ames Research Center/NAS (US)	Columbia	10160	52
GSIC Center (Japan)	TSUBAME	11088	47
Oak Ridge National Laboratories (US)	Jaguar	10424	43

• Aim: to use and to be efficient on these supercomputers.

3rd International ABINIT Developer Workshop, Liège (Belgium), 29-31 January 2007

F. Bottin

œ

Theoretical background

- The Norm-Conserving method (NC)
- The Projector Augmented-Wave method (PAW)
- Locally Optimal Block Preconditioned Conjugate Gradient method (LOBPCG)
- The self-consistent loop (SCF)

BandFFT parallelization
 Hypothesis – Requirements

- Implementation
- . Hypothesis
- Principles
- Results
- Benchmarks
- Optimisation
- NC
- PAW
- Howte
- Conclusio
- Conclusion
- Perspectives

- Standard implementation
- BandFFT Results
- Benchmarks
- Beyond the standard implementation Optimisations
- Norm-conserving Results
- PAW results
- How to use the BandFFT parallelization
- 4 Conclusion Perspectives
 - Conclusion
 - Perspectives

Theory

Theoretical background The Norm-Conserving method (NC) The Projector Augmented-Wave method (PAW) ۰ Locally Optimal Block Preconditioned Conjugate Gradient method (LOBPCG) The self-consistent loop (SCF) BandFFT parallelization Hypothesis – Requirements How to use the BandFFT parallelization



The Norm-Conserving method (NC)

Theory

NC

PAW

LOBPC

Implomente

Hypothesis

Principles

Results

Benchmark:

Optimisatio

NC

PAW

Howto

Conclusic

Conclusion

Perspectives



$\tilde{n}(\mathbf{r}) = \sum_{nk} f_{nk} |\tilde{\Psi}_{nk}(\mathbf{r})|^2$



The Projector Augmented-Wave method (PAW)

Theory

PAW

LOBPCG

SCF

mplementatio

Ĥ

Hypothesis Principles

Results

Benchmarks

Optimisation

NC

PAW

Howto

Conclusion

Conclusion

Perspectives

$$\tilde{\mathcal{U}} = -\frac{\Delta}{2} + v_{\rm loc} + \sum_{I} \sum_{ij \in I} \big| \tilde{\mathcal{P}}^{I}_{i} \big\rangle D^{I}_{ij} \big\langle \tilde{\mathcal{P}}^{I}_{j} \big| \quad {\rm with} \quad \tilde{\mathcal{H}} \tilde{\Psi}_{nk} = \varepsilon_{nk} \mathcal{O} \tilde{\Psi}_{nk}$$

Densities – Overlap operator

$$\hat{n}(\mathbf{r}) \quad \mathrm{and} \quad \rho^{I}_{ij} = \sum_{nk} f_{nk} \langle \tilde{\Psi}_{nk} | \tilde{\mathcal{P}}^{I}_{i} \rangle \langle \tilde{\mathcal{P}}^{I}_{j} | \tilde{\Psi}_{nk} \rangle$$

$$\mathcal{O} = \mathcal{I} + \sum_{I} \sum_{ij \in I} \big| \tilde{\mathcal{P}}^{I}_{i} \big\rangle (\langle \Phi^{I}_{i} | \Phi^{I}_{j} \rangle) - \langle \tilde{\Phi}^{I}_{i} | \tilde{\Phi}^{I}_{j} \rangle) \big\langle \tilde{\mathcal{P}}^{I}_{j} \big|$$

$$n(\mathbf{r}) = \tilde{n}(\mathbf{r}) + \sum_{I} (n^{1,I}(\mathbf{r}) - \tilde{n}^{1,I}(\mathbf{r}))$$

