z-Pares: A Complex Moment Based Parallel Sparse Eigenvalue Solver Package

Tetsuya Sakurai
University of Tsukuba

Joint work with
Hiroto Tadano, Du Lei, Ikuro Yamazaki, Yasunori Futamura

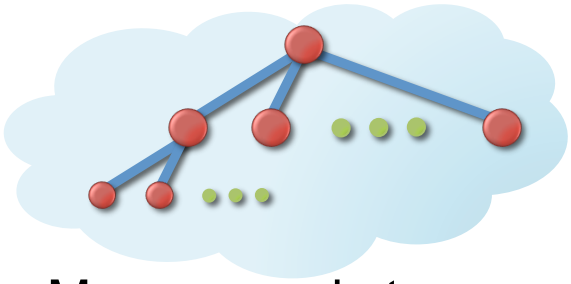
Contents

- An eigensolver for sparse matrices
  - Using contour integrals
- Systems of linear equations with multiple RHS
  - Block Krylov methods and shifted Krylov methods
- Numerical examples
  - SEP: Band calculation with real space density functional theory
  - GEP: Molecular orbital computation of protein

Motivation

- For post-petascale supercomputers, we need to consider hierarchical structures:

Computing nodes and cores



Memory and storage

Networks

Algorithms

- New algorithm design: New concept of algorithm design to adapt to the change of architectures

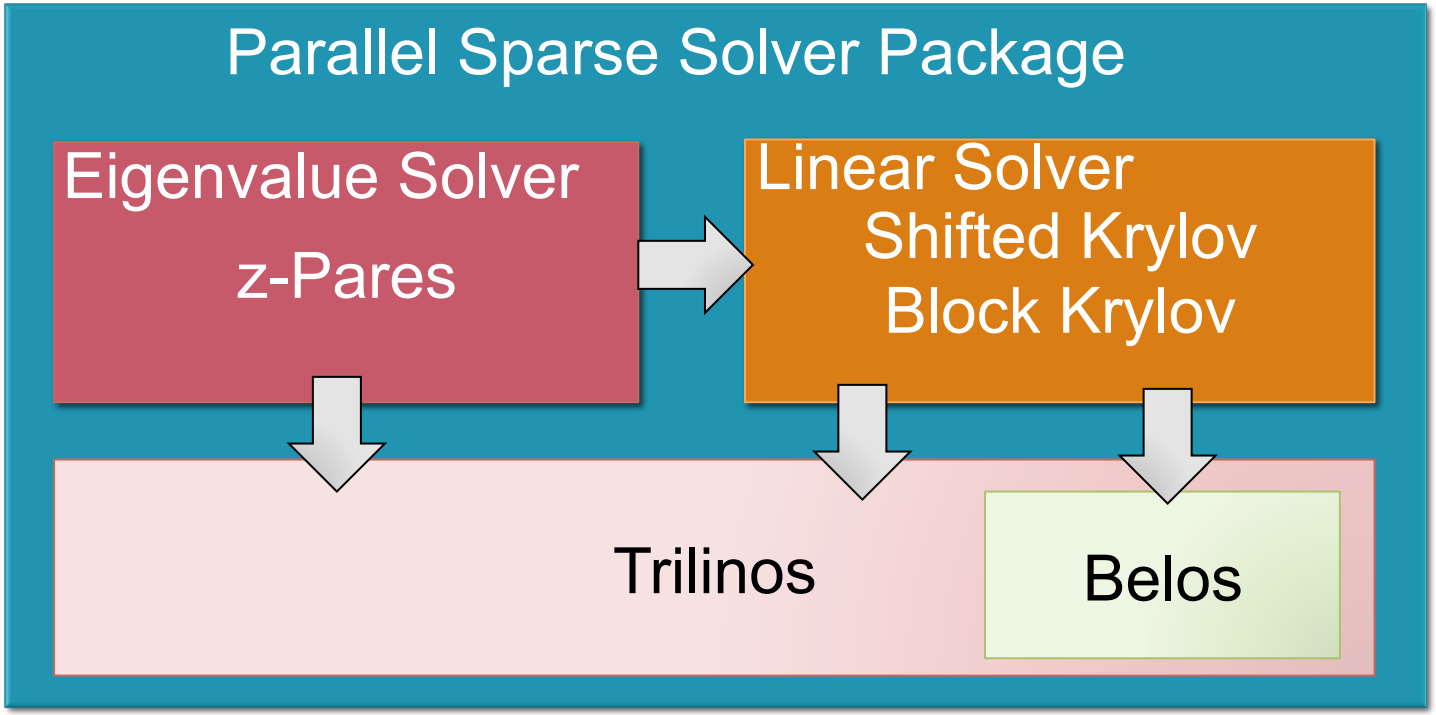
Development of Parallel Eigensolver Software

