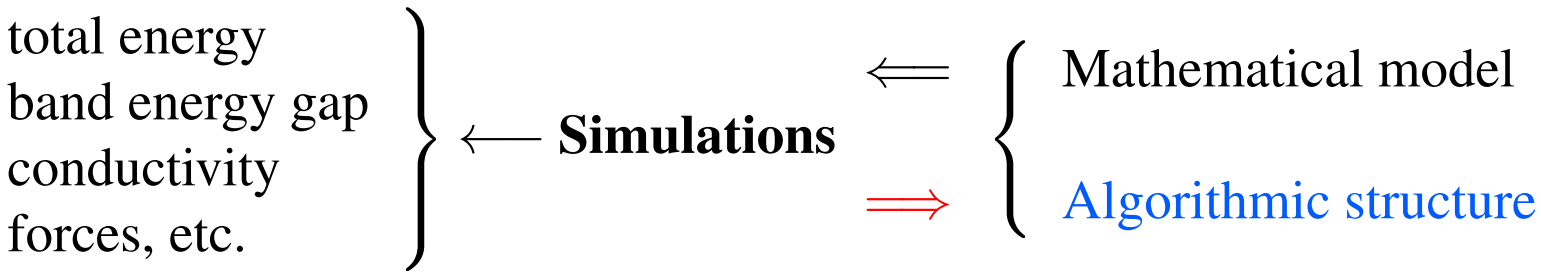
JÜLICH
FORSCHUNGSZENTRUM

Block Iterative Eigensolvers for Sequences of Dense Correlated Eigenvalue Problems

Birkbeck University, London, June the 29th 2012 | Edoardo Di Napoli

Motivation and Goals

Reverse Simulation



Goal

Increasing the **performance** of large legacy codes by exploiting physical information extracted from the simulations that can be used to speed-up the algorithms used in such codes

Outline

Stating the problem: how sequences of generalized eigenproblems arise in all-electron computations

Eigenvectors angle evolution

ALGORITHM $\Leftarrow$ SIM – Exploiting eigenvector collinearity: block iterative eigensolvers

Outline

Stating the problem: how sequences of generalized eigenproblems arise in all-electron computations

Eigenvectors angle evolution

ALGORITHM $\Leftarrow$ SIM – Exploiting eigenvector collinearity: block iterative eigensolvers

The Foundations

Investigative framework

Quantum Mechanics and its ingredients

- $H = -\frac{\hbar^2}{2m}\sum_{i=1}^{n}\nabla_i^2 - \sum_{i=1}^{n}\sum_{\alpha}\frac{Z_\alpha}{|x_i - a_\alpha|} + \sum_{i<j}\frac{1}{|x_i - x_j|}$ Hamiltonian
- $\Phi(x_1; s_1, x_2; s_2, \ldots, x_n; s_n)$ Wavefunction

$\Phi : \left(\mathbb{R}^3 \times \{\pm\frac{1}{2}\}\right)^n \longrightarrow \mathbb{R}$ high-dimensional anti-symmetric function – describes the orbitals of atoms and molecules. In the *Born-Oppenheimer* approximation, it is the solution of the

Electronic Schrödinger Equation

$$H\Phi(x_1; s_1, x_2; s_2, \ldots, x_n; s_n) = \mathcal{E}\Phi(x_1; s_1, x_2; s_2, \ldots, x_n; s_n)$$

Density Functional Theory (DFT)

1. $\Phi(x_1;s_1,x_2;s_2,\ldots,x_n;s_n) \Longrightarrow \Lambda_{i,a}\phi_a(x_i;s_i)$
2. **density of states** $n(\mathbf{r}) = \sum_a |\phi_a(\mathbf{r})|^2$
3. In the Schrödinger equation the exact Coulomb interaction is substituted with an effective potential $V_0(\mathbf{r}) = V_{\mathrm{I}}(\mathbf{r}) + V_{\mathrm{H}}(\mathbf{r}) + V_{\mathrm{xc}}(\mathbf{r})$

Hohenberg-Kohn theorem

- $\exists$ one-to-one correspondence $n(\mathbf{r}) \leftrightarrow V_0(\mathbf{r}) \quad \Longrightarrow \quad V_0(\mathbf{r}) = V_0(\mathbf{r})[n]$
- $\exists!$ a functional $E[n]$: $\quad E_0 = \min_n E[n]$