	Alg	gorithm 1 LOBPCG
e	Re	quire: $\Psi^0 = \{\Psi^0_1, \dots, \Psi^0_m\}$ close to the minimum and K a pre-
heory		conditioner; $\mathbf{P} = \{\mathbf{P}_1^{(0)}, \dots, \mathbf{P}_m^0\}$ is initialized to 0.
NC PAW	1:	for $i=0,1,\ldots,\kappa$ do
LOBPCG	2:	$\Upsilon^{(\mathrm{i})}=\Upsilon(oldsymbol{\Psi}^{(\mathrm{i})})$
scr mplementation	3:	$\mathbf{R}^{(\mathrm{i})} = \mathcal{H} \mathbf{\Psi}^{(\mathrm{i})} - \Upsilon^{(\mathrm{i})} \mathcal{O} \mathbf{\Psi}^{(\mathrm{i})}$
Hypothesis	4:	$\mathbf{W}^{(\mathrm{i})} = \mathrm{K} \mathbf{R}^{(\mathrm{i})}$
Principles Results	5:	The Rayleigh-Ritz method is applied within the subspace $\Xi =$
Benchmarks Optimisations		$\left\{ \mathbf{P}_{1}^{(i)}, \dots, \mathbf{P}_{m}^{(i)}, \Psi_{1}^{(i)}, \dots, \Psi_{m}^{(i)}, \mathbf{W}_{1}^{(i)}, \dots, \mathbf{W}_{m}^{(i)} \right\}$
NC PAW	6:	$\mathbf{\hat{\Psi}}^{(i+1)} = \Delta^{(i)} \mathbf{\Psi}^{(i)} + \Lambda^{(i)} \mathbf{W}^{(i)} + \Gamma^{(i)} \mathbf{P}^{(i)}$
Howto	7:	$\mathbf{P}^{(i+1)} = \Lambda^{(i)} \mathbf{W}^{(i)} + \Gamma^{(i)} \mathbf{P}^{(i)}$
Conclusion	8:	end for

More efficient than CG in many cases 1

¹A. Knyazev Toward the Optimal Preconditioned Eigensolver: Locally Optimal Block Preconditioned Conjugate Gradient Method, SIAM Journal on Scientific Computing **23**, 517 (2001).

 $\tilde{\Psi}$

NC NC

(H)

PAW

LOBPCO

SCF

Implementation

Hypothesis

Principles

Results

Benchmarks

Optimisations

NC

PAW

Howto

Conclusion

Conclusion

Perspectives

$\mathbf{r}_{n}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}}$	
$[\tilde{\mathbf{n}} + \hat{\mathbf{n}}](\mathbf{r}) $ and $\rho_{ij} \leftarrow \{\mathbf{c}_{\mathbf{n}}(\mathbf{G}); \epsilon_{\mathbf{n}}\}$	
$ert \ \ \bigvee \qquad $	
$ \downarrow \qquad \qquad \uparrow \\ \langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \tilde{\mathcal{H}} \tilde{\Psi}_n \rangle = \epsilon_n \langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \mathcal{O} \tilde{\Psi}_n \rangle $	

Theory

DAW

LOBPC

SCF

Implementatio

Hypothesis

Principles

Results

Benchmarks

Optimisation

NC

PAW

Howto

Conclusio

Conclusion

$$\begin{split} \tilde{\Psi}_{n}(\mathbf{r}) &= \sum_{\mathbf{G}} c_{n}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}} \\ & \downarrow \\ & [\tilde{n}+\hat{n}](\mathbf{r}) \quad \text{and} \quad \rho_{ij} \quad \longleftarrow \quad \{c_{n}(\mathbf{G}); \epsilon_{n}\} \\ & \downarrow \qquad \qquad \uparrow \\ v_{loc}(\mathbf{r}) \quad \text{and} \quad v_{nl}(\mathbf{r}) \qquad | \quad \tilde{\mathcal{H}} - \epsilon_{n}\mathcal{O} \mid = 0 \\ & \downarrow \qquad \qquad \uparrow \\ \langle e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}} \mid \tilde{\mathcal{H}} \mid \tilde{\Psi}_{n} \rangle &= \epsilon_{n} \langle e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}} \mid \mathcal{O} \mid \tilde{\Psi}_{n} \rangle \end{split}$$

Time consuming parts

• The non-local terms (nonlop).

Theory

DAW

LOBPC

SCF

Implementation

Hypothesis

Principles

Results

Benchmarks

Optimisation

NC

PAW

Howto

Conclusio

Conclusion

$$\begin{split} \tilde{\Psi}_{n}(\mathbf{r}) &= \sum_{\mathbf{G}} c_{n}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}} \\ & \downarrow \\ & [\tilde{n}+\hat{n}](\mathbf{r}) \quad \text{and} \quad \rho_{ij} \quad \longleftarrow \quad \{c_{n}(\mathbf{G});\epsilon_{n}\} \\ & \downarrow \\ v_{loc}(\mathbf{r}) \quad \text{and} \quad v_{nl}(\mathbf{r}) \qquad | \quad \tilde{\mathcal{H}} - \epsilon_{n}\mathcal{O} \mid = 0 \\ & \downarrow \\ & \langle e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}} | \quad \tilde{\mathcal{H}} | \quad \tilde{\Psi}_{n} \rangle = \epsilon_{n} \langle e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}} | \quad \mathcal{O} | \quad \tilde{\Psi}_{n} \rangle \end{split}$$

Time consuming parts

- The non-local terms (nonlop).
- The resolution of the KS equations (lobpcg).