- CREST Project: Development of an Eigen-Supercomputing Engine using a Post-Petascale Hierarchical Model (FY2011 – FY2015)
  - For both dense and sparse matrices
    - In this talk, we focus on a sparse solver (z-Pares)

- Goal: Provide high performance eigensolver software for post-petascale architectures
  - Scalable and reliable parallel algorithms for eigenvalue computations
  - High performance implementation for various applications
  - Provide practical and portable software

Implementation of z-Pares

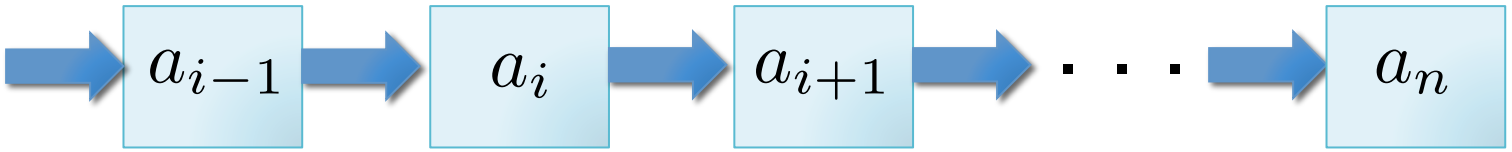
- A parallel eigensolver package for sparse matrices
  - Implemented on Trilinos
  - Uses Block Krylov methods and Shifted Kyrlov methods for spectral transformation



Our Approach for Parallel Scalability

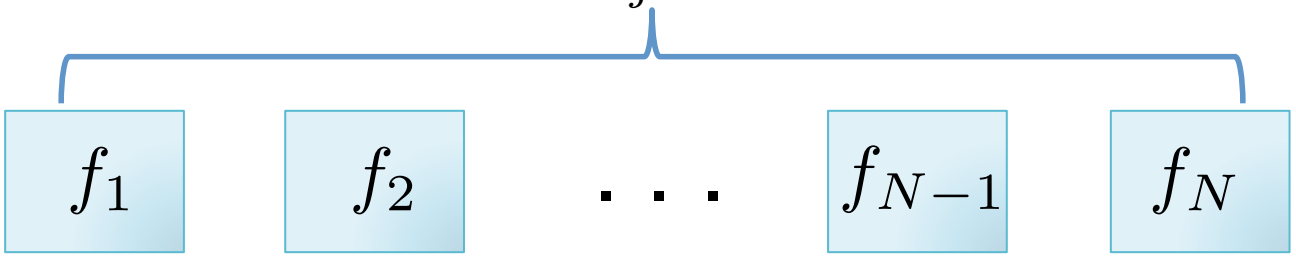
- Avoid recurrence calculations
  - Algorithm described by a recurrence relation (ex. Krylov methods):

$$a_{i+1} = b_i\, a_i + c_i\, a_{i-1}$$



  - Algorithm described by a summation (ex. Numerical quadrature):

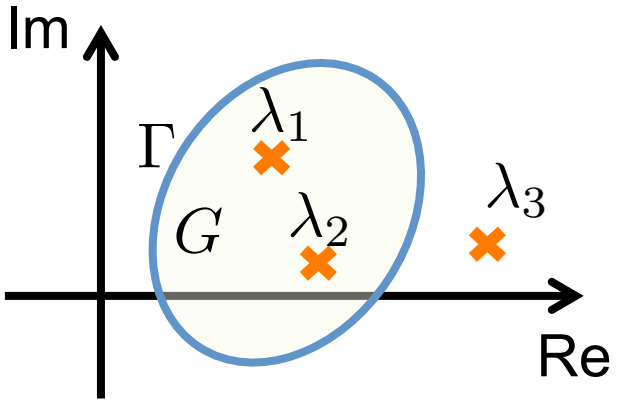
$$I_N = \sum_{j=1}^{N} w_j f_j$$



Basic Idea

- Contour integral for a rational function

$$\frac{1}{2\pi \mathrm{i}} \oint_{\Gamma} \sum_{i=1}^{n} \frac{\nu_i}{z-\lambda_i} \mathrm{d}z = \sum_{\lambda_i \in G} \nu_i$$



- Spectral decomposition of the resolvent of A

$$(zI - A)^{-1} = \sum_{i=1}^{n} \frac{P_i}{z-\lambda_i}, \quad P_i = \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{H}}$$

λ_i : eigenvalue, $\boldsymbol{u}_i$: eigenvector
(for simplicity, we considered the case that A is real symmetric)

- Contour integral for the resolvent of A

$$\frac{1}{2\pi \mathrm{i}} \oint_{\Gamma} (zI - A)^{-1} \mathrm{d}z = \sum_{\lambda_i \in G} P_i$$