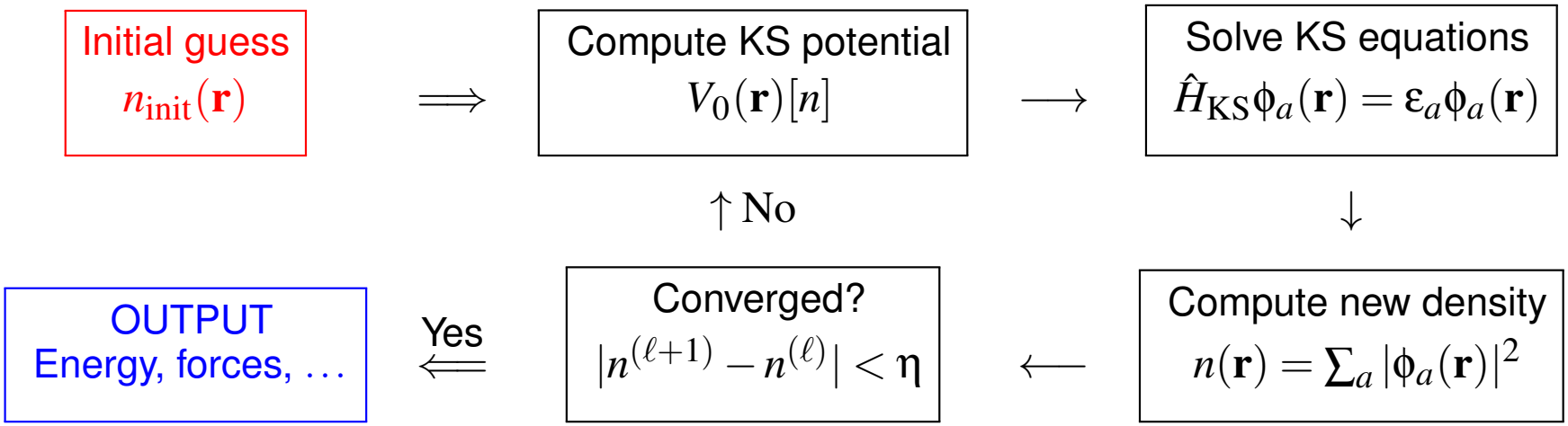
The high-dimensional Schrödinger equation translates into a set of coupled non-linear low-dimensional self-consistent Kohn-Sham (KS) equation

$$\forall\, a \quad \text{solve} \quad \hat{H}_{\mathrm{KS}}\phi_a(\mathbf{r}) = \left(-\frac{\hbar^2}{2m}\nabla^2 + V_0(\mathbf{r})\right)\phi_a(\mathbf{r}) = \varepsilon_a\phi_a(\mathbf{r})$$

Kohn-Sham scheme and DFT

Self-consistent cycle

Typically this set of equations is solved using an iterative self-consistent cycle



In practice this iterative cycle is much more computationally challenging and requires some form of broadly defined discretization

Generalized eigenvalue problems $Ax = \lambda Bx$

A common way of discretizing the KS equations is to expand the wave functions $\phi_a(\mathbf{r})$ on a basis set

$$\phi_a(\mathbf{r}) \longrightarrow \phi_{\mathbf{k},\nu}(\mathbf{r}) = \sum_{|\mathbf{G}+\mathbf{k}|\leq \mathbf{K}_{max}} c^{\mathbf{G}}_{\mathbf{k},\nu}\psi_{\mathbf{G}}(\mathbf{k},\mathbf{r})$$

This expansion is then inserted in the KS equations

$$\psi^*_{\mathbf{G}}(\mathbf{k},\mathbf{r})\sum_{\mathbf{G}'}\hat{H}_{\mathrm{KS}}\ c^{\mathbf{G}'}_{\mathbf{k},\nu}\ \psi_{\mathbf{G}'}(\mathbf{k},\mathbf{r}) = \lambda_{\mathbf{k}\nu}\ \psi^*_{\mathbf{G}}(\mathbf{k},\mathbf{r})\sum_{\mathbf{G}'}c^{\mathbf{G}'}_{\mathbf{k},\nu}\ \psi_{\mathbf{G}'}(\mathbf{k},\mathbf{r}),$$

and, by defining the matrix entries for the left and right hand side respectively as Hamiltonian $A_{\mathbf{k}}$ and overlap matrices $B_{\mathbf{k}}$,