3rd International ABINIT Developer Workshop, Liège (Belgium), 29-31 January 2007

Theor

DA1A/

LOBPCG

SCF

Implementatior

Hypothesis

Principles

Results

Benchmarks

Optimisation

NC

PAW

Howto

Conclusio

Conclusion

$$\begin{split} \tilde{\Psi}_{n}(\mathbf{r}) &= \sum_{\mathbf{G}} c_{n}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \\ & \downarrow \\ & [\tilde{n}+\hat{n}](\mathbf{r}) \quad \text{and} \quad \rho_{ij} \quad \longleftarrow \quad \{c_{n}(\mathbf{G}); \epsilon_{n}\} \\ & \downarrow \\ & \downarrow \\ v_{loc}(\mathbf{r}) \quad \text{and} \quad v_{nl}(\mathbf{r}) \qquad \qquad | \quad \tilde{\mathcal{H}} - \epsilon_{n} \mathcal{O} \mid = 0 \\ & \downarrow \\ & \langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \mid \tilde{\mathcal{H}} \mid \tilde{\Psi}_{n} \rangle = \epsilon_{n} \langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \mid \mathcal{O} \mid \tilde{\Psi}_{n} \rangle \end{split}$$

Time consuming parts

- The non-local terms (nonlop).
- The resolution of the KS equations (lobpcg).
- The diagonalisation within the sub-space (subdiago+orthon).

 Ψ

Theor NC

PAW

LOBPCG

SCF

Implementation

Hypothesis

Principles

Results

Benchmarks

Optimisation

NC

 $P\!AW$

Howto

Conclusio

Conclusion

$$\begin{split} \mathbf{r}_{n}(\mathbf{r}) &= \sum_{\mathbf{G}} \mathbf{c}_{n}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}} \\ & \downarrow \\ \begin{bmatrix} \mathbf{\tilde{n}} + \hat{n} \end{bmatrix}(\mathbf{r}) & \text{and} \quad \rho_{ij} & \longleftarrow \quad \{\mathbf{c}_{n}(\mathbf{G}); \epsilon_{n}\} \\ & \downarrow & \uparrow \\ \mathbf{v}_{loc}(\mathbf{r}) & \text{and} \quad v_{nl}(\mathbf{r}) & | \tilde{\mathcal{H}} - \epsilon_{n}\mathcal{O} | = 0 \\ & \downarrow & \uparrow \\ \langle e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}} | \tilde{\mathcal{H}} | \tilde{\Psi}_{n} \rangle &= \epsilon_{n} \langle e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}} | \mathcal{O} | \tilde{\Psi}_{n} \rangle \end{split}$$

Time consuming parts

- The non-local terms (nonlop).
- The resolution of the KS equations (lobpcg).
- The diagonalisation within the sub-space (subdiago+orthon).
- The calculation of the density and local potential (FFT).

 $^{3^{}m rd}$ International ABINIT Developer Workshop, Liège (Belgium), 29-31 January 2007

Theory

DAIA

LOBPCO

SCF

Implementation

Hypothesis

Principles

Results

Benchmarks

Optimisation

NC

PAW

Howto

Conclusio

Conclusion

Perspectives

$$\begin{split} \tilde{\Psi}_{n}(\mathbf{r}) &= \sum_{\mathbf{G}} c_{n}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}} \\ & \downarrow \\ & [\tilde{n}+\hat{n}](\mathbf{r}) \quad \text{and} \quad \rho_{ij} \quad \longleftarrow \quad \{c_{n}(\mathbf{G}); \epsilon_{n}\} \\ & \downarrow \\ & \downarrow \\ v_{loc}(\mathbf{r}) \quad \text{and} \quad v_{nl}(\mathbf{r}) \qquad | \quad \tilde{\mathcal{H}} - \epsilon_{n}\mathcal{O} \mid = 0 \\ & \downarrow \\ & \langle e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}} | \quad \tilde{\mathcal{H}} | \quad \tilde{\Psi}_{n} \rangle = \epsilon_{n} \langle e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}} | \quad \mathcal{O} | \quad \tilde{\Psi}_{n} \rangle \end{split}$$

amous parallelizations

- ${\scriptstyle \bullet}$ over the ${\bf k}\text{-points}.$
- \bullet over the bands \rightarrow require a block eigensolver.
- over the plane-waves \rightarrow require a parallel 3dim-FFT algorithm.

