Damped Jacobi-type Preconditioner for Supernova Simulation

今倉 暁 / Akira IMAKURA

Center for Computational Sciences, University of Tsukuba

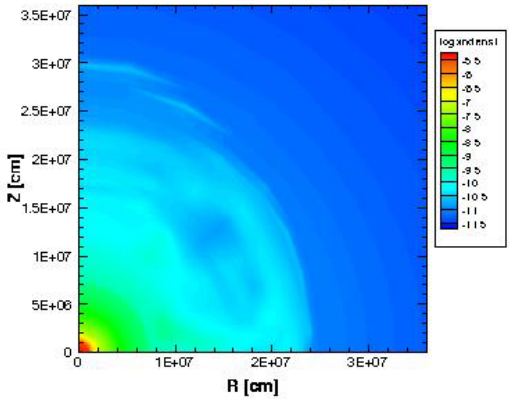
Joint work with

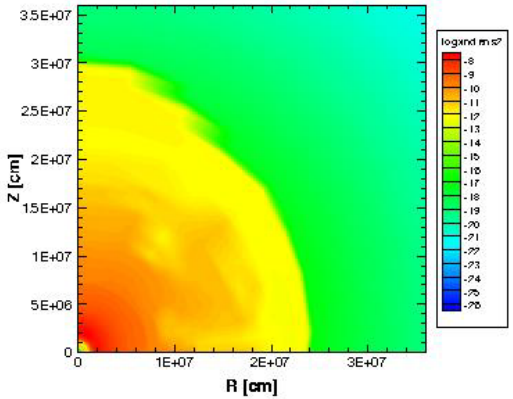
櫻井鉄也(筑波大学), 住吉光介(沼津高専), 松古栄夫(KEK)

Introduction

- Target simulation
  - ***Supernova simulation***
    -- 3D Neutrino transfer (time evolution of neutrino distribution)