Numerical Quadrature

- Approximated by a numerical quadrature

$$\sum_{\lambda_i \in G} P_i = \frac{1}{2\pi \mathrm{i}} \oint_{\Gamma} (zI - A)^{-1} \mathrm{d}z \approx \sum_{j=1}^{N} w_j (z_j I - A)^{-1}$$

z_j : quadrature point, w_j : quadrature wait

- Apply for an arbitrary vector $\boldsymbol{v}$

$$\sum_{\lambda_i \in G} P_i \boldsymbol{v} \approx \sum_{j=1}^{N} w_j (z_j I - A)^{-1} \boldsymbol{v}$$

Systems of linear equations at shift points z_j

$$(z_j I - A)\boldsymbol{x}_j = \boldsymbol{v}, \ \ j = 1, \ldots, N$$

Eigensolver using Contour Integrals

- SS method (Sakurai, et. al. (2003), Asakura, et al. (2009))
  - Eigenvalue problem

$$T(\lambda)\boldsymbol{x} = \boldsymbol{0}$$

where $T(\lambda)$ is an analytic matrix function.

  - Complex moments are calculated.

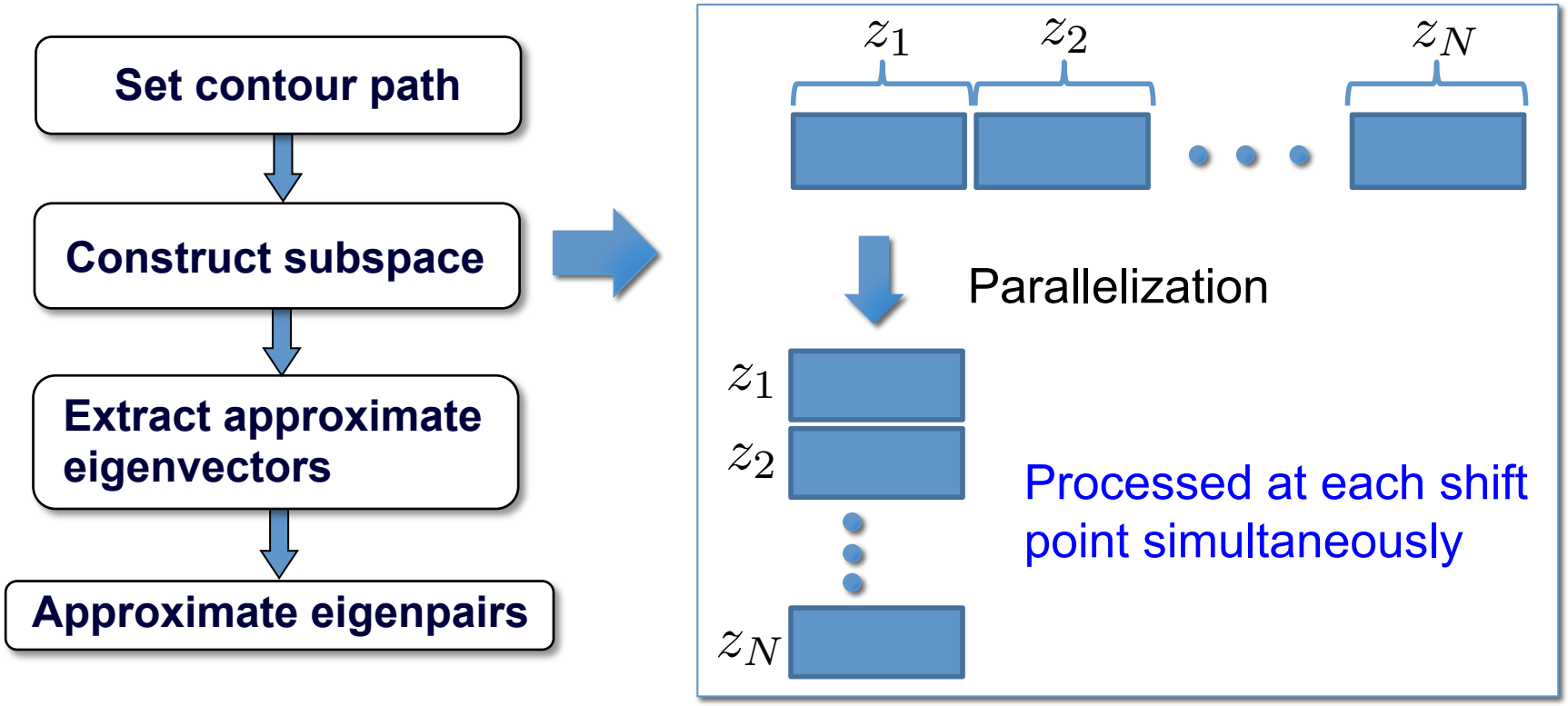
$$\boldsymbol{s}_k = \frac{1}{2\pi \mathrm{i}} \oint_{\Gamma} z^k T(z)^{-1} \mathrm{d}z\, \boldsymbol{v} \approx \sum_{j=1}^{N} w_j z_j^k T(z_j)^{-1} \boldsymbol{v}$$