F. Bottin

œ

Theoretical background

- The Norm-Conserving method (NC)
- The Projector Augmented-Wave method (PAW)
- Locally Optimal Block Preconditioned Conjugate Gradient method (LOBPCG)
- The self-consistent loop (SCF)

Implementation

- Hypothesis
- Principles
- Results
- Benchmarks
- Optimisation
- NC
- PAW
- Howt
- Conclusion
- Conclusion

- BandFFT parallelization
- Hypothesis Requirements
- Standard implementation
 - andFFT Results
 - Benchmarks
- Beyond the standard implementation Optimisations
- Norm-conserving Results
- PAW results
- How to use the BandFFT parallelization
- 4 Conclusion Perspectives
 - Conclusion
 - Perspectives

œ

Hypothesis – Requirements

Theory	• M bands and a P^3 FFT grid
NC	
PAW	• Two-dimensional mxp grid of processors with m
LOBPCG	• Two dimensional mxp grid of processors with m
SCF	"band-processors" and p "FFT-processors".
Implementation	r i i r r
Hypothesis	
Principles	
Results	
Benchmarks	
Optimisations	
NC	
PAW	
Howto	• m and p have to be commensurate with M and P, respectively
Conclusion	
Conclusion	• For the 3dim-FFT one have to ensure: $p < \frac{p}{2}$
Perspectives	p = 2

$ilde{\Psi}_n(\mathbf{r}) = \sum_{\mathbf{G}} c_n(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}}$ $[\tilde{\mathbf{n}} + \hat{\mathbf{n}}](\mathbf{r})^{*}$ and $\rho_{ij} \leftarrow \{\mathbf{c}_{\mathbf{n}}(\mathbf{G}); \epsilon_{\mathbf{n}}\}$ $|\tilde{\mathcal{H}} - \epsilon_{n}\mathcal{O}| = 0$ Principles $v_{loc}(\mathbf{r})$ and $v_{nl}(\mathbf{r})$ $\langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} | \tilde{\mathcal{H}} | \tilde{\Psi}_n \rangle = \epsilon_n \langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} | \mathcal{O} | \tilde{\Psi}_n^{\dagger} \rangle$ All the coefficients are on the proc. 0 $c_n(\mathbf{g})$

$\tilde{\Psi}_n(\mathbf{r}) = \sum_{\mathbf{C}} c_n(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}}$ $[\tilde{\mathbf{n}} + \hat{\mathbf{n}}](\mathbf{r})^{\downarrow}$ and $\rho_{ij} \leftarrow \{c_n(\mathbf{G}); \epsilon_n\}$ $|\tilde{\mathcal{H}} - \epsilon_{\mathrm{n}}\mathcal{O}| = 0$ Principles $v_{loc}(\mathbf{r})$ and $v_{nl}(\mathbf{r})$ $\langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} | \tilde{\mathcal{H}} | \tilde{\Psi}_n \rangle = \epsilon_n \langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} | \mathcal{O} | \tilde{\Psi}_n \rangle$ Distribution over the p "FFT-processors" $c_n(g_1)$ $c_n(g_2)$... $c_n(g_p)$

F.	Bottin

Standard implementation

 $\tilde{\Psi}_{r}$

Theory

NC

PAW

LOBPC

SCF

Implementatior

Hypothesis

Principles

Results

Benchmarks

Optimisation:

NC

PAW

Howto

Perspectives

$\mathbf{r}_{\mathbf{h}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{\mathbf{n}}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}}$	
$[\tilde{\mathbf{n}} + \hat{\mathbf{n}}](\mathbf{r}) \downarrow$ and $\rho_{\mathrm{ij}} \leftarrow$	- $\{c_n(\mathbf{G}); \epsilon_n\}$
$ert \mathop{\bigvee}\limits_{\mathrm{v}_{\mathrm{loc}}}(\mathbf{r}) \mathrm{and} \mathrm{v}_{\mathrm{nl}}(\mathbf{r})$	$ \tilde{\mathcal{H}} - \epsilon_{\mathbf{n}} \mathcal{O} = 0$
$ \langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \tilde{\mathcal{H}} \tilde{\Psi}_n \rangle = \epsilon_n \langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \tilde{\mathcal{H}} \tilde{\mathcal{H}}$	$ + \mathbf{G} \rangle \cdot \mathbf{r} \left \mathcal{O} \right \tilde{\Psi}_{\mathrm{n}} angle$

Distribution over the m "band-processors"