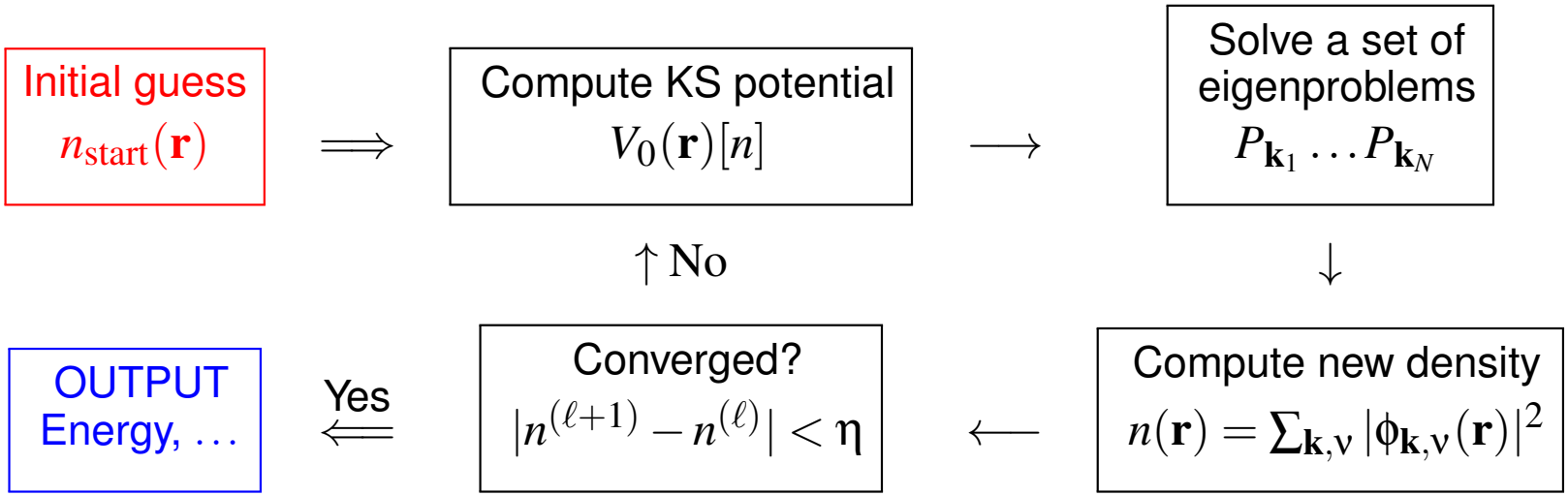
$$[A_{\mathbf{k}}\ B_{\mathbf{k}}]_{\mathbf{G}\mathbf{G}'} = \sum_{\alpha}\int d\mathbf{r}\ \psi^*_{\mathbf{G}}(\mathbf{k},\mathbf{r})\left[\hat{H}_{\mathbf{KS}}\ \hat{\mathbb{1}}\right]\psi_{\mathbf{G}'}(\mathbf{k},\mathbf{r})$$

one arrives at generalized eigenvalue equations parametrized by **k**

$$P_{\mathbf{k}}: \quad \sum_{\mathbf{G}'}(A_{\mathbf{k}})_{\mathbf{G}\mathbf{G}'}\ c^{\mathbf{G}'}_{\mathbf{k}\nu} = \lambda_{\mathbf{k}\nu}\sum_{\mathbf{G}'}(B_{\mathbf{k}})_{\mathbf{G}\mathbf{G}'}\ c^{\mathbf{G}'}_{\mathbf{k}\nu}. \quad \equiv \quad A_{\mathbf{k}}x_i = \lambda_i B_{\mathbf{k}}x_i.$$

Discretized Kohn-Sham scheme

Self-consistent cycle



Observations:

1. A and B are respectively hermitian and hermitian positive definite
2. eigenproblems across **k** index have different size and we consider them independent from each other (for the moment)
3. eigenvectors of problems of same **k** are seemingly uncorrelated across iterations i
4. **k**= 1:10-100 ; i = 1:20-50

Outline

Stating the problem: how sequences of generalized eigenproblems arise in all-electron computations

Eigenvectors angle evolution

ALGORITHM ⇐ SIM – Exploiting eigenvector collinearity: block iterative eigensolvers

Sequences of eigs: an ALGORITHM $\Leftarrow$ SIM case

Sequences of eigenproblems

- Consider the set of generalized eigenproblems $P^{(1)} \ldots P^{(\ell)} P^{(\ell+1)} \ldots P^{(N)} \neq (P)^N$
- Could this sequence $\left\{ P^{(\ell)} \right\}$ of eigenproblems evolve following a convergence pattern in line with the convergence of $n(\mathbf{r})$?