  - When multiple input vectors $V = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_L]$ are used, systems with multiple right-hand sides are solved.

$$T(z_j)X_j = V,\ \ j = 1, \ldots, N$$

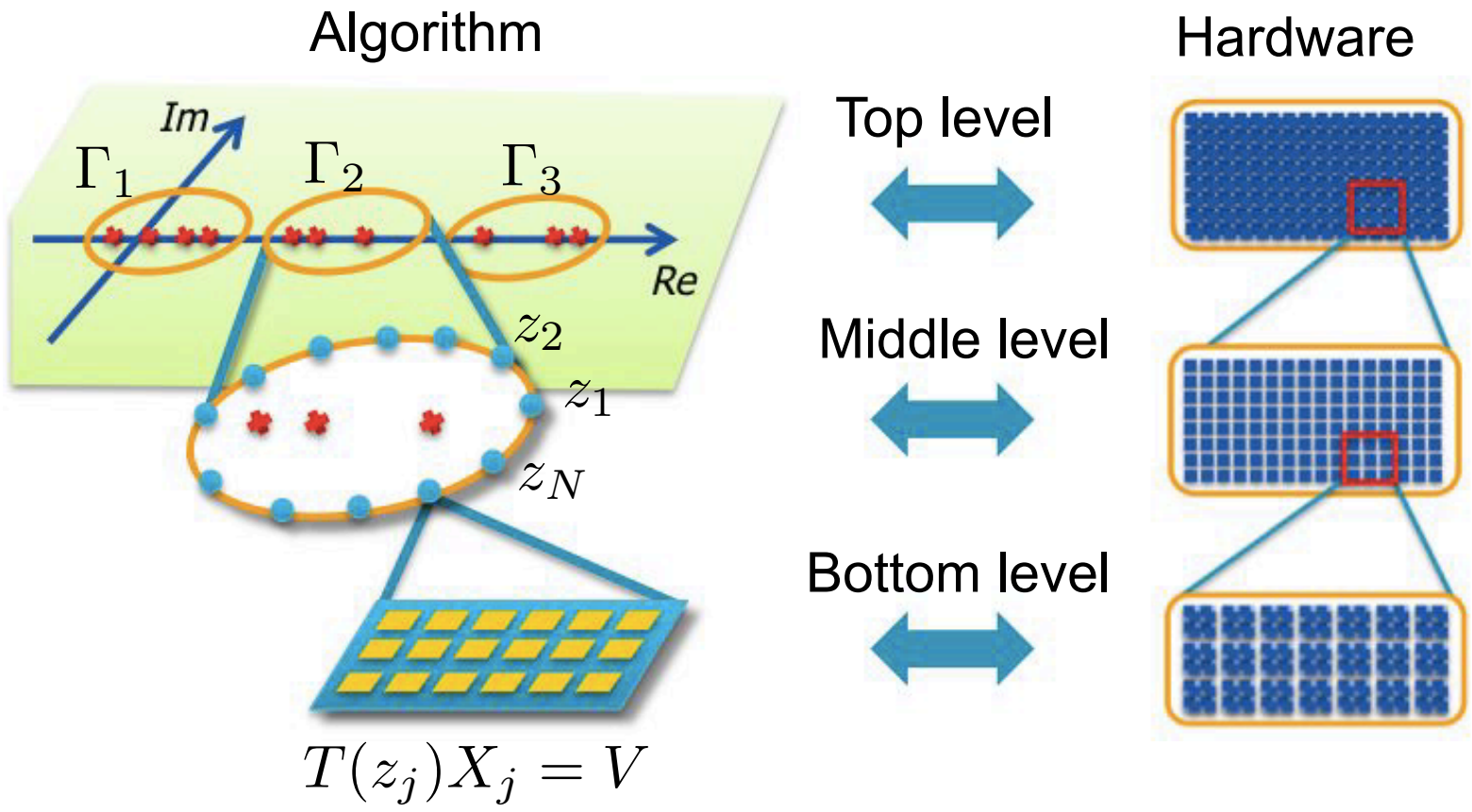
Parallelization

- The SS method requires no successive process in the subspace construction step



Hierarchical Parallel Structure

- Hardware is grouped according to a hierarchical structure of the algorithm



Shifted Krylov Subspace Method for SEP

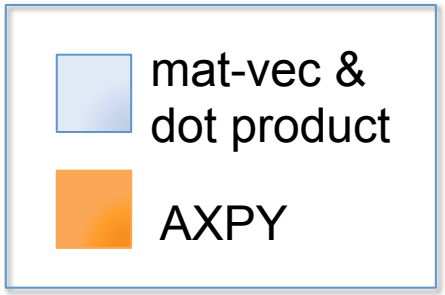
- Shifted linear systems:
  - In case of standard eigenvalue problems, shifted linear systems are solved.