 $\begin{array}{ccccc} c_n(g_{11}) & c_n(g_{12}) & \dots & c_n(g_{1p}) \\ c_n(g_{21}) & c_n(g_{22}) & \dots & c_n(g_{2p}) \\ \vdots & \vdots & \ddots & \vdots \\ c_n(g_{m1}) & \dots & \dots & c_n(g_{mp}) \end{array}$

CE CE	Standard implementation	
Theory	$ ilde{\Psi}_{\mathrm{n}}(\mathbf{r}) = \sum_{\mathbf{G}} \mathrm{c}_{\mathrm{n}}(\mathbf{G}) \mathrm{e}^{\mathrm{i}(\mathbf{k}+\mathbf{G}).\mathbf{r}}$	
PAW LOBPCG SCF	$[ilde{\mathrm{n}}+ ilde{\mathrm{n}}](\mathbf{r})^{igcell} ext{ and } ho_{\mathrm{ij}} extsf{ } \longleftarrow extsf{ } \{\mathrm{c_n}(\mathbf{G});\epsilon_{\mathrm{n}}\}$	
Implementation Hypothesis Principles	$ \downarrow \qquad \qquad \uparrow \\ \mathbf{v}_{\mathrm{loc}}(\mathbf{r}) \text{and} \mathbf{v}_{\mathrm{pl}}(\mathbf{r}) \qquad \qquad \uparrow \\ \tilde{\mathcal{H}} - \epsilon_{\mathrm{p}}\mathcal{O} \models 0 $	
Results Benchmarks Optimisations NC PAW	$\langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \tilde{\mathcal{H}} \tilde{\Psi}_{n} \rangle = \epsilon_{n} \langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \mathcal{O} \tilde{\Psi}_{n} \rangle$	
Howto Conclusion Conclusion	In lobpcg: blocks of size m	

$$\begin{array}{ccccc} c_{1:m}(g_{11}) & c_{1:m}(g_{12}) & \dots & c_{1:m}(g_{1p}) \\ c_{1:m}(g_{21}) & c_{1:m}(g_{22}) & \dots & c_{1:m}(g_{2p}) \\ \vdots & \vdots & \ddots & \vdots \\ c_{1:m}(g_{m1}) & \dots & \dots & c_{1:m}(g_{mp}) \end{array}$$

3rd International ABINIT Developer Workshop, Liège (Belgium), 29-31 January 2007



$\tilde{\Psi}_n(\mathbf{r}) = \sum_{\mathbf{G}} c_n(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}}$ $[\tilde{\mathbf{n}} + \hat{\mathbf{n}}](\mathbf{r})^{\downarrow}$ and $\rho_{ij} \leftarrow \{c_{\mathbf{n}}(\mathbf{G}); \epsilon_{\mathbf{n}}\}$ $|\tilde{\mathcal{H}} - \epsilon_{\mathrm{n}}\mathcal{O}| = 0$ $v_{loc}(\mathbf{r})$ and $v_{nl}(\mathbf{r})$ Principles $\begin{array}{c|c} c_{1:m}(g_{11}) & c_{1:m}(g_{12}) & \dots & c_{1:m}(g_{1p}) \\ c_{1:m}(g_{21}) & c_{1:m}(g_{22}) & \dots & c_{1:m}(g_{2p}) \end{array}$ $\begin{array}{c|c} \vdots & \vdots & \ddots & \vdots \\ c_{1:m}(g_{m1}) & \dots & \dots & c_{1:m}(g_{mp}) \end{array}$

œ	Standard implementation	
Theory NC PAW LOBPCG	$\tilde{\Psi}_{n}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}$ \downarrow $[\tilde{n} + \hat{n}](\mathbf{r}) \text{and} o:\qquad \{c_{-}(\mathbf{G}): \epsilon_{-}\}$	
SCF Implementation Hypothesis Principles	$\begin{array}{c} \left[\mathbf{n} + \mathbf{n}_{\mathbf{j}}(\mathbf{r}) \right] & \text{and} \mathbf{v}_{\mathbf{n}\mathbf{j}} \\ \downarrow \\ \mathbf{v}_{\text{loc}}(\mathbf{r}) & \text{and} \mathbf{v}_{\mathbf{n}\mathbf{j}}(\mathbf{r}) \\ \end{array} \qquad \qquad$	
Results Benchmarks Optimisations NC PAW	$\langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \tilde{\mathcal{H}} \tilde{\Psi}_{n} \rangle = \epsilon_{n} \langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \mathcal{O} \tilde{\Psi}_{n} \rangle$	
Howto Conclusion Conclusion Perspectives	one have a m×m matrix $c_1(g_{11}) c_2(g_{11}) \dots c_m(g_{11})$	