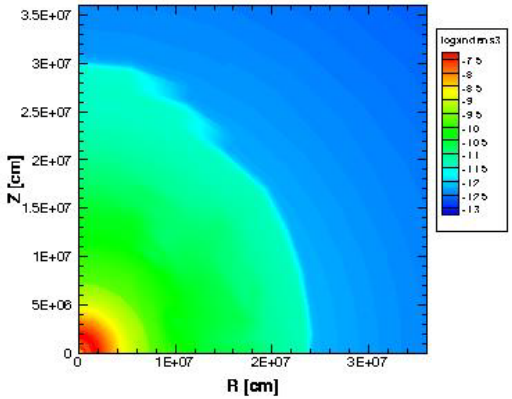
ν_e

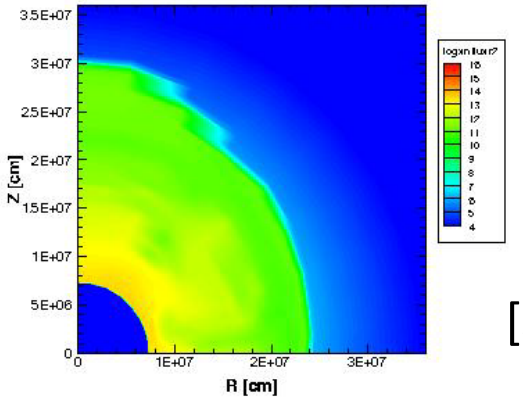




ν_e

ν_μ





ν_e
flux
r-direction

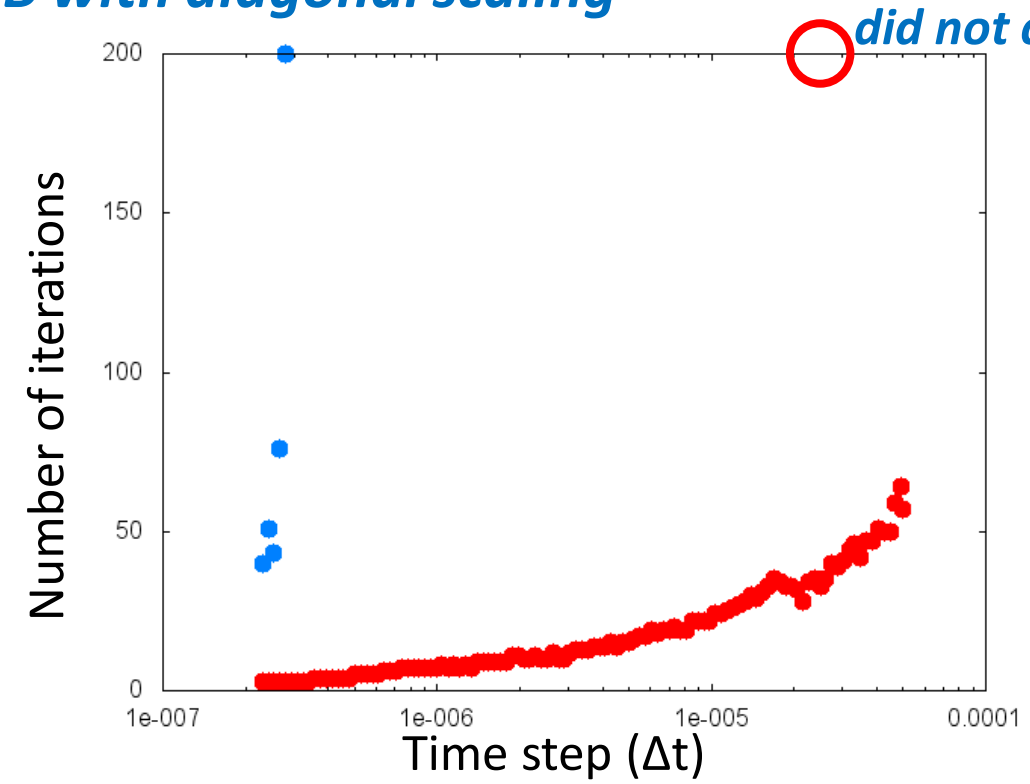
[Sumiyoshi, 2011]

Introduction

- Target simulation
  - Scale
    - -- $N_r x N_q x N_f = 200 \times 9 \times 9$, $N_e x N_{nq} x N_{nf} = 14 \times 6 \times 12$
    - -- 1/8 circle, r = 0~1140 km
  - Outline
    - -- Time evolution : Implicit method
      - → It is required to solve ***large and sparse linear systems***
    - -- The most time consuming part is to solve the systems
    - -- Required time evolution : t = 1 [sec.]
      - → Larger time step (Δt) is required ($\Delta t \approx 10^{-5}$ [sec.])
    - -- ***Larger Δt leads to more ill-conditioned systems***

Introduction

- Traditional code for linear systems
  - ***Bi-CGSTAB with diagonal scaling***



- Our main goal of this talk
  - Propose a ***Damped Jacobi-type preconditioner (with parameter estimation)*** for supernova simulation

Outline

- Introduction
- Solvers and preconditioners for linear systems
- Parameter estimation for damped Jacobi-type preconditioner
  - Damped Jacobi-type method
  - Parameter estimation
- Application to supernova simulation
- Conclusion

Solvers and preconditioners

- Target problem
  - Large, sparse, real and non-symmetric linear systems:

$$Ax = b, \quad A \in \mathbb{R}^{n \times n}, \quad x, b \in \mathbb{R}^n,$$

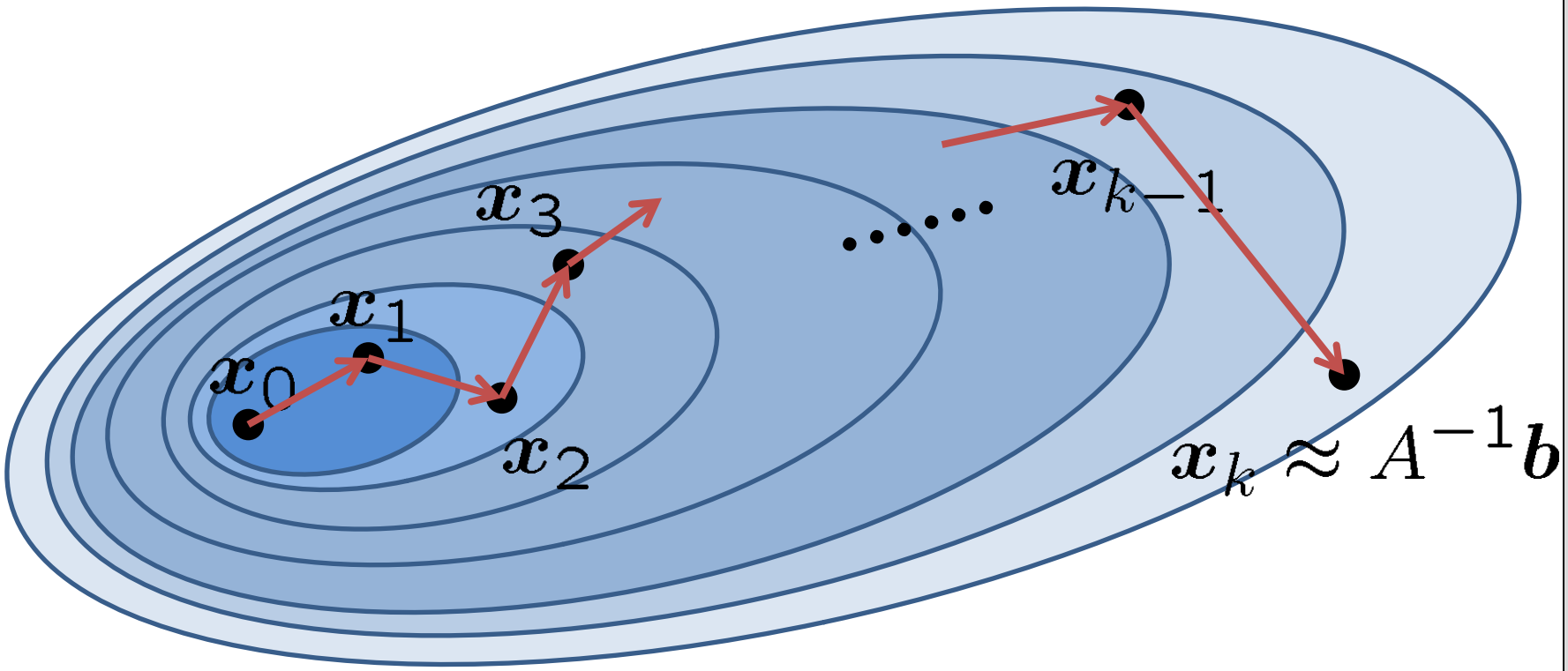
    obtained from ***supernova simulation***.

- Solvers for linear systems
  - Direct methods
    -- Gaussian elimination, LU decomposition, ...
  - Stationary iterative methods
    -- Jacobi, Gauss-Seidel, SOR, ...
  - ***Preconditioned*** Krylov subspace (iterative) methods
    -- Conjugate Gradient (CG) , GMRES, ***Bi-CGSTAB***, ...




Solvers and preconditioners

- Krylov subspace method
  - Basic ideal: Projection method based on the Krylov subspace
  - Rough sketch (initial guess: x_0)



Solvers and preconditioners

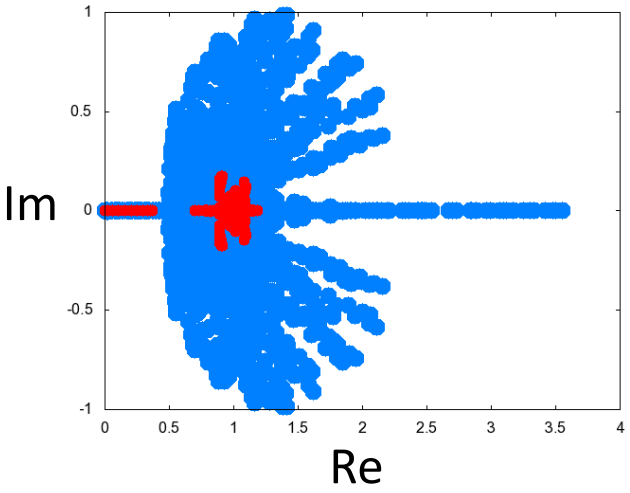
- Preconditioners for Krylov subspace methods
  - Convergence of Krylov subspace methods