$$(z_j I - A)X_j = V, \ \ j = 1, \ldots, N$$

- Reduction of number of Mat-Vec:
  - By using the shift invariance of the Krylov subspace, we can reduce the number of matrix-vector multiplications.

normal Krylov method for shifted systems

shifted Krylov method



Shifted Krylov Subspace Method for SEP

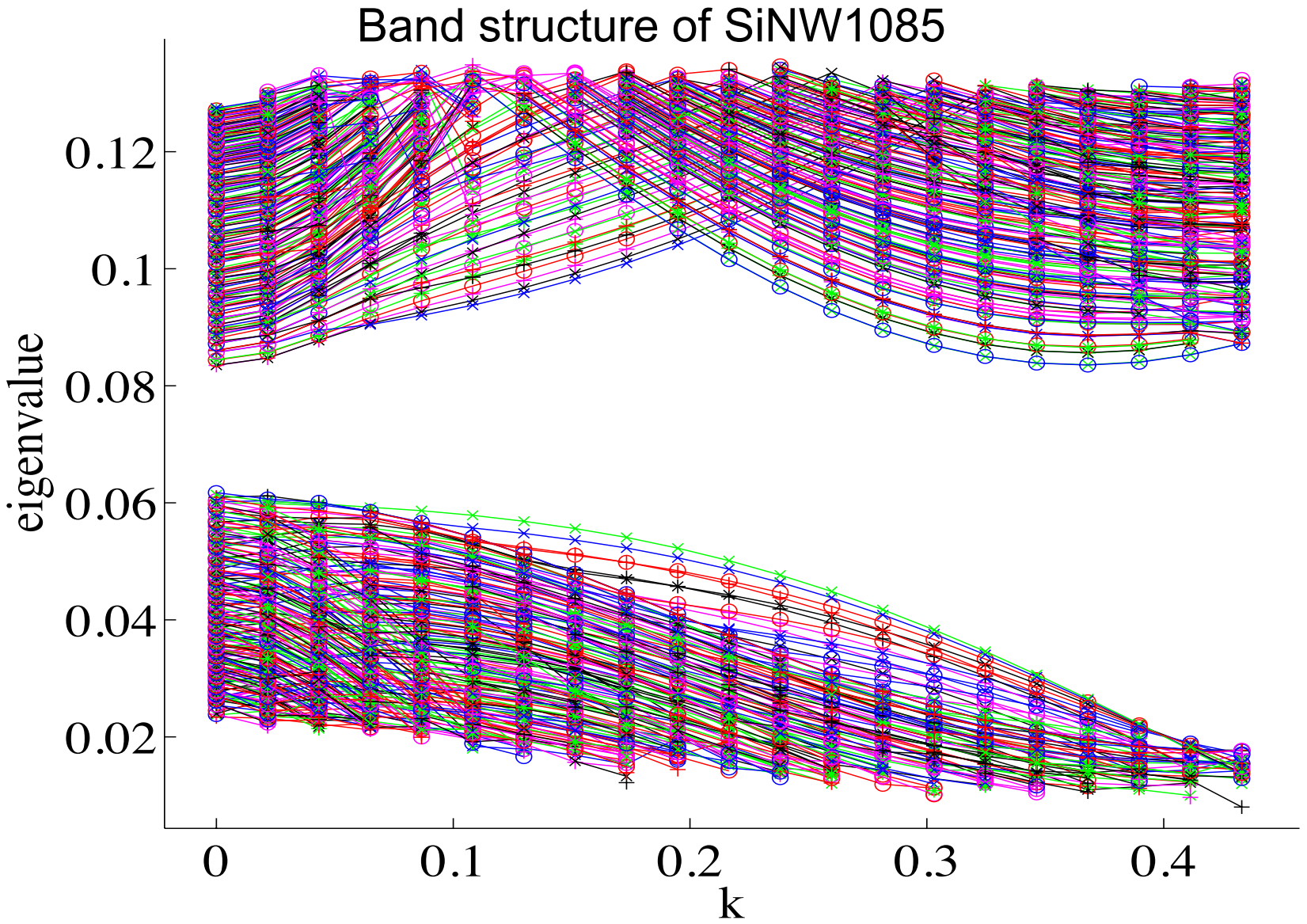
- In case of a lot of shift points:
  - If there are a lot of shifted systems, the computational cost for AXPY becomes dominant