 $c_1(g_{21})$ $c_2(g_{21})$... $c_m(g_{21})$

3rd International ABINIT Developer Workshop, Liège (Belgium), 29-31 January 2007

. $c_1(g_{m1})$

œ	Standard implementation
Theory NC PAW	$ ilde{\Psi}_{n}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}}$
LOBPCG SCF	$[\tilde{\mathbf{n}} + \hat{\mathbf{n}}](\mathbf{r}) \stackrel{\star}{\downarrow} ext{and} ho_{\mathrm{ij}} \qquad \longleftarrow \{c_{\mathbf{n}}(\mathbf{G}); \epsilon_{\mathbf{n}}\}$
Implementation Hypothesis Principles	$\downarrow \qquad \qquad$
Results Benchmarks Optimisations NC PAW	$\langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \tilde{\mathcal{H}} \tilde{\Psi}_{n} \rangle = \epsilon_{n} \langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \mathcal{O} \tilde{\Psi}_{n} \rangle$
Howto Conclusion Conclusion	which can be transposed
Perspectives	$c_1(g_{11}) c_1(g_{21}) \dots c_1(g_{m1})$

 $c_2(g_{11}) \quad c_2(g_{21}) \quad \dots \quad c_2(g_{m1})$

3rd International ABINIT Developer Workshop, Liège (Belgium), 29-31 January 2007

 $c_{m}(g_{11})$

CE)	Standard implementation
Theory NC PAW LOBPCG	$\tilde{\Psi}_{n}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}$ $\downarrow \qquad \qquad$
SCF Implementation Hypothesis Principles	$\begin{bmatrix} \mathbf{n} + \mathbf{n} \end{bmatrix}(\mathbf{r}) \text{and} p_{ij} \longleftarrow \{\mathbf{c}_{\mathbf{n}}(\mathbf{G}), \mathbf{c}_{\mathbf{n}}\} \\ \downarrow \qquad \qquad \uparrow \\ \mathbf{v}_{loc}(\mathbf{r}) \text{and} \mathbf{v}_{nl}(\mathbf{r}) \qquad \qquad \tilde{\mathcal{H}} - \epsilon_{n}\mathcal{O} = 0$
Results Benchmarks Optimisations NC PAW Howto	$\langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \tilde{\mathcal{H}} \tilde{\Psi}_{n} \rangle = \epsilon_{n} \langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \mathcal{O} \tilde{\Psi}_{n} \rangle$
Conclusion Conclusion	and gathered in one column
Perspectives	$\begin{vmatrix} c_1(g_1) \\ c_2(g_1) \end{vmatrix}$

 \vdots $c_m(g_1)$

 $3^{
m rd}$ International ABINIT Developer Workshop, Liège (Belgium), 29-31 January 2007

œ	Standard implementation
Theory	$ ilde{\Psi}_{\mathrm{n}}(\mathbf{r}) = \sum_{\mathbf{G}} \operatorname{cn}_{\mathrm{I}}(\mathbf{G}) \mathrm{e}^{\mathrm{i}(\mathbf{k}+\mathbf{G}).\mathbf{r}}$
PAW LOBPCG SCF	$[ilde{\mathbf{n}} + \hat{\mathbf{n}}](\mathbf{r}) \stackrel{\downarrow}{\longrightarrow} ext{and} ho_{\mathrm{ij}} \longleftarrow \{\mathrm{c}_{\mathrm{n}}(\mathbf{G}); \epsilon_{\mathrm{n}}\}$
Implementation Hypothesis Principles	$ \downarrow \qquad \uparrow \\ \mathbf{v}_{\mathbf{r}-}(\mathbf{r}) \text{and} \mathbf{v}_{-1}(\mathbf{r}) \qquad \tilde{\mathcal{H}} - \epsilon_{-}\mathcal{O} = 0 $
Results Benchmarks Optimisations	(i(k+Q) = i(k+Q) =
NC PAW Howto	$\langle \mathrm{e}^{\mathrm{i}(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \mathcal{H} \Psi_{\mathrm{n}} angle = \epsilon_{\mathrm{n}}\langle \mathrm{e}^{\mathrm{i}(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \mathcal{O} \Psi_{\mathrm{n}} angle$
Conclusion Conclusion	to give the final distribution.

