Iterative Solvers for Large Linear Systems

Part Vb: Variants of BiCG

Andreas Meister

University of Kassel, Department of Analysis and Applied Mathematics

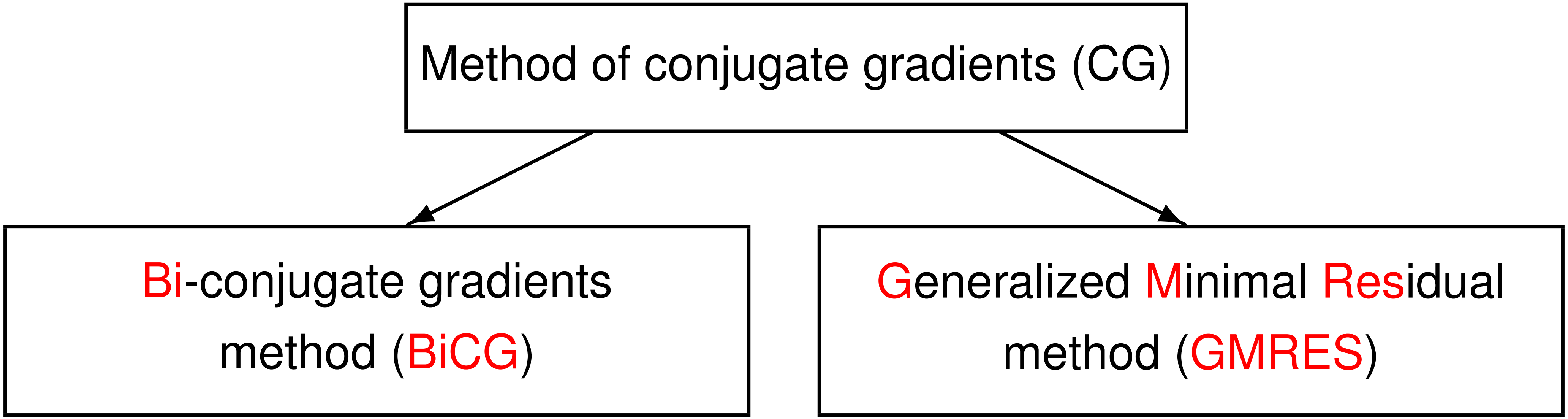
Outline

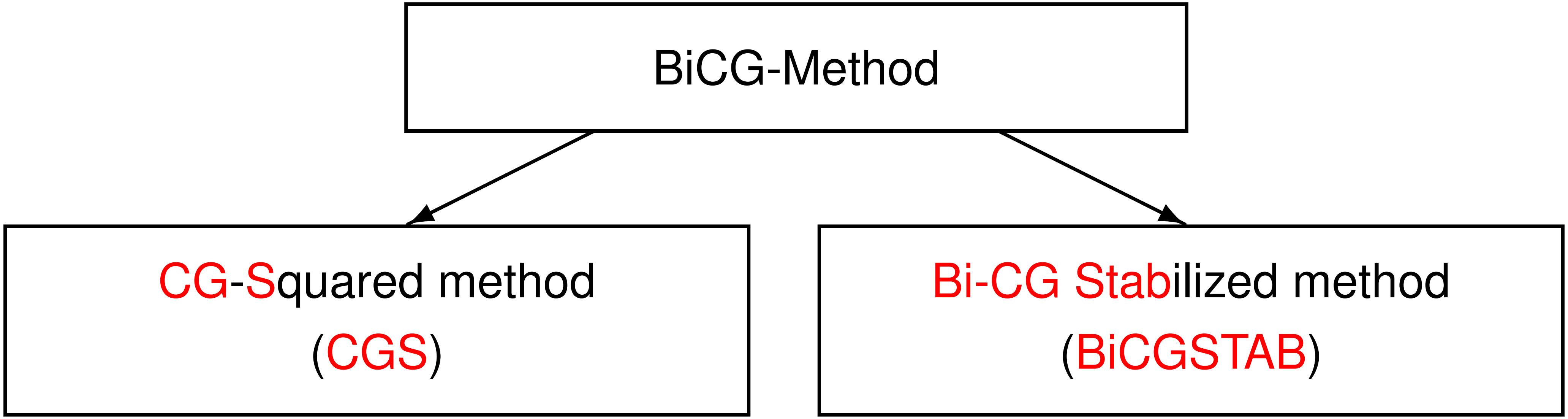
- Basics of Iterative Methods
- Splitting-schemes
  - Jacobi- u. Gauß-Seidel-scheme
  - Relaxation methods
- Methods for symmetric, positive definite Matrices
  - Method of steepest descent
  - Method of conjugate directions
  - CG-scheme