Convergence depend on ***distribution of eigenvalues***

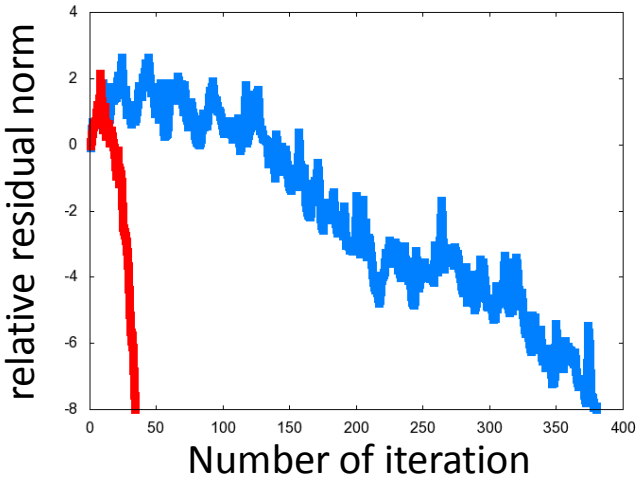
  - Preconditioners

$$Ax = b \quad \Leftrightarrow \quad AK^{-1}y = b, \quad x = K^{-1}y$$

Distribution of eigenvalues



Convergence history



Solvers and preconditioners

- Preconditioners for Krylov subspace methods
  - Incomplete factorization type
    -- IC, ILU, ...
  - Approximate inverse type
    -- SPAI, polynomial preconditioner, ***(diagonal) scaling***, ...
  - Variable type (use some iterative method)
    -- GMRES, Jacobi, ***Damped Jacobi-type method***
- Our goals of this talk
  - Propose a parameter estimation technique for damped Jacobi-type preconditioner
  - Examine the performance of the proposed preconditioner for linear systems from supernova simulation

Outline

- Introduction
- Solvers and preconditioners for linear systems
- Parameter estimation for damped Jacobi-type preconditioner
  - Damped Jacobi-type method
  - Parameter estimation
- Application to supernova simulation
- Conclusion




Parameter estimation for Damped Jacobi-type prec.

- Stationary iterative methods
  - Initial partition

$$A = M - N,$$

where M is assumed to be nonsingular

  - Recurrence formula

$$\boldsymbol{x}_k = M^{-1}N\boldsymbol{x}_{k-1} + M^{-1}\boldsymbol{b}$$

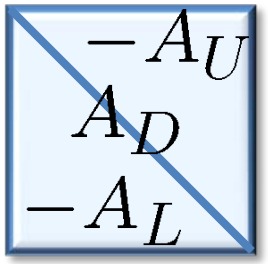
  - Convergence of stationary iterative method

Convergence depend on ***spectral radius of iteration matrix***:

$$\rho(G) := \rho(M^{-1}N) := \left|\max_i |\lambda_i(M^{-1}N)|\right|$$

Parameter estimation for Damped Jacobi-type prec.

- Ideal Stationary iterative methods
  - ***High convergence***
    -- It has small $\rho(G)$
  - ***High speed***
    -- It requires small computational costs per iteration
  - ***High parallelization efficiency***
    -- It has less sequential process (in each iteration)



- Traditional stationary iterative methods

	Convergence	Speed	Parallelization
Jacobi	Bad	Good	Good
Gauss-Seidel	Good	Good	Bad
SOR	Bad -- Excellent	Good	Bad

Jacobi method (suitable for parallelization)

Parameter estimation for Damped Jacobi-type prec.

