The Algorithm of Multiple Relatively Robust Representations for Multi-core Processors

Matthias Petschow and **Paolo Bientinesi**

AICES, RWTH Aachen
pauldj@aices.rwth-aachen.de

PARA 2010: State of the Art in Scientific and Parallel Computing
June 6–9, 2010
Reykjavik, Iceland

$AX = X\Lambda$

- Input: $A \in \mathcal{C}^{n \times n}$, $A^H = A$; #eigenpairs: $1 \leq k \leq n$
- Output: $X \in \mathcal{C}^{n \times k}$ eigenvectors
  $\Lambda \in \mathcal{R}^{k \times k}$ eigenvalues

Problem

$AX = X\Lambda$

- Input: $A \in \mathcal{C}^{n \times n}$, $A^H = A$; #eigenpairs: $1 \leq k \leq n$
- Output: $X \in \mathcal{C}^{n \times k}$ eigenvectors
  $\Lambda \in \mathcal{R}^{k \times k}$ eigenvalues

Approach

- $T = Q^H A Q$ Reduction to tridiagonal form $O(n^3)$
- $TZ = Z\Lambda$ Tridiagonal eigenproblem $O(kn)$–$O(n^3)$
- $X = QZ$ Backtransformation $O(kn^2)$

Problem

RWTH AACHEN UNIVERSITY

$AX = X\Lambda$

- Input: $A \in \mathcal{C}^{n\times n}$, $A^H = A$; #eigenpairs: $1 \le k \le n$
- Output: $X \in \mathcal{C}^{n\times k}$ eigenvectors
  $\Lambda \in \mathcal{R}^{k\times k}$ eigenvalues

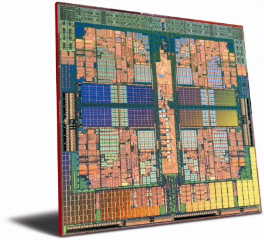
Approach

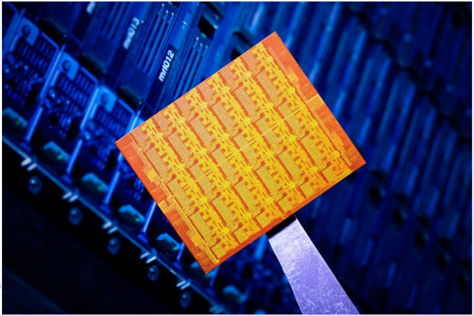
- $T = Q^H A Q$ Reduction to tridiagonal form $O(n^3)$
- $TZ = Z\Lambda$ Tridiagonal eigenproblem $O(kn)$–$O(n^3)$
- $X = QZ$ Backtransformation $O(kn^2)$

Algorithms

- Inverse Iteration (1958): subsets $O(kn^2)$
- QR (1961): high-accuracy $O(n^3)$
- Divide & Conquer (1981): parallel, BLAS3 $O(n^3)$
- MRRR (1997): subsets, no re-orth. $\mathbf{O(kn)}$

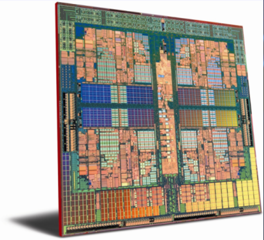
AI ces

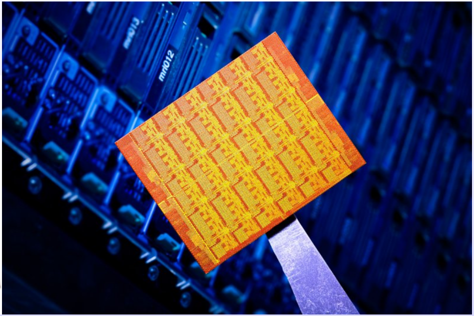




Multi-cores invasion: 499/500 entries of the Top 500

4 cores, 8 cores, . . . 24 cores, . . .