Outline

- Multigrid Method
  - Smoother, Prolongation, Restriction
  - Twogrid Method and Extension
- Methods for non-singular Matrices
  - GMRES
  - BiCG, CGS and BiCGSTAB
- Preconditioning
  - ILU, IC, GS, SGS, ...

Methods for non-singular Matrices





CGS-Algorithm

Aim:

- Accelerate the BiCG-method
- Avoid multiplications with A^T

Monitoring:

- Polynomial representation

$$r_j = \varphi_j(A) r_0 \,, \quad p_j = \psi_j(A) r_0 \quad \Longrightarrow \quad r_j^* = \varphi_j(A^T) r_0 \,, \quad p_j^* = \psi_j(A^T) r_0$$

- Occurence of r_j^* and p_j^*

  Solely to calculate the scalar values α_j and β_j

$$(r_j, r_j^*) \;\text{ and }\; (A p_j, p_j^*)$$

Basic Idea:

$$(x, A^T y) = x^T A^T y = (Ax)^T y = (Ax, y) \Longrightarrow (Ax, A^T y) = (A^2 x, y)$$

CGS-Algorithm

Reformulation:

Using

$$r_j = \varphi_j(A)r_0\ ,\quad p_j = \psi_j(A)r_0 \quad \text{and} \quad r_j^* = \varphi_j(A^T)r_0\ ,\quad p_j^* = \psi_j(A^T)r_0$$

yields

$$(r_j, r_j^*) = (\varphi_j(A)r_0\ , \varphi_j(A^T)r_0) = (\underbrace{\varphi_j^2(A)r_0}_{=:\ \hat{r}_j}\ , r_0) = (\hat{r}_j, r_0)$$

as well as

$$(A\,p_j, p_j^*) = (A\,\psi_j(A)r_0\ , \psi_j(A^T)r_0) = (A\ \underbrace{\psi_j^2(A)r_0}_{=:\hat{p}_j}\ , r_0) = (A\,\hat{p}_j, r_0)$$

Technical Exercise:

Express the scalar values α_j and β_j by means of $\hat{r}_j$ and $\hat{p}_j$

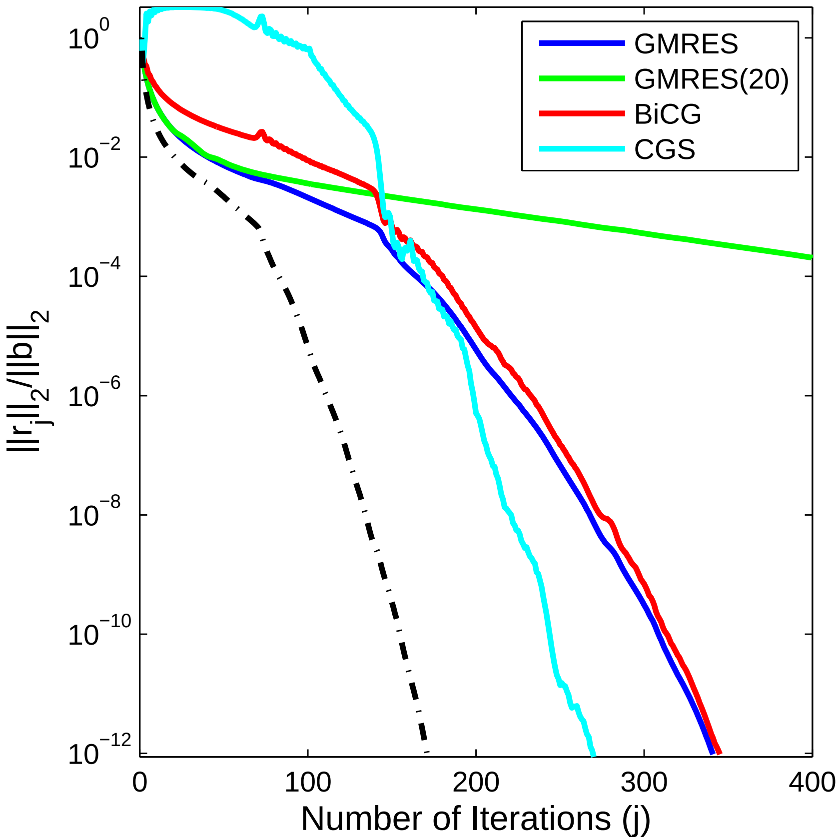
CGS-Algorithm

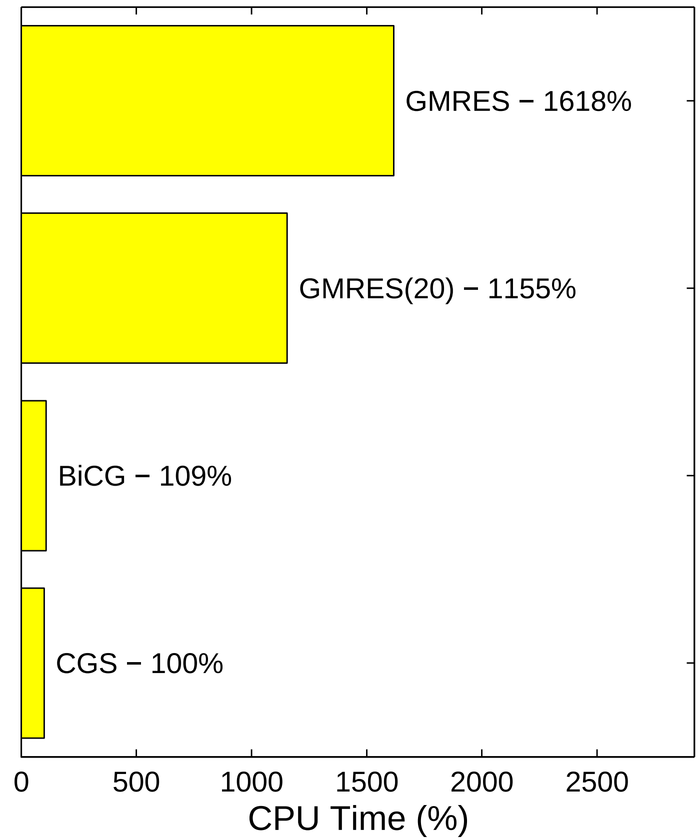
CGS-Algorithm

CGS-Algorithmus —

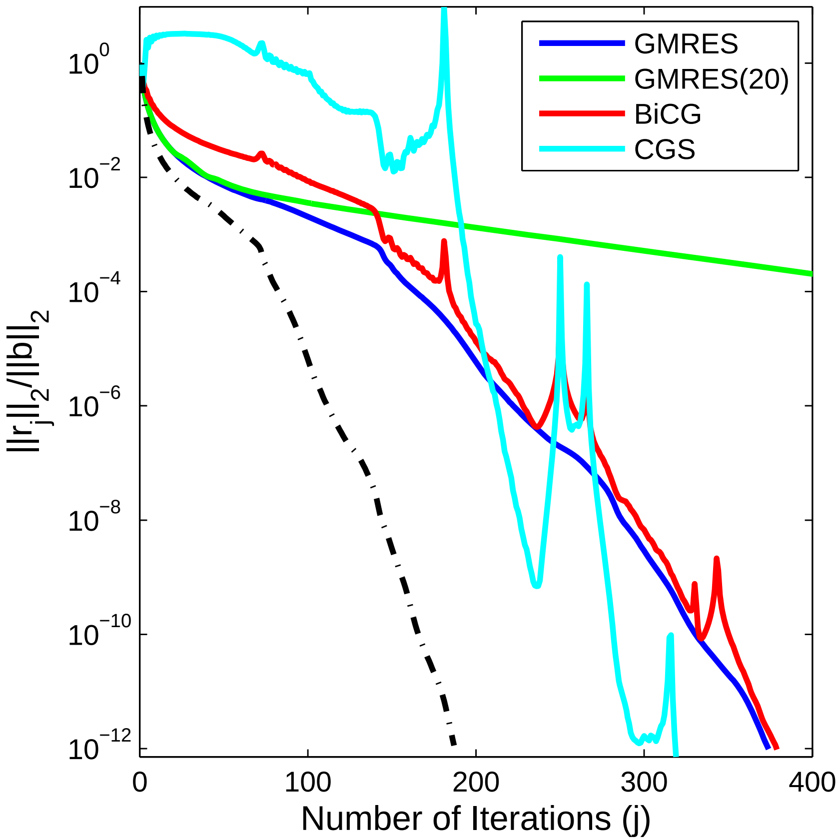
Wähle $\boldsymbol{x}_0 \in \mathbb{R}^n$ und $\varepsilon > 0$
$\boldsymbol{u}_0 = \boldsymbol{r}_0 = \boldsymbol{p}_0 := \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}_0\,,\ j := 0$
Solange $\|\boldsymbol{r}_j\|_2 > \varepsilon$
$\boldsymbol{v}_j := \boldsymbol{A}\boldsymbol{p}_j\,,\ \alpha_j := \frac{(\boldsymbol{r}_j,\boldsymbol{r}_0)_2}{(\boldsymbol{v}_j,\boldsymbol{r}_0)_2}$