- Damped Jacobi methods
  - Basic ideal:

Damped Jacobi method transforms the Jacobi's recurrence formula:

$$\begin{aligned} \boldsymbol{x}_k &= A_D^{-1}(A_L + A_U)\boldsymbol{x}_{k-1} + A_D^{-1}\boldsymbol{b} \\ &= \boldsymbol{x}_{k-1} + A_D^{-1}(\boldsymbol{b} - A\boldsymbol{x}_{k-1}) \end{aligned}$$

into

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha A_D^{-1}(\boldsymbol{b} - A\boldsymbol{x}_k)$$

by using damping parameter $\alpha \in \mathbb{R}$ in order to improve the convergence rate of Jacobi method.

Parameter estimation for Damped Jacobi-type prec.

- Damped Jacobi-**type** methods
  - Basic ideal:

Damped Jacobi method transforms the Jacobi's recurrence formula:

$$\begin{aligned} \boldsymbol{x}_k &= A_D^{-1}(A_L + A_U)\boldsymbol{x}_{k-1} + A_D^{-1}\boldsymbol{b} \\ &= \boldsymbol{x}_{k-1} + A_D^{-1}(\boldsymbol{b} - A\boldsymbol{x}_{k-1}) \end{aligned}$$

into

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha D^{-1}(\boldsymbol{b} - A\boldsymbol{x}_k)$$

by using damping parameter $\alpha \in \mathbb{R}$ and some diagonal matrix D in order to improve the convergence rate of Jacobi method.

Users need to set the parameter α,
and the convergence depend strongly on the parameter α

Parameter estimation for Damped Jacobi-type prec.

- Convergence theorem of damped Jacobi-type method

Theorem 1: Optimal damping parameter

Let $C(\gamma, \rho)$ ***be the inner region of the circle with center*** $\gamma \in \mathbb{R}$ ***and radius*** $\rho \in \mathbb{R}$ ***on the complex plane, and let*** $\gamma^*, \rho^* \in \mathbb{R}$ ***be defined by***

$$\gamma^*, \rho^* := \arg\min_{\gamma,\rho} \left|\frac{\rho}{\gamma}\right|,$$

$$s.t. \quad \lambda_i(D^{-1}A) \in C(\gamma, \rho), \quad i = 1, 2, \ldots, n.$$

Then the spectral radius $\rho(G_\alpha)$ ***of iteration matrix of damped Jacobi-type method is minimized by***

$$\alpha = \frac{1}{\gamma^*}$$ ***, and the minimized value is*** $$\rho(G_\alpha) = \left|\frac{\rho^*}{\gamma^*}\right|$$

Parameter estimation for Damped Jacobi-type prec.

- Convergence theorem of damped Jacobi-type method

Theorem 2: Convergence condition

The minimized spectral radius of iteration matrix of damped Jacobi-type method satisfies the following inequality:

$$\rho(G_{\alpha_{\mathrm{opt}}}) = \left|\frac{\rho^*}{\gamma^*}\right| < 1$$

if and only if

$$\mathcal{R}e\left(\lambda_i(D^{-1}A)\right) > 0, \quad i = 1, 2, \ldots, n$$

or

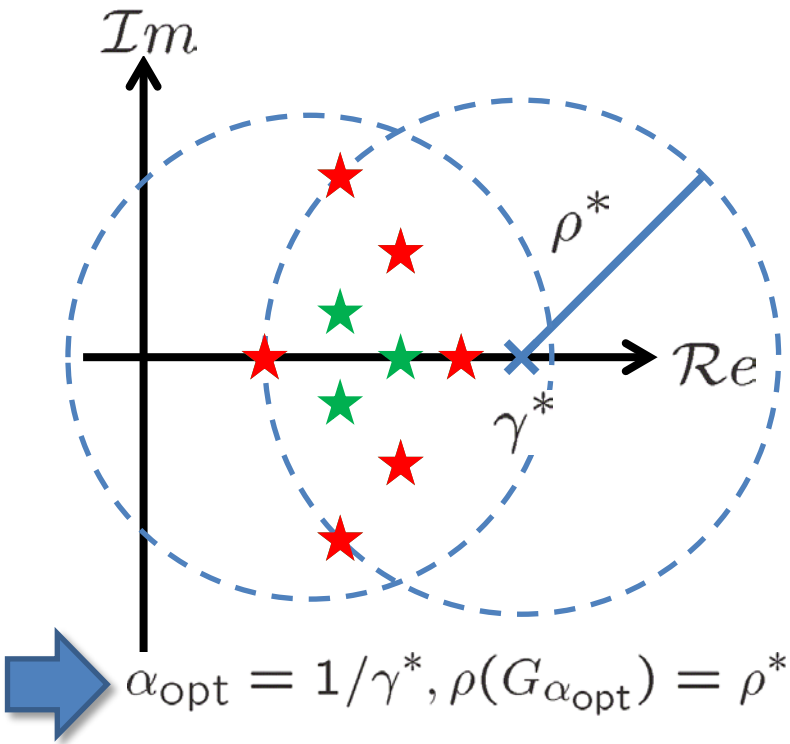
$$\mathcal{R}e\left(\lambda_i(D^{-1}A)\right) < 0, \quad i = 1, 2, \ldots, n$$