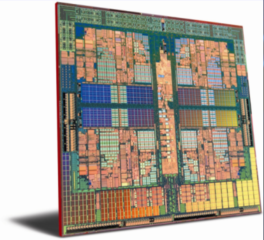


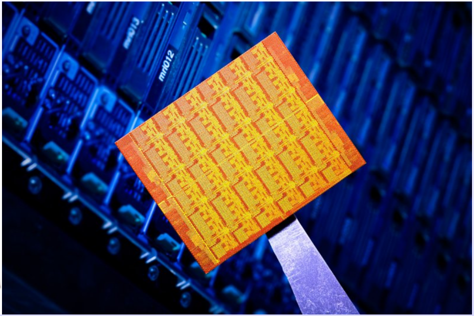


Multi-cores invasion: 499/500 entries of the Top 500

4 cores, 8 cores, . . . 24 cores, . . .

- More parallelism than we know what to do with?
- Is multi-threaded BLAS the solution for LA libs?
- Linear solvers $\neq$ Eigensolvers



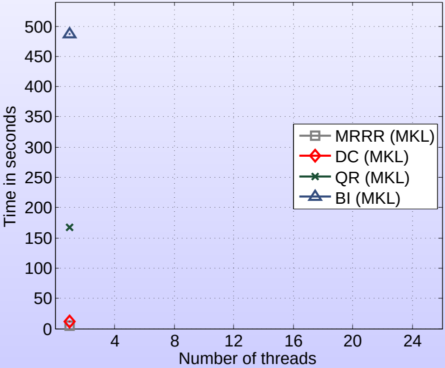


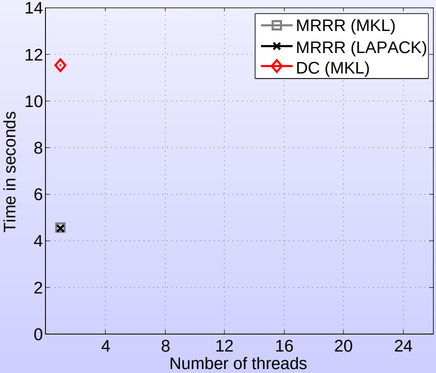
Multi-cores invasion: 499/500 entries of the Top 500

4 cores, 8 cores, . . . 24 cores, . . .

- More parallelism than we know what to do with?
- Is multi-threaded BLAS the solution for LA libs?
- Linear solvers $\neq$ Eigensolvers

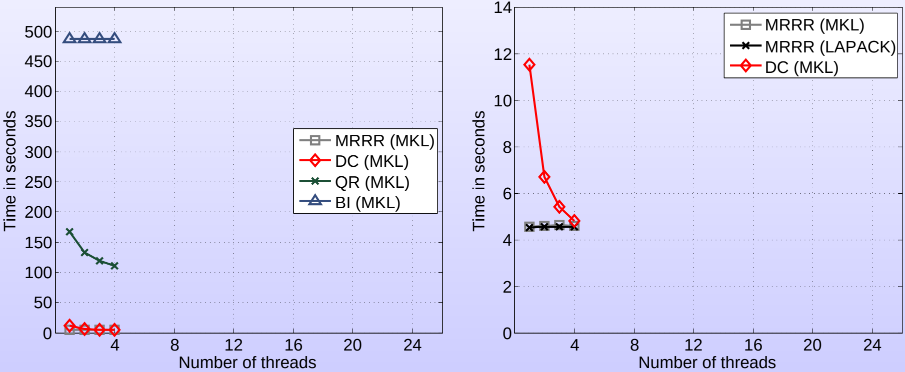
Multi-threaded BLAS?





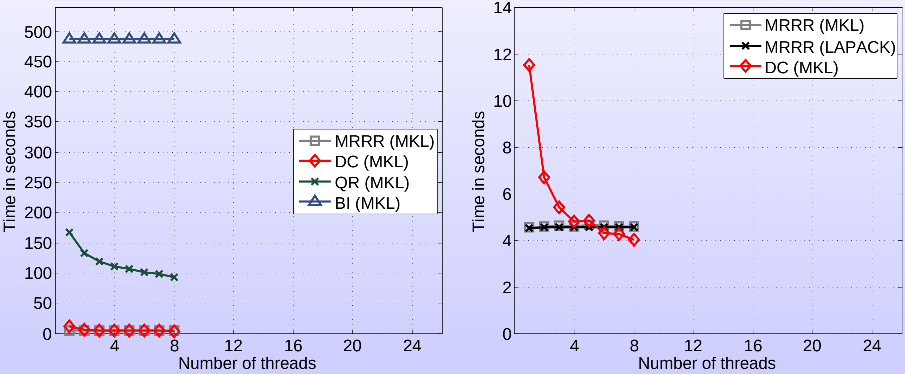
Tridiagonal eigensolver, matrix size=4289, from DFT.