$\boldsymbol{q}_j := \boldsymbol{u}_j - \alpha_j \boldsymbol{v}_j$
$\boldsymbol{x}_{j+1} := \boldsymbol{x}_j + \alpha_j\left(\boldsymbol{u}_j + \boldsymbol{q}_j\right)$
$\boldsymbol{r}_{j+1} := \boldsymbol{r}_j - \alpha_j \boldsymbol{A}\left(\boldsymbol{u}_j + \boldsymbol{q}_j\right)$
$\beta_j := \frac{(\boldsymbol{r}_{j+1},\boldsymbol{r}_0)_2}{(\boldsymbol{r}_j,\boldsymbol{r}_0)_2}$
$\boldsymbol{u}_{j+1} := \boldsymbol{r}_{j+1} + \beta_j \boldsymbol{q}_j$
$\boldsymbol{p}_{j+1} := \boldsymbol{u}_{j+1} + \beta_j\left(\boldsymbol{q}_j + \beta_j \boldsymbol{p}_j\right),\ j := j+1$

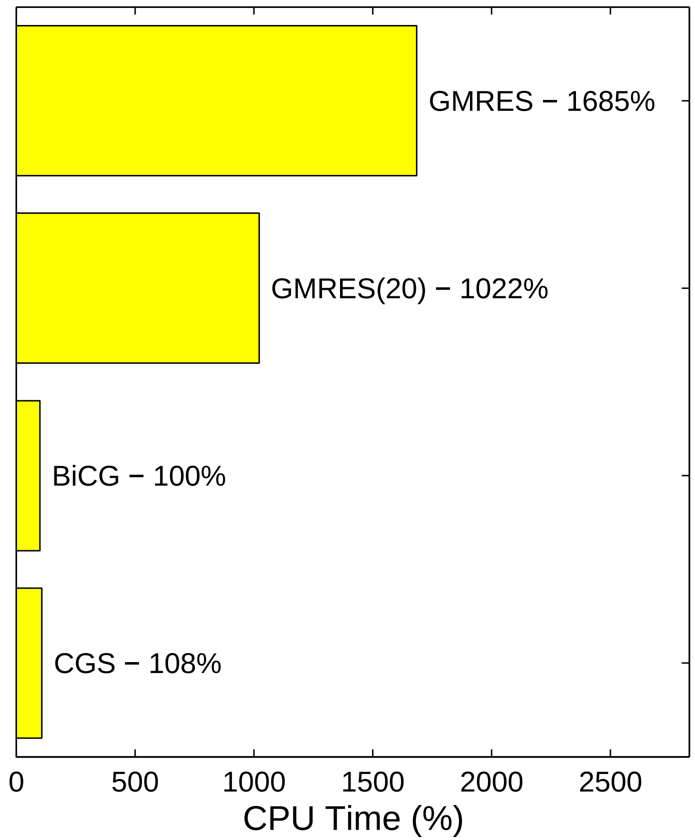
Test 1: Pure Diffusion ($\alpha = 0$, $\epsilon = 1$)