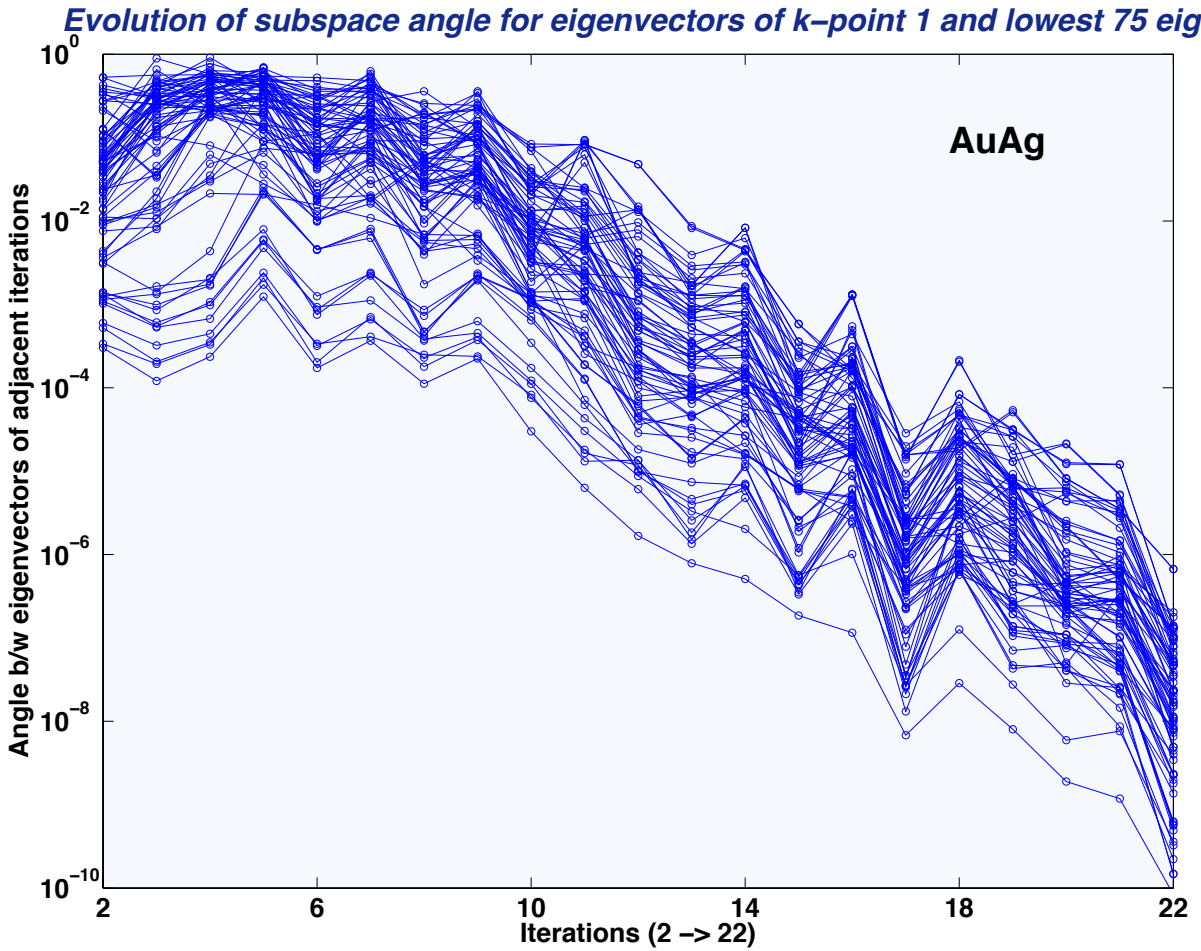
Numerical study:

- studied the evolutions of the angles b/w eigenvectors of successive iterations
- developed a method that establishes systematically a one-to-one correspondence b/w eigenvectors
- collected data on eigenvectors deviation angles
  1. analyzed deviation angles at fixed λ for all **k**s
  2. analyzed deviation angles at fixed **k** for all λs below Fermi Energy

Angle evolution

fixed $\mathbf{k}$

Example: a metallic compound at fixed $\mathbf{k}$



Correlation and its exploitation

- $\exists$ correlation between successive eigenvectors $x^{(\ell-1)}$ and $x^{(\ell)}$
- Angles decrease monotonically with some oscillation
- Majority of angles are small after the first few iterations

Note: Mathematical model $\not\Rightarrow$ correlation.
Correlation $\Leftarrow$ **systematic analysis** of the simulation.

Algorithm $\Leftarrow$ Sim

The stage is favorable to an iterative eigensolver where the eigenvectors of $P^{(\ell-1)}$ are fed to the solve $P^{(\ell)}$

Next stages of the investigation:

1. Establish which eigensolvers can exploit the evolution (Implicit Restarted Arnoldi, Krylov-Schur, Subspace Iteration, Davidson-like, etc.)
2. Investigate if approximate eigenvectors can speed-up iterative solvers
3. Understand if iterative methods be competitive with direct methods for **dense** problems

Outline

Stating the problem: how sequences of generalized eigenproblems arise in all-electron computations

Eigenvectors angle evolution

ALGORITHM $\Leftarrow$ SIM – Exploiting eigenvector collinearity: block iterative eigensolvers

Block Iterative Eigensolvers

Two essential properties the iterative algorithms have to comply with:

1. the ability to receive as input a **sizable set** of approximate eigenvectors;
2. the capacity to **solve simultaneously** for a substantial portion of eigenpairs.

Block iterative methods constitutes the natural choice:

- they accept a variable set of multiple starting vectors;
- these methods have a faster convergence rate and avoid stalling when facing small clusters of eigenvalues;
- when augmented with polynomial accelerators their performance is further improved.