Multi-threaded BLAS?



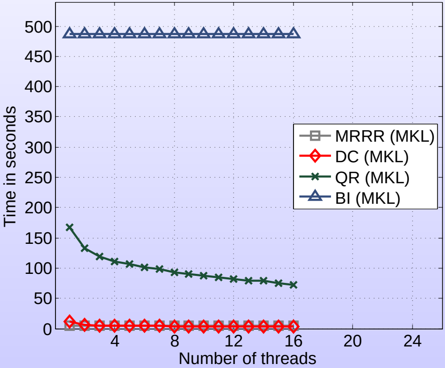
Tridiagonal eigensolver, matrix size=4289, from DFT.

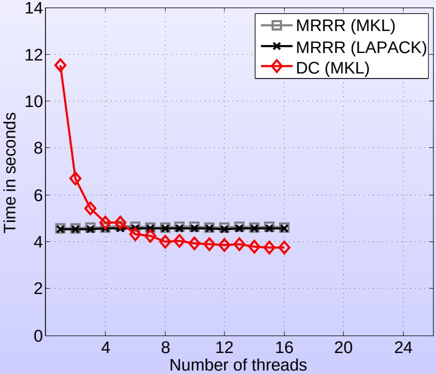
Multi-threaded BLAS?



Tridiagonal eigensolver, matrix size=4289, from DFT.

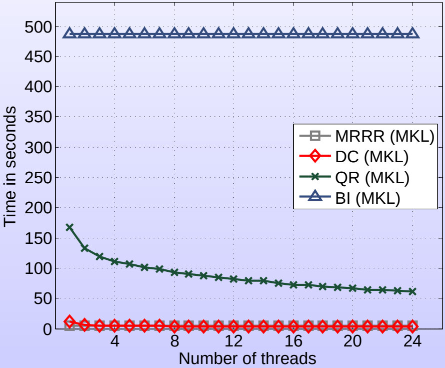
Multi-threaded BLAS?

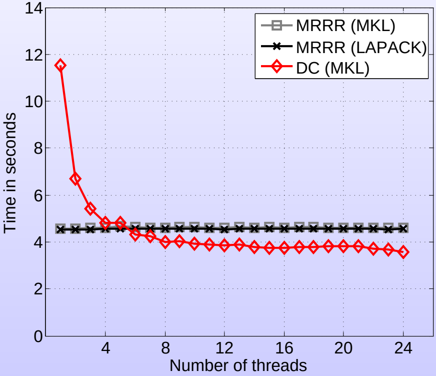




Tridiagonal eigensolver, matrix size=4289, from DFT.

Multi-threaded BLAS?





Tridiagonal eigensolver, matrix size=4289, from DFT.

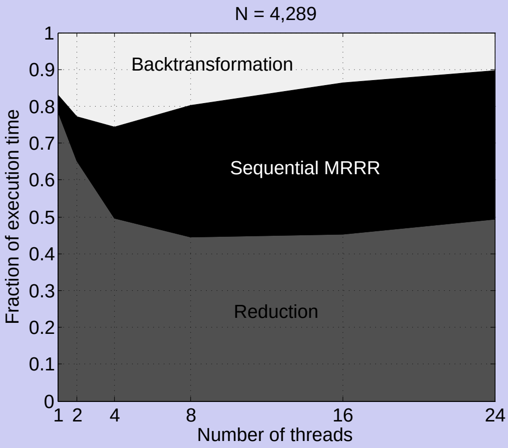
More motivation?

. . . it's $O(n^2)$ anyway

More motivation?

RWTH AACHEN UNIVERSITY

. . . it's $O(n^2)$ anyway



If not properly parallelized, even $O(n^2)$ dominates!