$c_1(g_1)$	$c_1(g_{12})$	• • •	$c_1(g_p)$
$c_2(g_1)$	$c_2(g_{22})$		$c_2(g_p)$
:	:	·	÷
$c_{\rm m}({ m g_1})$			$c_{\rm m}(g_{\rm p})$

 $3^{\rm rd}$ International ABINIT Developer Workshop, Liège (Belgium), 29-31 January 2007

œ	Standard implementation
Theory NC PAW	$ ilde{\Psi}_{n}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}}$
lobpcg SCF	$[\tilde{\mathbf{n}} + \hat{\mathbf{n}}](\mathbf{r}) \stackrel{\star}{\downarrow} \text{ and } \rho_{ij} \longleftarrow \{c_{\mathbf{n}}(\mathbf{G}); \epsilon_{\mathbf{n}}\}$
Implementation Hypothesis Principles	$\begin{array}{c} \downarrow \\ \mathbf{v}_{\mathrm{loc}}(\mathbf{r}) \text{and} \mathbf{v}_{\mathrm{pl}}(\mathbf{r}) \qquad \qquad \tilde{\mathcal{H}} - \epsilon_{\mathrm{p}} \mathcal{O} = 0 \end{array}$
Results Benchmarks	
Optimisations NC PAW Howto	$\langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \tilde{\mathcal{H}} \tilde{\Psi}_n \rangle = \epsilon_n \langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \mathcal{O} \tilde{\Psi}_n \rangle$
Conclusion	We can apply a parallel 3dim-FFT on each line

$\begin{array}{c} c_1(g_1) \\ c_2(g_1) \end{array}$	$\begin{array}{c} c_1(g_2) \\ c_2(g_2) \end{array}$	 	$\begin{array}{c} c_1(g_p) \\ c_2(g_p) \end{array}$
:	÷	·	:
$c_m(g_1)$		•••	$c_m(g_p)$

3rd International ABINIT Developer Workshop, Liège (Belgium), 29-31 January 2007

F. Bottin

œ

Theoretical background

- The Norm-Conserving method (NC)
- The Projector Augmented-Wave method (PAW)
- Locally Optimal Block Preconditioned Conjugate Gradient method (LOBPCG)
- The self-consistent loop (SCF)

mplementation

- Hypothesis
- Principles

Results

- Benchmarks Optimisations NC PAW Howto Conclusion
- Conclusion

BandFFT parallelization

- Hypothesis Requirements
- Standard implementation

BandFFT Results

- Benchmarks
- Beyond the standard implementation Optimisations
- Norm-conserving Results
- PAW results
- How to use the BandFFT parallelization
- 4 Conclusion Perspectives
 - Conclusion
 - Perspectives

Benchmarks – Test cases

œ
000.

Theory	• Gold system with 108 atoms, 648 bands, 108^3 FFT grid, 1
NC PAW	k -point and $E_{cut}=24$ Ha.
LOBPCG SCF	• We stop the SCF for nstep=15
mplementation	
Hypothesis Principles	• A two-dimensional grid of processors with nproc=1, 4, 18, 36, 54, 108, 162 and 216
Results	51, 100, 102 and 210.
Benchmarks Optimisations NC	• Example: for nproc=108, we can choose $m \times p=108 \times 1, 54 \times 2, \overline{36 \times 3, 27} \times 4, 18 \times 6, 12 \times 9, 9 \times 12, 6 \times 18, 4 \times 27 \text{ and } 3 \times 36$
PAW Howto	• Test are performed on 2 supercomputers:

Supercomputer	Node	Interconnection
Tantale (CCRT)	4-procs AMD OPTERON 2.4 GHz	Infiniband
TERA-10 (CEA/DIF)	Novascale 16-procs Intel Itanium	Quadrics

The LAPACK and BLAS libraries

• In lobpcg: timing of zgemm for 54 processors

	zgemm/Total	Ratio
Standard libraries	80000/150000	53%
Optimized libraries	10000/100000	10%

Beyond the standard implementation - Optimisations $\tilde{\Psi}_{n}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}$