ALGORITHMS: Block Krylov-Schur – Block Chebyshev-Davidson – Chebyshev Subspace Iteration – LOBPCG

Experimental tests setup

Matrix sizes: 2,000 ÷ 5,700
Num of fixed-point iterations: 15 ÷ 30
Num of k-points: 6 ÷ 27

B ill-conditioned

B is in general almost singular.

Examples: $\text{size}(A) = 50 \rightarrow \kappa(A) \approx 10^4$ $\quad$ $\text{size}(A) = 500 \rightarrow \kappa(A) \approx 10^7$

We used the standard form for the problem

$Ax = \lambda Bx \quad \longrightarrow \quad A'y = \lambda y$ with $A' = L^{-1}AL^{-T}$ and $y = L^T x$

Numerical Study:

- Approx. vs Random solutions against Iteration Index
- Approx. vs Random solutions against Spectrum Fraction

Numerical tests were performed using Matlab version R2011b (7.13.0.564) running on an Intel i7 CPU with 8 cores at 2.93 GHz. Four cores and 8 Gb of RAM were fully dedicated to computations

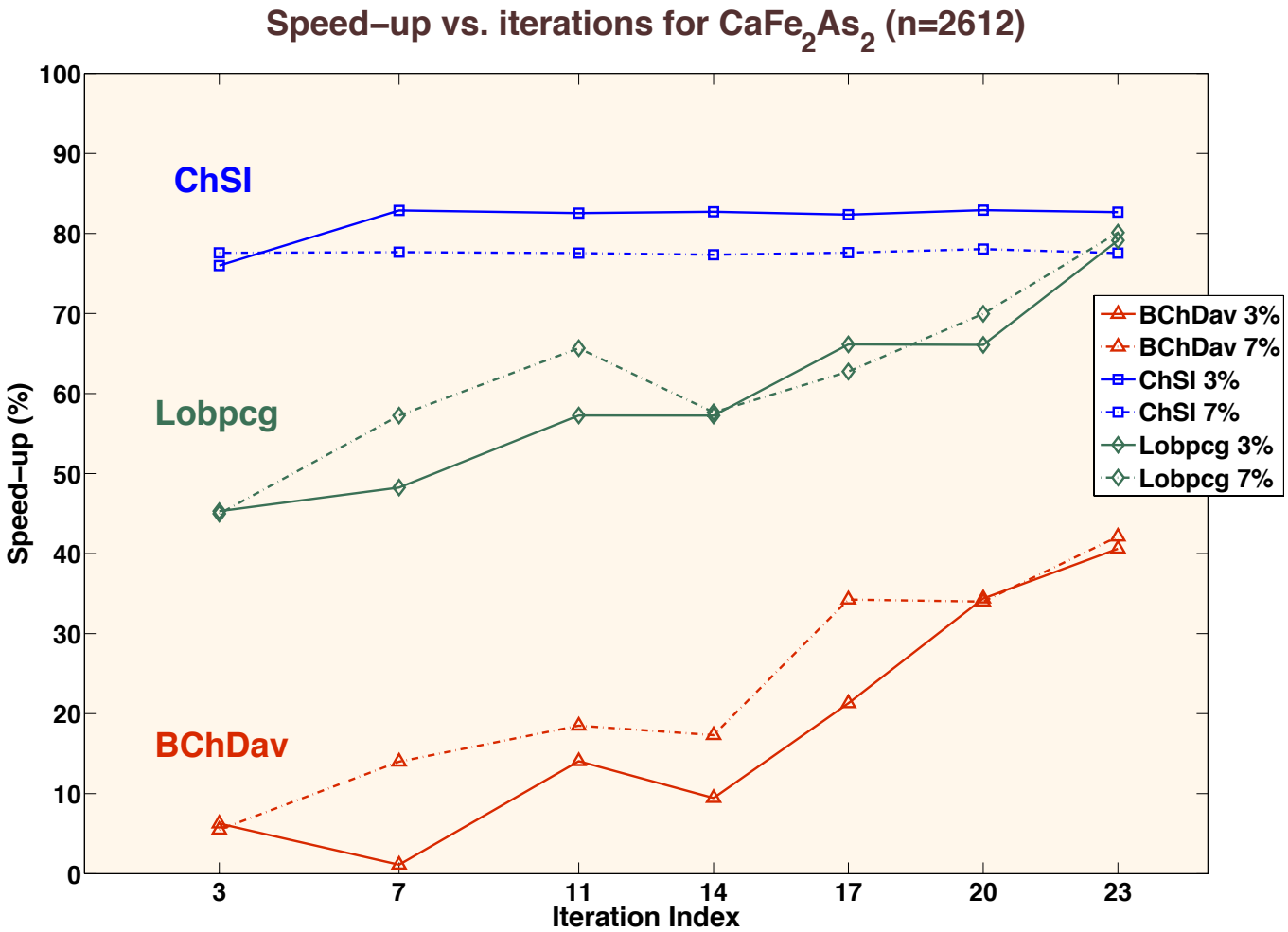


Figure: *Comparison between the 3 most effective block iterative eigensolvers for $CaFe_2As_2$ with respect to the outer-iteration index*