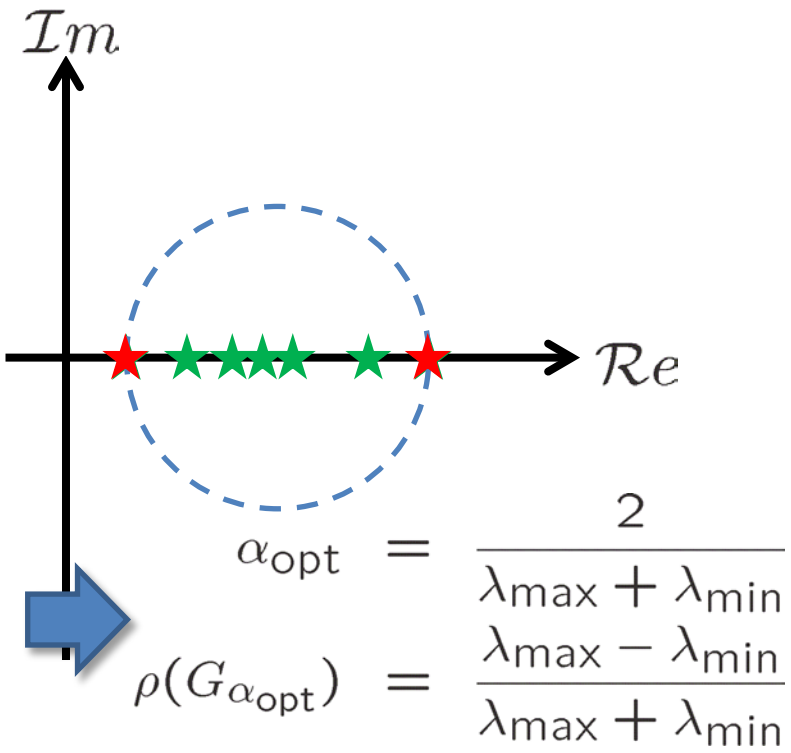
Parameter estimation for Damped Jacobi-type prec.

- Convergence theorem of damped Jacobi-type method

Remark: Distribution of eigenvalues and Optimal parameter

Distribution of $\lambda_i(D^{-1}A)$





Damping parameter can be optimized ***distribution of eigenvalues***

Parameter estimation for Damped Jacobi-type prec.

- Parameter estimation (***proposal preconditioner***)
  - Basic ideal: ***Offline tuning***

-- Damping parameter is estimated before algorithms run.

Input
Initial guess x_0
Diagonal matrix D

Parameter estimation I
Approximate extreme eigenvalues of $D^{-1}A$
by some iterations of Arnoldi method

Parameter estimation II
Estimate $\alpha_{\mathrm{opt}} = 1/c^*, \rho(G_{\alpha_{\mathrm{opt}}}) = r^*/c^*$
from approximated extreme eigenvalues

Preconditioned Krylov
Run damped Jacobi-type preconditined Krylov subspace method with estimated damping parameter α_{opt}




Outline

- Introduction
- Solvers and preconditioners for linear systems
- Parameter estimation for damped Jacobi-type preconditioner
  - Damped Jacobi-type method
  - Parameter estimation
- Application to supernova simulation
- Conclusion