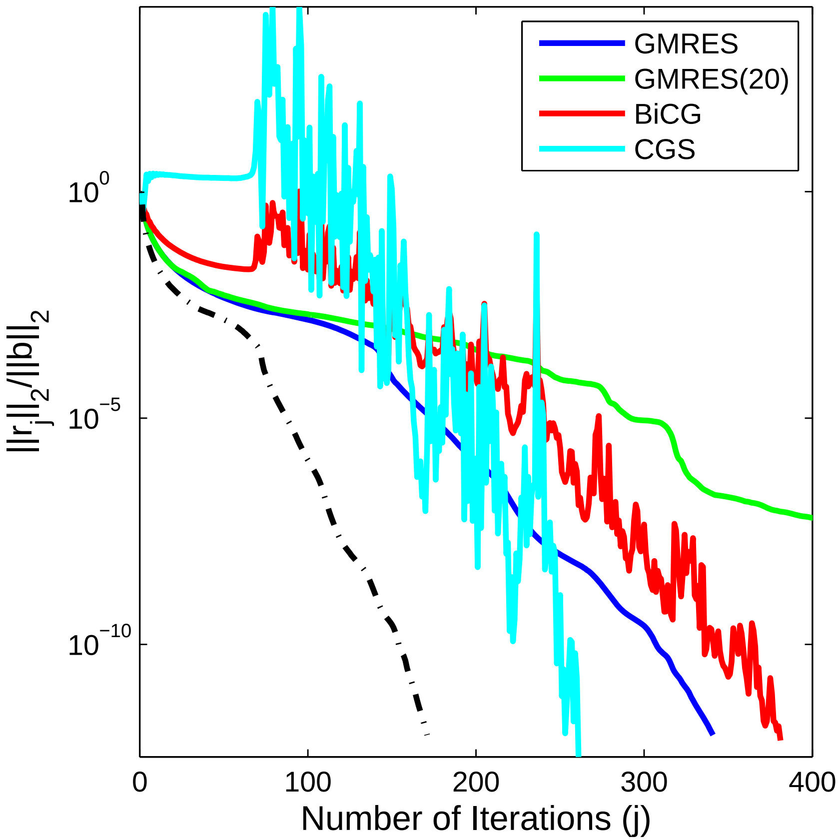
Test 2: Weak Convection-Diffusion ($\alpha = 0.1$, $\epsilon = 1$)

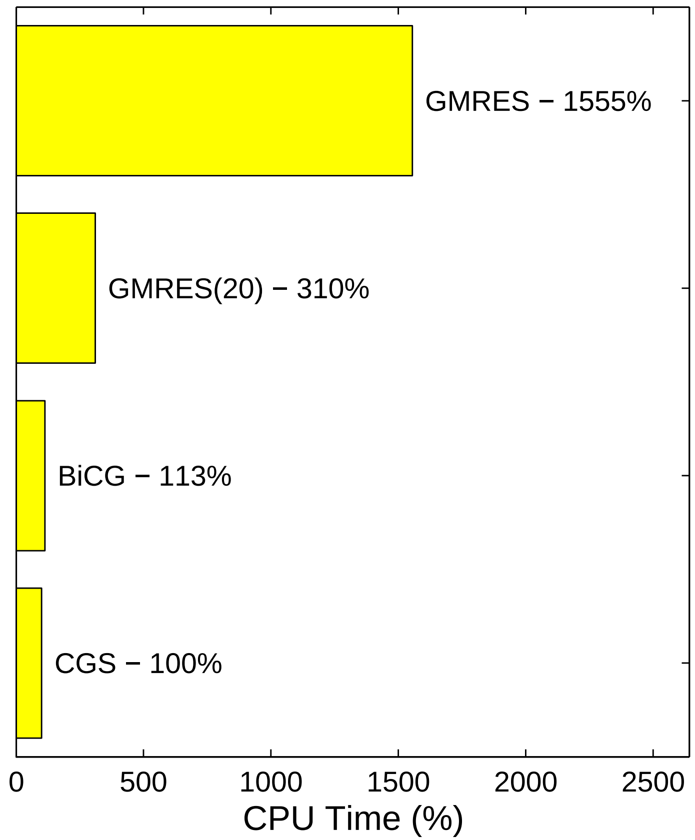




Comparison of GMRES, GMRES(m), BiCG and CGS

Test 3: Convection-Diffusion ($\alpha = 1$, $\epsilon = 0.1$)