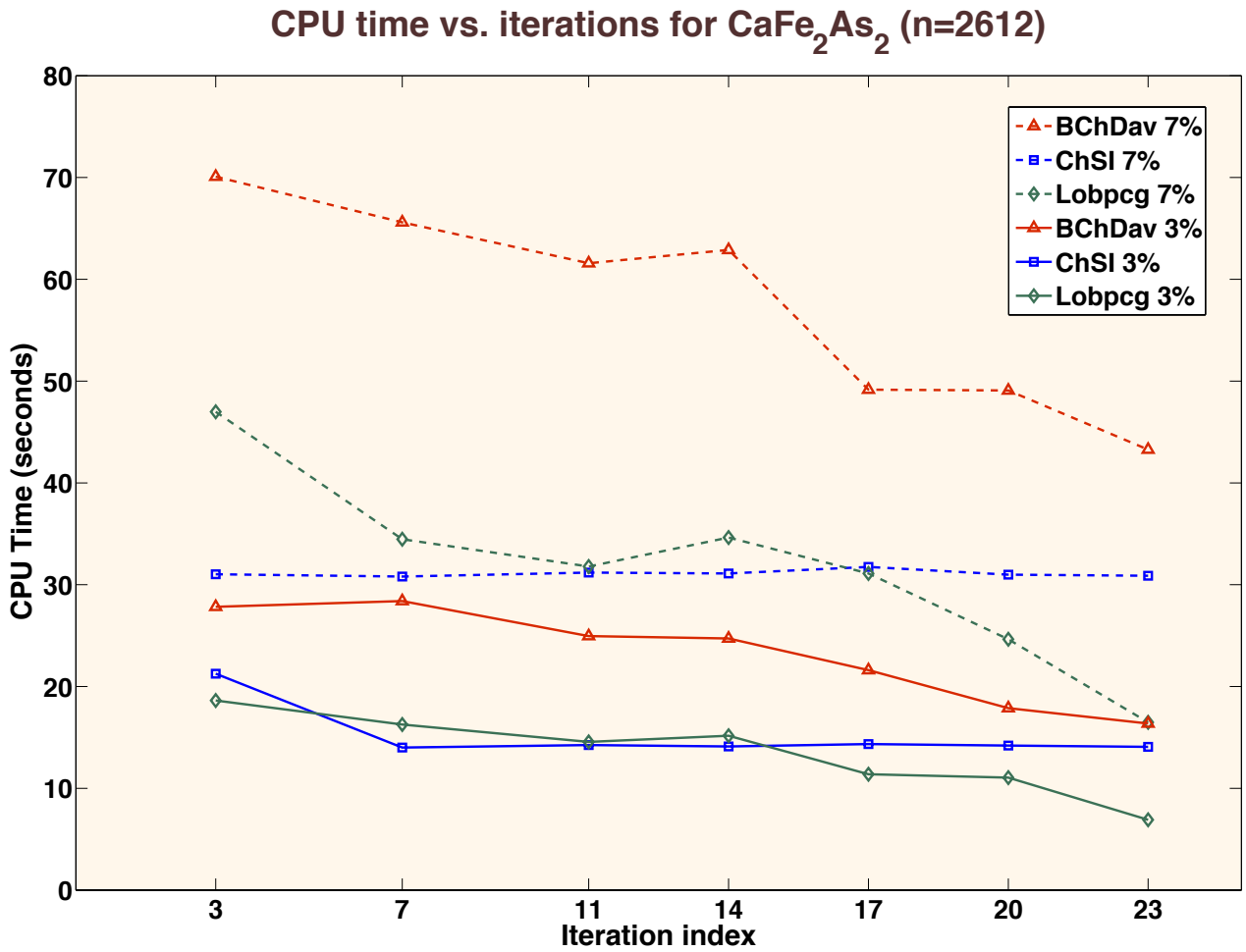


Figure: *Comparison between the 3 most effective block iterative eigensolvers for $CaFe_2As_2$ with respect to the outer-iteration index*