Multiple Relatively Robust Representations

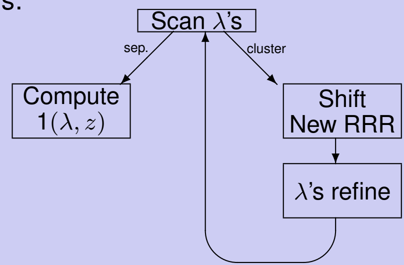
- first stable algorithm to compute k eigenpairs in $O(nk)$ ops

Multiple Relatively Robust Representations

- first stable algorithm to compute k eigenpairs in $O(nk)$ ops
- no reorthogonalization

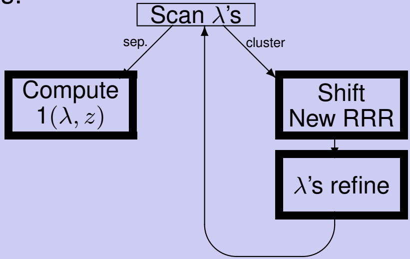
Multiple Relatively Robust Representations

- first stable algorithm to compute k eigenpairs in $O(nk)$ ops
- no reorthogonalization
- 1) eigenvalues $\rightarrow$ 2) eigenvectors + eigenvalues
- eigenvalues: *Bisection* or *dqds*
- eigenvectors:

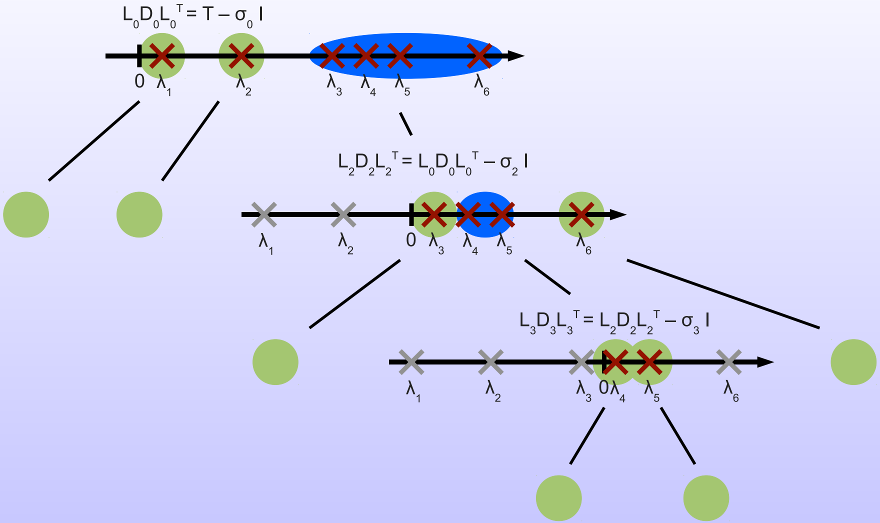


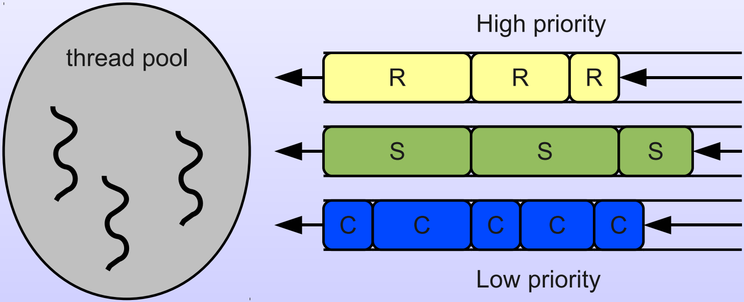
Multiple Relatively Robust Representations

- first stable algorithm to compute k eigenpairs in $O(nk)$ ops
- no reorthogonalization
- 1) eigenvalues $\rightarrow$ 2) eigenvectors + eigenvalues
- eigenvalues: *Bisection* or *dqds*
- eigenvectors:



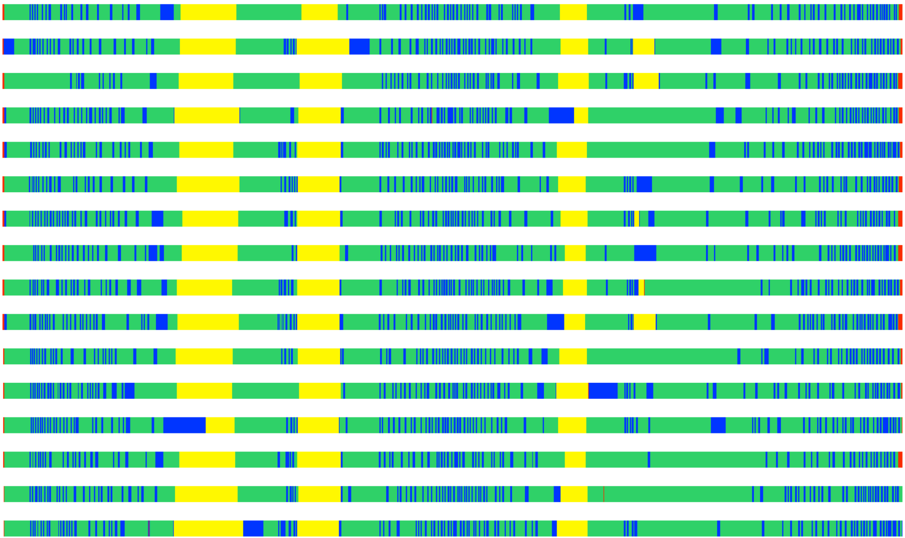
Representation Tree

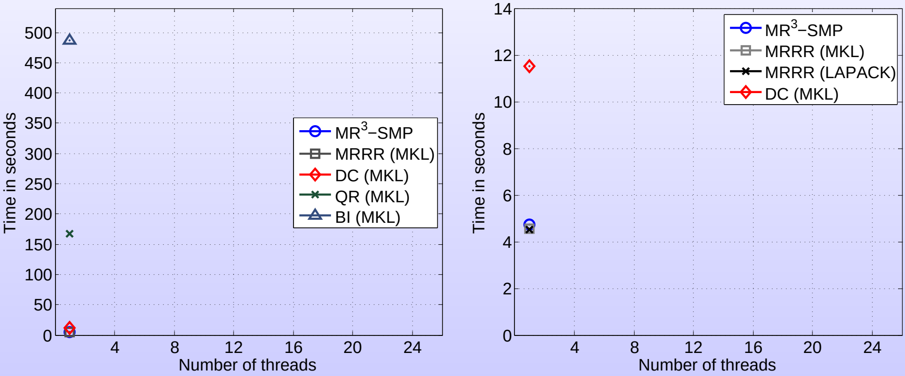



thread pool
High priority
R
R
R
S
S
S
C
C
C
C
C
Low priority

Example trace: 16 cores—eigenvectors

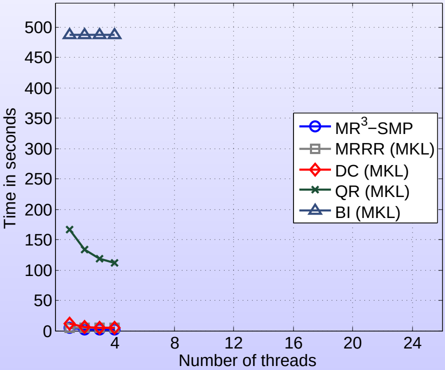
Matrix size: 12387   Execution time: 3.3s   Sequential: 49.3s (LAPACK)

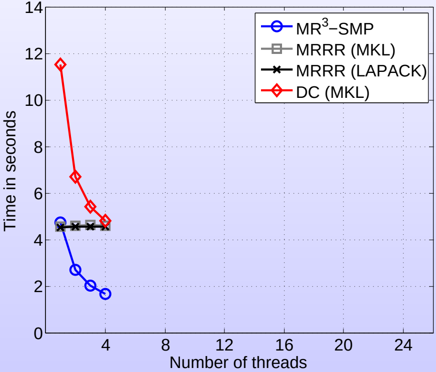




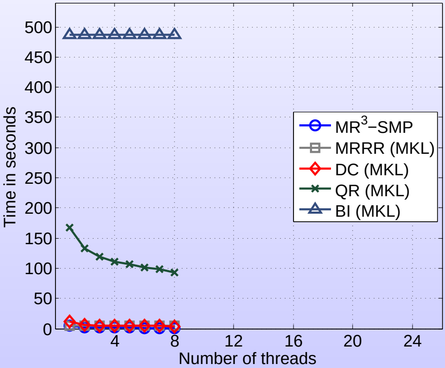
Matrix size: 4289.