- Shifted Block Krylov method:
  - AXPY for multiple RHSs can be implemented by GEMM
    - Improves performance

Numerical Examples for SEPs

- Application for band calculation with real space density functional theory (RSDFT)
  - An interior standard eigenvalue problem (SEP)
    - Eigenvalues around the band gap
    - Solved several matrices: $A(\mathrm{k})\,\boldsymbol{u} = \lambda\,\boldsymbol{u}, \mathrm{k} = 0, \Delta\mathrm{k}, 2\Delta\mathrm{k},...$
- Material: Silicon nanowire
- Test Environment: T2K-Tsukuba System
  - 2,048 cores for 1,085 atoms
  - 8,192 cores for 1,654 atoms

Numerical Examples for SEPs



Numerical Examples for SEPs

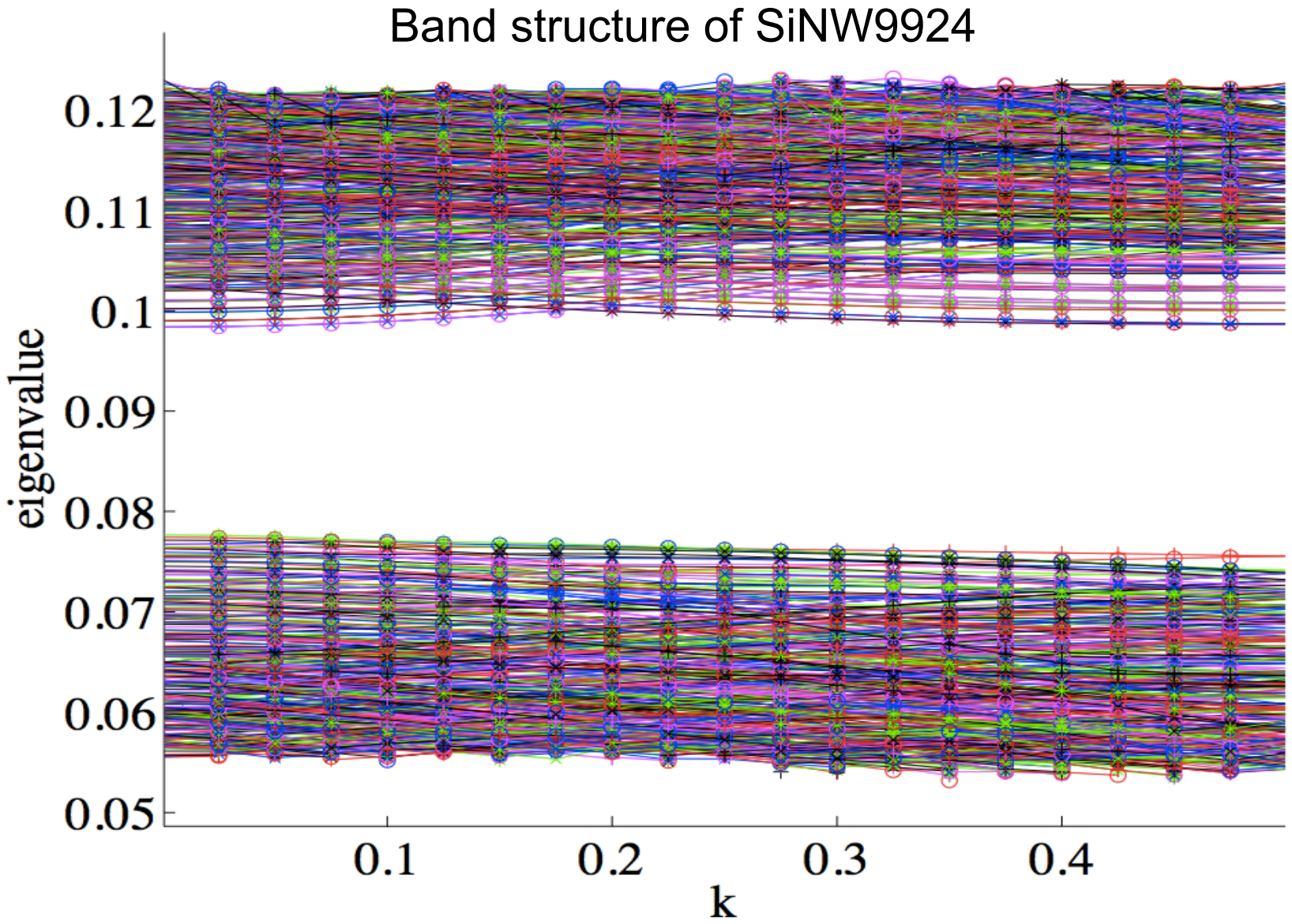
- Eigensolver:
  CG: Conjugate Gradient method
  SS: SS method with a shifted Block Krylov linear solver