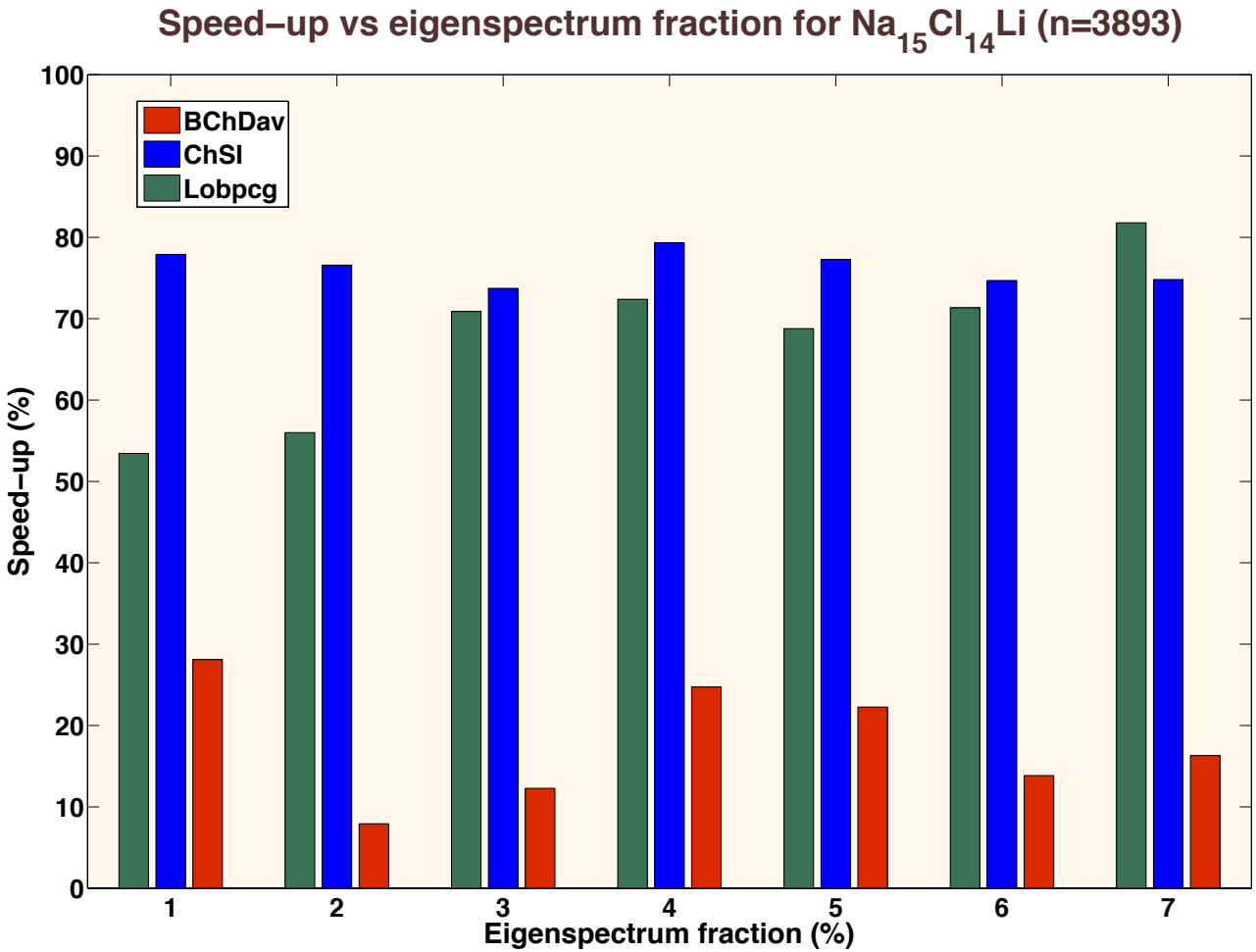


Figure: *Comparison between the 3 most effective block iterative eigensolvers for $Na_{15}Cl_{14}Li$ with respect to eigenspectrum fraction*