LOBPCG

SCF

Implementation

Hypothesis

Principles

Results

Benchmarks

Optimisations

NC PAW

Howto

Conclusion

Conclusion

Perspectives

$(\mathbf{r}) = \sum_{\mathbf{G}} c_n(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}$	
$[\tilde{\mathbf{n}} + \hat{\mathbf{n}}](\mathbf{r}) \downarrow$ and ρ_{ij}	$\longleftarrow \{c_n(\mathbf{G}); \epsilon_n\} \\ \uparrow$
$egin{array}{cc} & \downarrow & \ \mathrm{v}_{\mathrm{loc}}(\mathbf{r}) & \mathrm{and} & \mathrm{v}_{\mathrm{nl}}(\mathbf{r}) & \ & \ & \ & \ & \ & \ & \ & \ & \ & $	$ \tilde{\mathcal{H}} - \epsilon_{n}\mathcal{O} = 0$ \uparrow
$\langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \tilde{\mathcal{H}} \tilde{\Psi}_n \rangle = \epsilon_n \langle e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \tilde{\mathcal{H}} \tilde{\Psi}_n \rangle$	$e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}} \mathcal{O} ilde{\Psi}_{\mathrm{n}} angle$

The SCALAPACK library

• In vtowfk: timing of subdiago for 216 processors.

	648 bands	1296 bands
Standard implementation	8%	23%
SCALAPACK library	3%	2%

Generalisation of the transposition principle

- Before each call to nonlop we transpose the distribution and use a "FFT distribution".
- Minimize the number of collective communications.
- Permits to avoid bad load balancing.

Norm-conserving Results



- 108 atoms, 648 bands, 108^3 FFT grid, 1 k-point and $E_{cut}=24$ Ha.



In sequential: 90 000 sec. (with 90% in lobpcg).Linear scaling up to 100 for ABINIT and 200 for lobpcg.

PAW results (with 1 projector per angular momentum)





In sequential: 60 000 sec. (with 90% in lobpcg).
We loose efficiency between lobpcg, vtowfk and ABINIT.

Number of processors

Number of processors

PAW results (with 2 projectors per angular momentum)



SCE

Implomontatio

Hypothesi

Principles

Results

Benchmarks

Optimisatio

NC

PAW

HOWI

Conclusio

Conclusion

Perspectives





In sequential: 75 000 sec. (with 90% in lobpcg).Due to: nonlop_ylm, the double grid and spherical terms.



Compilation and input variables

Theory
NC
PAW
LOBPCG
SCF
mplementat
Hypothesis
Principles
Results
Benchmarks

Howto

- Compile with the -DMPL_FFT and -DSCALAPACK preprocessing flags.
- Link to the optimized librairies (LAPACK, BLAS and SCALAPACK) on your plateform.
- npband=m and npfft=p for the band- and FFT-processors.
- wfoptalg=4: the lobpcg 1 method
- fft_opt_lob=2: the generalisation of the transposition principle.
- fftalg=401: the 3dim parallel FFT of Goedecker *et al.* 2 .

²S. Goedecker, M. Boulet and T. Deutsch An efficient 3-dim FFT for plane wave electronic structure calculations on massively parallel machines composed of multiprocessor nodes, Comput. Phys. Comm. **154**, 105 (2003).

Restrictions: ground state only calculations

- œ
- Theory NC PAW Losecco SCF Implementation Hypothesis Principles Results Benchmarks Optimisations NC PAW Howo Conclusion Conclusion
- iprcch=0. Other values are allowed for testing purposes but the parallelization is then very poor. For large systems the correction of forces (iprcch=1) and the prediction of the density at each MD step (iprcch=2) lead to memory overflow and swap.
- istwfk=nkpt*1. The time-reversal symetry is not yet compatible with the bandFFT parallelization.
- nsppol=nspden=1. No spin-polarized (or antiferromagnetic) calculations.
- useylm=0. No PAW calculations.
- $\bullet\,$ is cf> 0. Non self-consistent calculations are not tested.
- $\bullet~{\rm I/O}$ is not yet available.

Theoretical background BandFFT parallelization Hypothesis – Requirements How to use the BandFFT parallelization

- Conclusion Perspectives
 Conclusion
- Perspectives

Conclusion

Conclusion

- œ
- Theory NC
- PAW
- LOBPO
- SCF
- mplementation
- Hypothesis
- Principles
- Results
- Benchmarks
- Optimisations
- NC
- PAW
- Howto
- Conclusion

Conclusion

Perspectives

- ABINIT scales lineary up to 100 processors in NC and PAW.
- lobpcg is responsible for the superlinear behaviour and scales perfectly up to 200 processors.
- The BandFFT parallelization have to be optimized in the framework of PAW calculations.
- TODO: Retrieve all the functionalities of ABINIT (spin polarization, NSCF calculations, I/O ...)



Perspectives

Perspectives: the triple nkG parallelization (in 5.4.x)

108 atoms, 648 bands, 108^3 FFT grid, 10 k-point and $E_{cut}=24$ Ha.



Linear scaling of ABINIT up to 1000 processors.