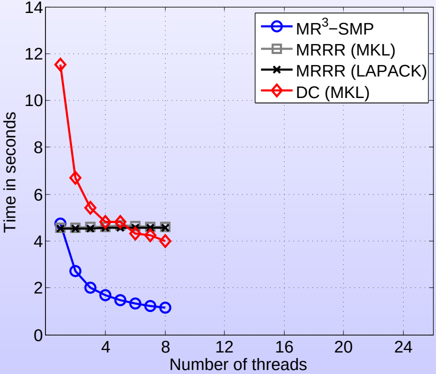
Tridiagonal stage





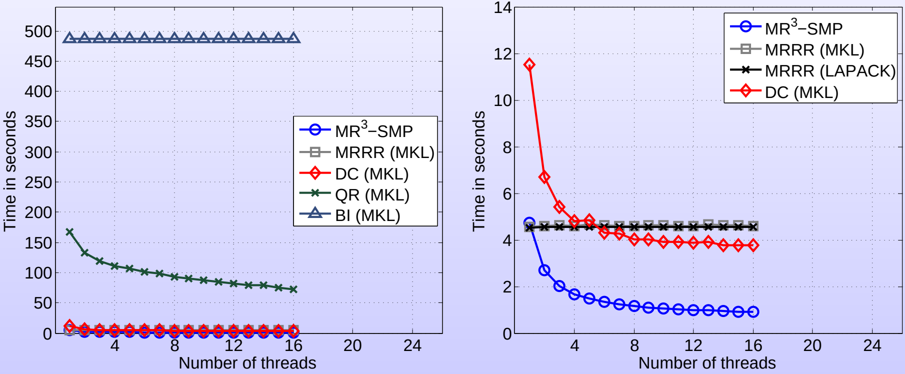
Matrix size: 4289.



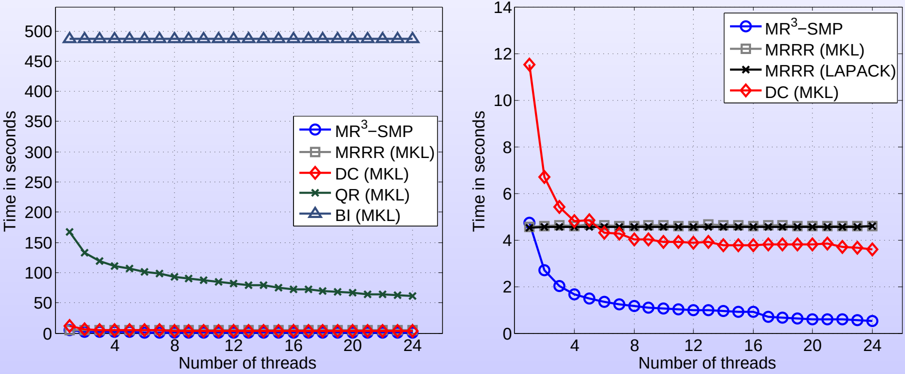


Matrix size: 4289.

Tridiagonal stage



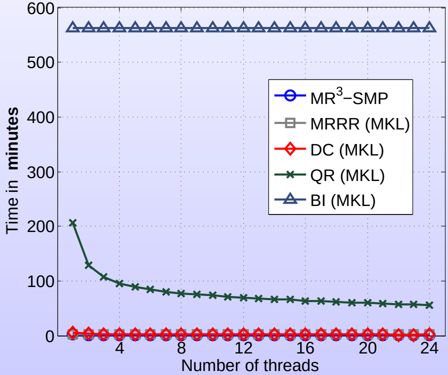
Matrix size: 4289.

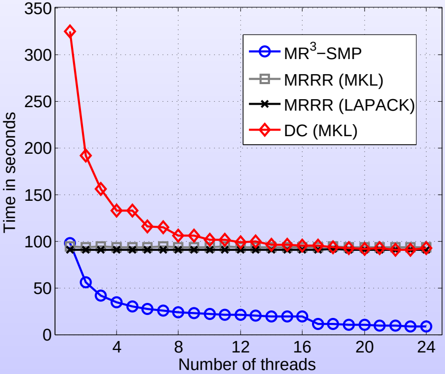


Matrix size: 4289.