#atoms	Mat. dim.	#core	Time (sec)		SS/CG Time ratio
			CG	SS	
1,085	576,000	2,048	2,373	1,534	0.65
1,654	1,453,248	8,192	1,074*	563	0.52

*certain eigenvalues were not obtained in the interval to be computed

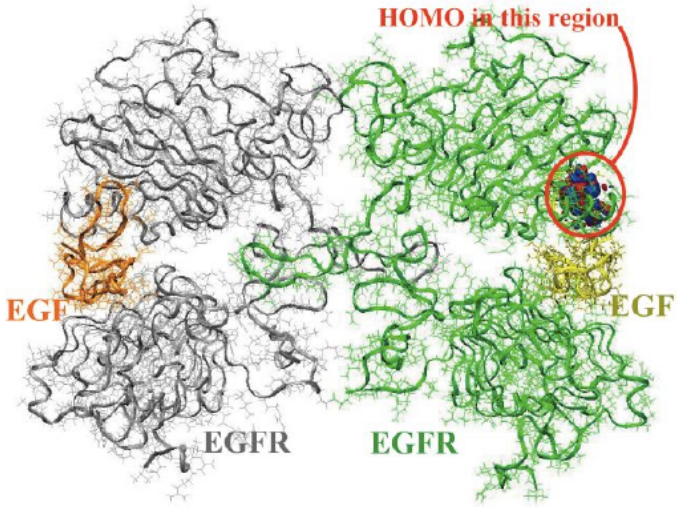
We have executed our code on the K computer with 6,144 cores for the case of silicon nanowire 9,924 atoms.

Numerical Examples for SEPs



Numerical Examples for GEPs

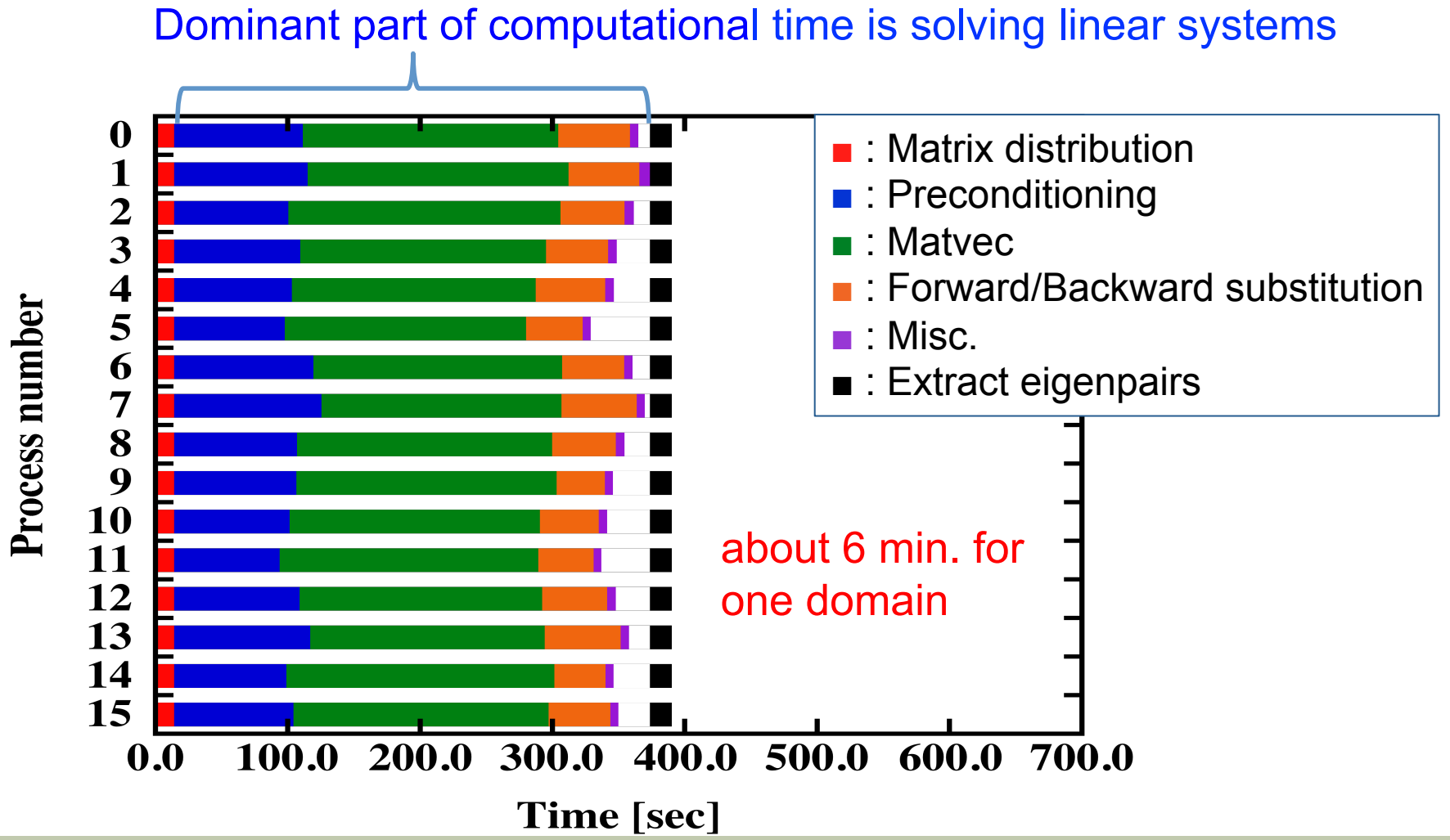
- Application for molecular orbital (MO) computation of protein
  - Generalized eigenvalue problem (GEP)