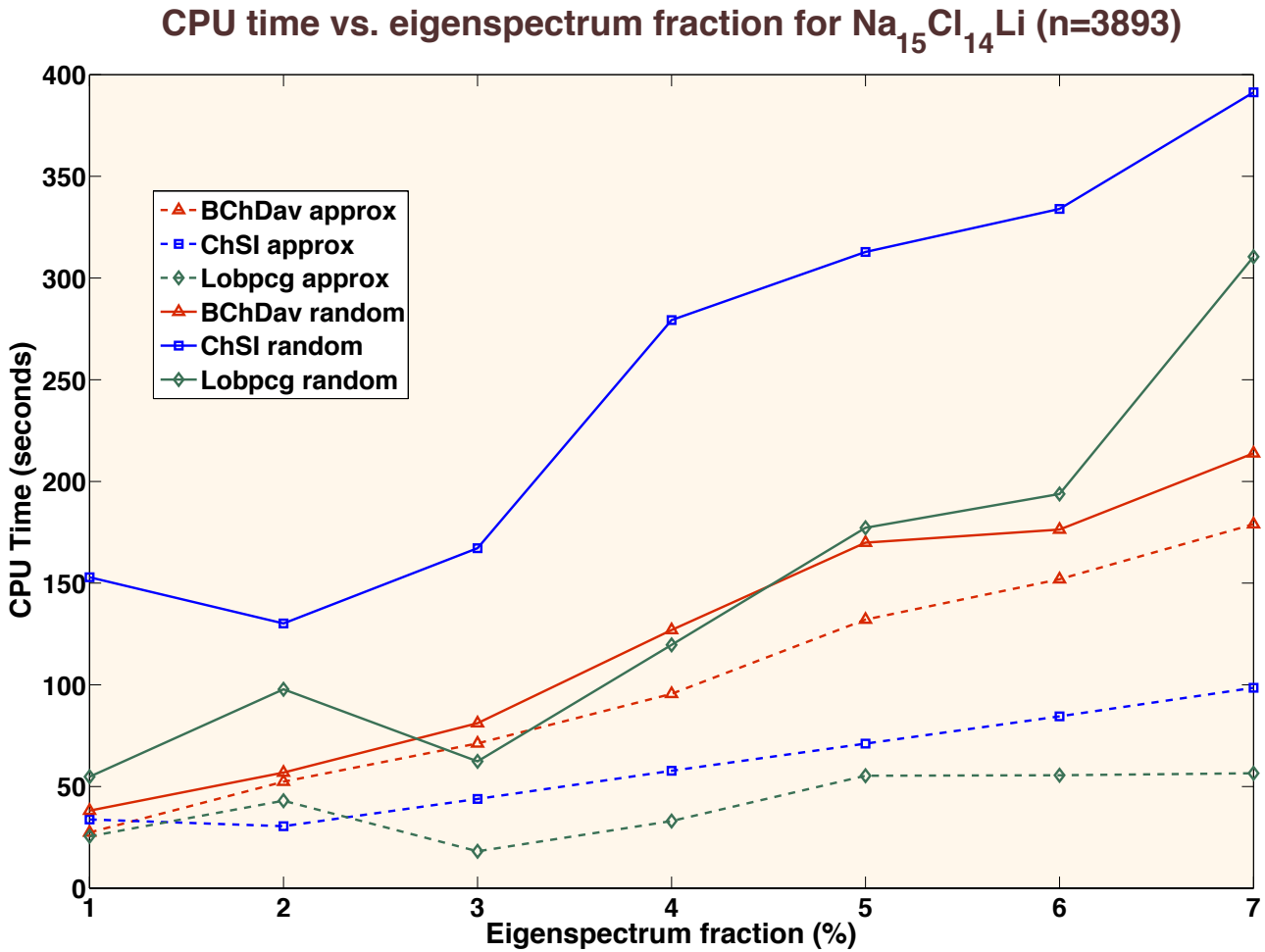


Figure: *Comparison between the 3 most effective block iterative eigensolvers for $Na_{15}Cl_{14}Li$ with respect to eigenspectrum fraction*

In summary

All methods allow the possibility of “seeding” the eigensolver with **as many** eigenvectors as needed

1. **Chebyshev Subspace Iteration and LOBPCG methods:**
   - Experience between 55% and 85% speed-up
   - The speed-up has little dependence on the eigesprectrum fraction or the iteration index
   - Their efficiency highly depend on the efficiency of Matrix-Matrix multiplication sub-routines $\Rightarrow$ GPUs

2. **Block Chebyshev-Davidson method:**
   - Experiences a speed-up up to 45 %
   - The increased “feeding” capacity allows speed-ups even when a substantial fraction of the spectrum is sought after
   - Evolution of the sequence implies an increase in speed-ups towards the end of the sequence

Conclusions and future work

- Feeding eigenvectors of $P^{(\ell-1)}$ speed-ups the iterative solver for $P^{(\ell)}$;
- The algorithmic structure of FLAPW could be re-thought to take into consideration our results $\rightarrow$ **"reverse simulation" can substantially influence the computational paradigm of an application**;
- Conduct a more low-level analysis of speed-up on variants of the investigated algorithms
- Planning the generalization of a block Chebyshev-filtered Subspace Iteration eigensolver with a general interface for the use of approximate solutions;
- Analysis on the structure of the entries of A and B across adjacent iteration seems to suggest an exploitation of low-rank updates for sequences of eigenproblems;
- Correlation may open the road to a reduction of eigenproblem complexity through a sequence of reductions to tridiagonal form.

References

1 EDN
Block Iterative Eigensolvers for Sequences of Correlated Eigenvalue Problems
Submitted to SIAM Journal on Scientific Computing [arXiv:1206.3768]

2 EDN, P. Bientinesi, and S. Blügel,
Correlation in sequences of generalized eigenproblems arising in Density Functional Theory,
Comp. Phys. Comm. 183 (2012), pp. 1674-1682, [arXiv:1108.2594].