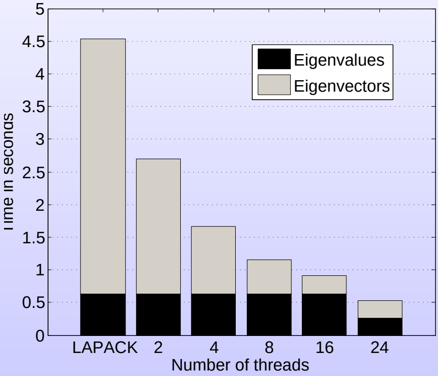
RWTH AACHEN UNIVERSITY

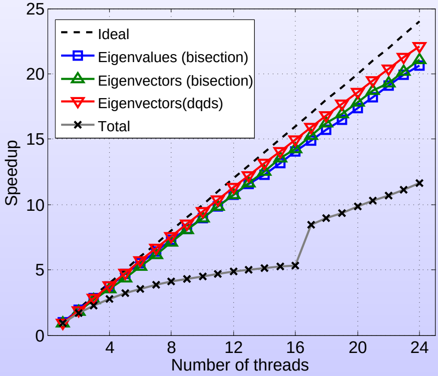
Matrix size: 16023. Frequency response analysis of automobiles.



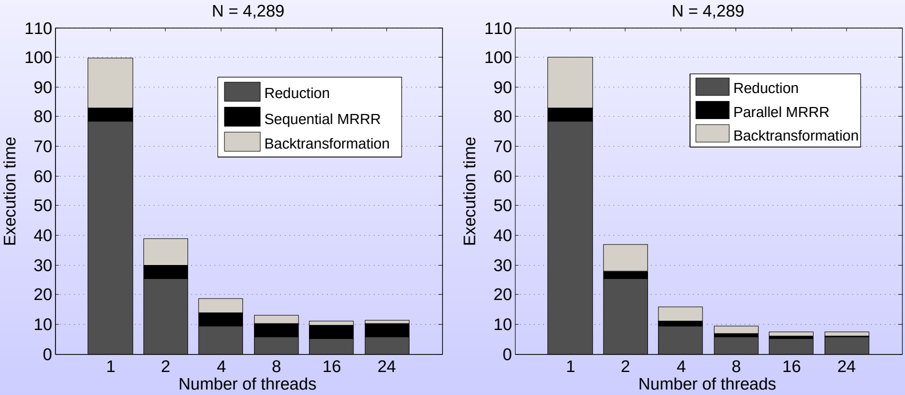


From almost 10 hours to 8.3 seconds.

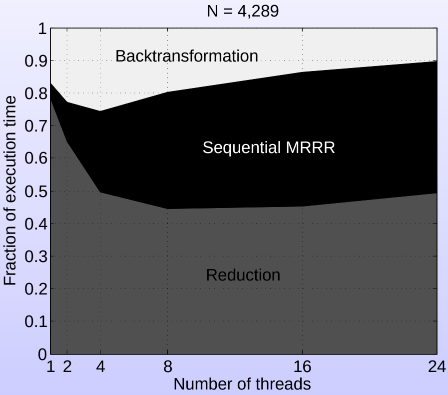

Time in seconds
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
Eigenvalues
Eigenvectors
LAPACK
2
4
8
16
24
Number of threads

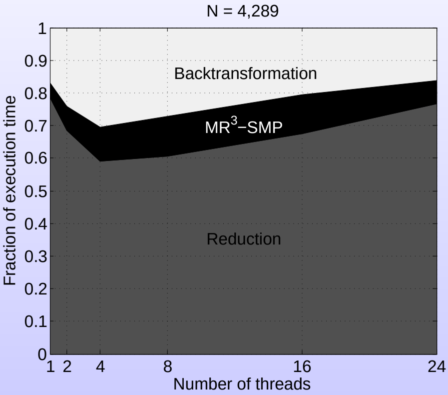

Speedup
25
20
15
10
5
0
Ideal
Eigenvalues (bisection)
Eigenvectors (bisection)
Eigenvectors(dqds)
Total
4
8
12
16
20
24
Number of threads

3 stages: before and after



3 stages: before and after





Conclusions

Multi-threaded BLAS for eigensolvers: not THAT good.

MRRR-SMP

- eigensolver tailored for multi-cores
- almost perfect speedups
- routines are available

Thank you for the attention.

Deutsche Forschungsgemeinschaft DFG

Financial support from the Deutsche Forschungsgemeinschaft (German Research Association) through grant GSC 111 is gratefully acknowledged.