CGS-Algorithm - Summary

Derivation:

- Based on BiCG-Algorithm
- Squaring the polynomial representation

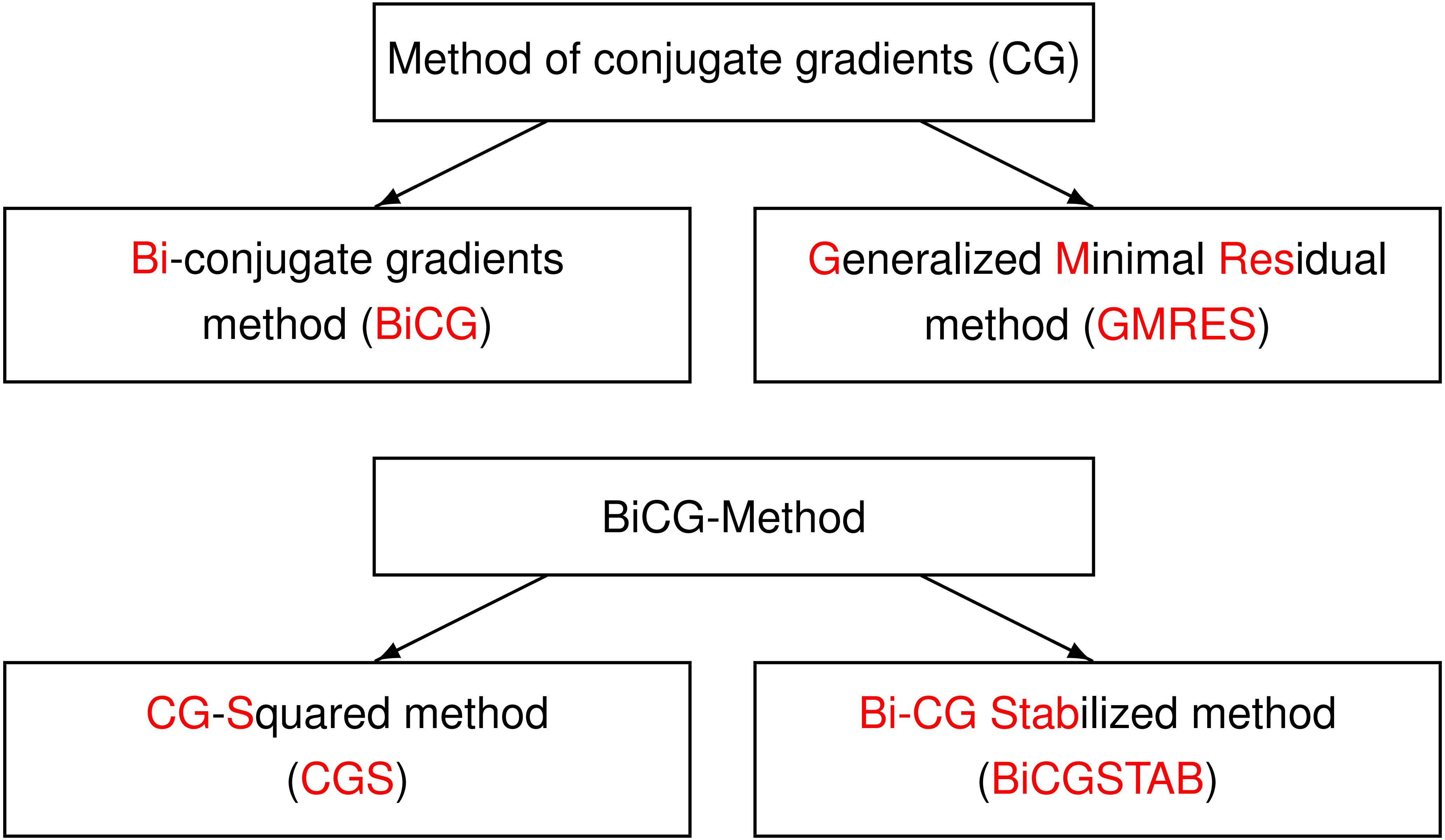
Advantages:

- Keenly less storage requirements (compared to GMRES)
- No symmetry constraint on A (compared to CG)
- Requires no multiplications with A^T (compared to BiCG)

Disadvantages:

- No minimization of an underlying functional
  $\rightarrow$ Oszillations in the convergence history
- Possible break down due to division by $(A\,p_j, r_0)$

Methods for non-singular Matrices



BiCGSTAB-Algorithm

Aim:

- Improving the BiCG- and CGS-method
- Avoid multiplications with A^T
- Introducing a minimization of the residual

Procedure:

- Polynomial representation

$$r_j = \varphi_j(A) r_0 \;, \quad p_j = \psi_j(A) r_0$$

- Employ

$$\tilde{r}_j^* = \Phi_j(A^T) r_0 \;, \quad \tilde{p}_j^* = \Phi_j(A^T) r_0$$

with

$$\Phi_0(A^T) = I, \quad \Phi_{j+1}(A^T) = (I - \omega_j A^T)\Phi_j(A^T)$$

BiCGSTAB-Algorithm

Reformulation:

Introducing $\tilde{r}_j^* = \Phi_j(A^T)r_0$ and $\tilde{p}_j^* = \Phi_j(A^T)r_0$ into BiCG yields

$$(r_j, \tilde{r}_j^*) = (\varphi_j(A)r_0, \Phi_j(A^T)r_0) = (\Phi_j(A)\varphi_j(A)r_0, r_0)$$

and

$$(A\,p_j, \tilde{p}_j^*) = (A\,\psi_j(A)r_0, \Phi_j(A^T)r_0) = (\Phi_j(A)A\,\psi_j(A)r_0, r_0)$$

Minimization of the residual:

- $r_{j+1} = (I - \omega_j A)s_j$
- Define $f_j(\omega) = \|(I - \omega A)s_j\|_2^2$

$$\Rightarrow \quad f_j'(\omega) = -2(As_j, s_j) + 2\omega(As_j, As_j)$$

$$f_j''(\omega) = 2(As_j, As_j) \geq 0$$

$$\Rightarrow \quad \omega_j = \arg\min_{\omega\in\mathbb{R}} f_j(\omega) = \frac{(As_j, s_j)}{(As_j, As_j)}$$