Application to supernova simulation

- Supernova simulation
  - Scale
    - -- $N_r x N_q x N_f = 200 \times 9 \times 9$, $N_e x N_{nq} x N_{nf} = 14 \times 6 \times 12$
    - -- 1/8 circle, r = 0~1140 km
  - Obtained linear systems
    - -- Dimension ≈ 16 million
    - -- Number of nonzeros ≈ 1.28 billion
      (Average nonzeros per row or column ≈ 80)
- Experimental condition
  - KEK SR16000/M1
    - -- 1 node 32 cores (logical CPU 64 cores)

Application to supernova simulation

- Proposed preconditioner
  - Bi-CGSTAB with Damped Jacobi-type preconditioner
    - -- Iteration for preconditioner : 20
    - -- Approximate extreme eigenvalues : 20 iters of Arnoldi
    - -- Cutoff parameter : 0.01
    - -- Diagonal matrix :

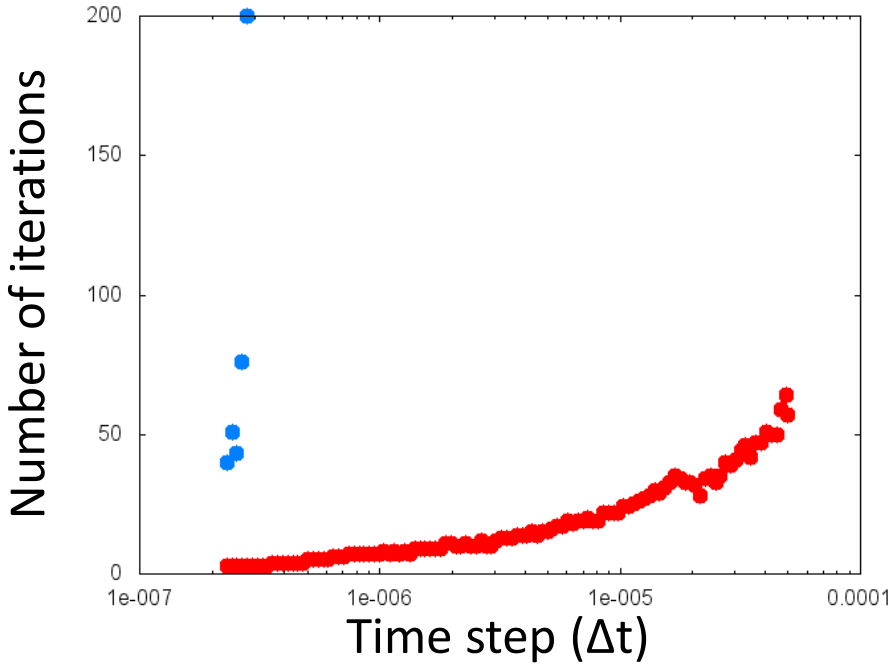
$$d_i = a_{ii}$$

$$d_i = \sum_j |a_{ij}|$$

$$d_i = \sqrt{\sum_j a_{ij}^2}$$

Application to supernova simulation

- Numerical results
  - Time step (Δt) v.s. number of iterations
    -- Bi-CGSTAB with ***Diagonal scaling*** / ***Proposed precond.***



  - Computational time [sec.]
    -- ***Diagonal scaling***: 0.3 / iter
    -- ***Proposed***: 3.0 / iter (estimation: 3.0 / 1 diagonal matrix)

Outline

- Introduction
- Solvers and preconditioners for linear systems
- Parameter estimation for damped Jacobi-type preconditioner
  - Damped Jacobi-type method
  - Parameter estimation
- Application to supernova simulation
- Conclusion

Conclusion

- Conclusion
  - In this talk, we proposed
    Damped Jacobi-type preconditioner with parameter estimation.
  - From our numerical experiments, we leaned that the proposed preconditioner will be ***efficient for supernova simulation***.
- Further investigation
  - Further improvement of the proposed preconditioner
    -- Implementation
    -- Another acceleration techniques