- Target protein:
  Epidamal Growth Factor Receptor
  17,246 atoms
- Matrix dimension: 96,234,
  #nonzeros: 456,807,648
- Test environment:
  T2k-Tsukuba System
- Linear solver: Block COCG method
- Preconditioner:
  MUMPS (sparse direct solver) for a dropped coefficient matrix

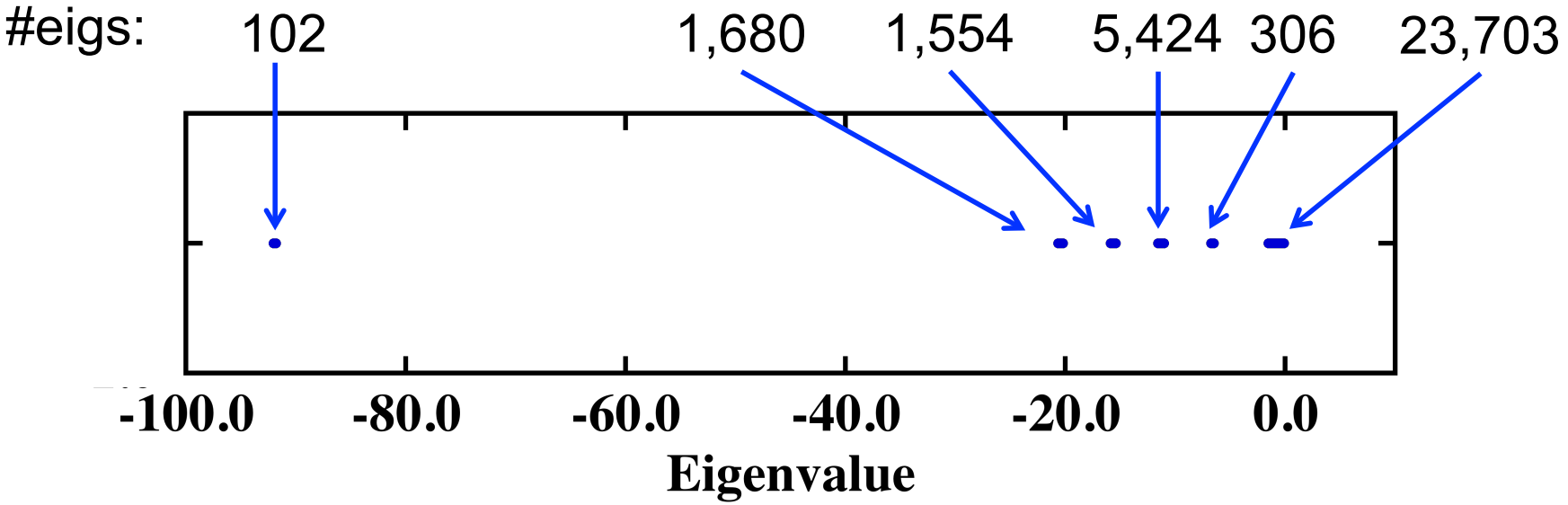
Numerical Examples for GEPs

- One contour path: 16 linear systems are solved with 128 cores



Numerical Examples for GEPs

- Multiple contour paths
  - Eigenvalues from lowest orbital to highest occupied orbital are calculated for self consistent field (SCF) iteration
  - 314 contour paths are set (128 cores for each contour path)
  - 32,769 eigenvalues and corresponding eigenvectors are obtained



Conclusions

- Development of parallel eigensolver and its software for post-petascale computing environment
  - Contour integral based method
    - Scalable algorithm
    - Hierarchical parallel structure
  - Several number of linear systems with multiple RHSs
    - Block Krylov methods and Shifted Krylov methods
    - Time for solving linear systems is dominant

- Looking for
  - efficient linear solvers and preconditioners
  - applications