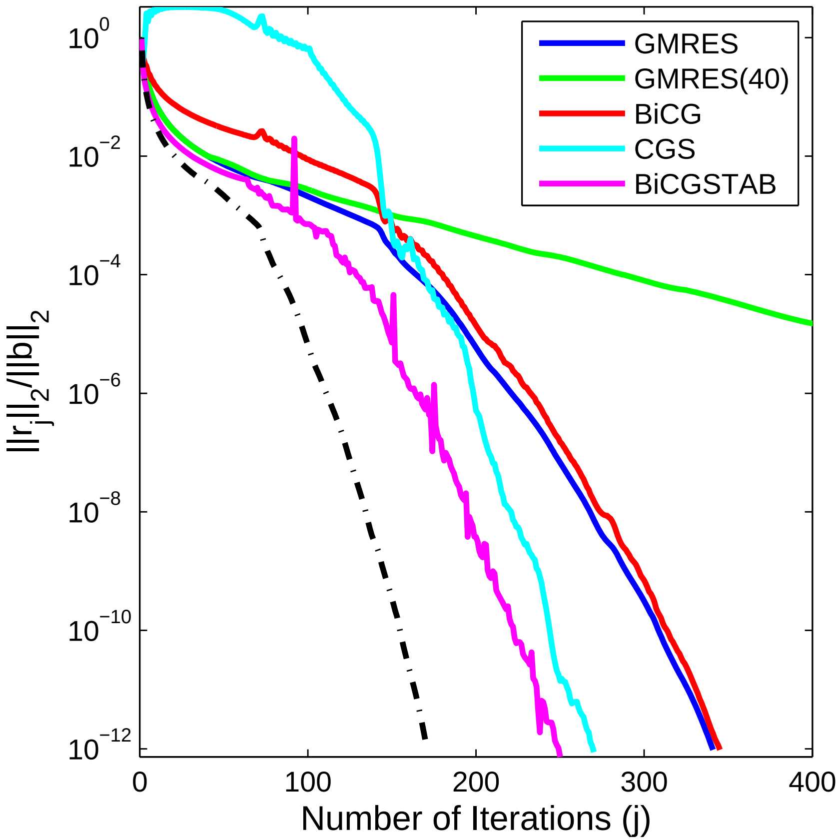
BiCGSTAB-Algorithm

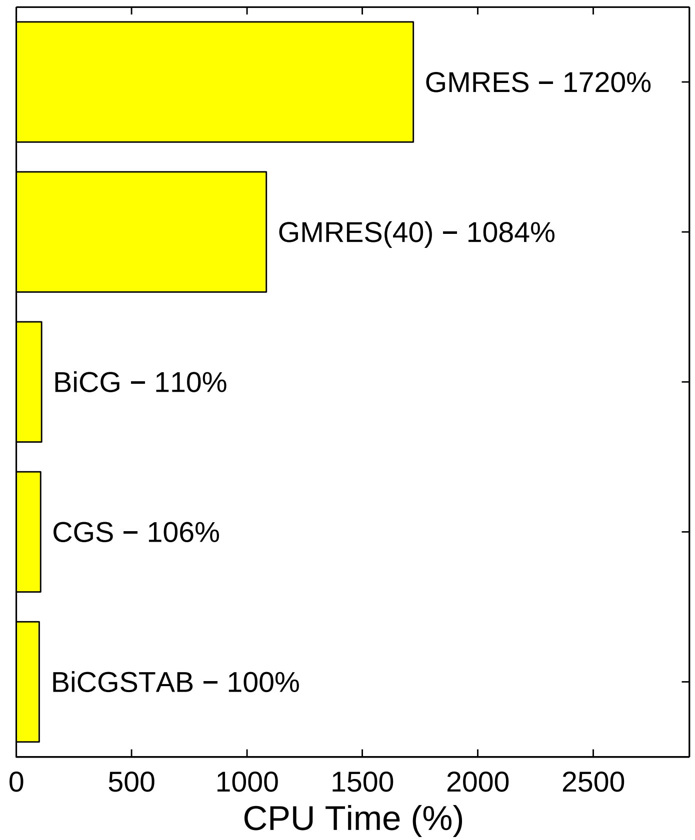
BiCGSTAB-Algorithmus —

Wähle $\boldsymbol{x}_0 \in \mathbb{R}^n$ und $\varepsilon > 0$
$\boldsymbol{r}_0 := \boldsymbol{p}_0 := \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}_0\,,\ \rho_0 := (\boldsymbol{r}_0, \boldsymbol{r}_0)_2\,,\ j := 0$
Solange $\|\boldsymbol{r}_j\|_2 > \varepsilon$
$\quad \boldsymbol{v}_j := \boldsymbol{A}\boldsymbol{p}_j\,,\ \alpha_j := \dfrac{\rho_j}{(\boldsymbol{v}_j, \boldsymbol{r}_0)_2}$
$\quad \boldsymbol{s}_j := \boldsymbol{r}_j - \alpha_j \boldsymbol{v}_j\,,\ \boldsymbol{t}_j := \boldsymbol{A}\,\boldsymbol{s}_j$
$\quad \omega_j := \dfrac{(\boldsymbol{t}_j, \boldsymbol{s}_j)_2}{(\boldsymbol{t}_j, \boldsymbol{t}_j)_2}$
$\quad \boldsymbol{x}_{j+1} := \boldsymbol{x}_j + \alpha_j \boldsymbol{p}_j + \omega_j \boldsymbol{s}_j$
$\quad \boldsymbol{r}_{j+1} := \boldsymbol{s}_j - \omega_j \boldsymbol{t}_j$
$\quad \rho_{j+1} := (\boldsymbol{r}_{j+1}, \boldsymbol{r}_0)_2\,,\ \beta_j := \dfrac{\alpha_j}{\omega_j}\dfrac{\rho_{j+1}}{\rho_j}$
$\quad \boldsymbol{p}_{j+1} := \boldsymbol{r}_{j+1} + \beta_j\,(\boldsymbol{p}_j - \omega_j \boldsymbol{v}_j)\,,\ j := j+1$

Comparison of GMRES, BiCG, CGS and BiCGSTAB

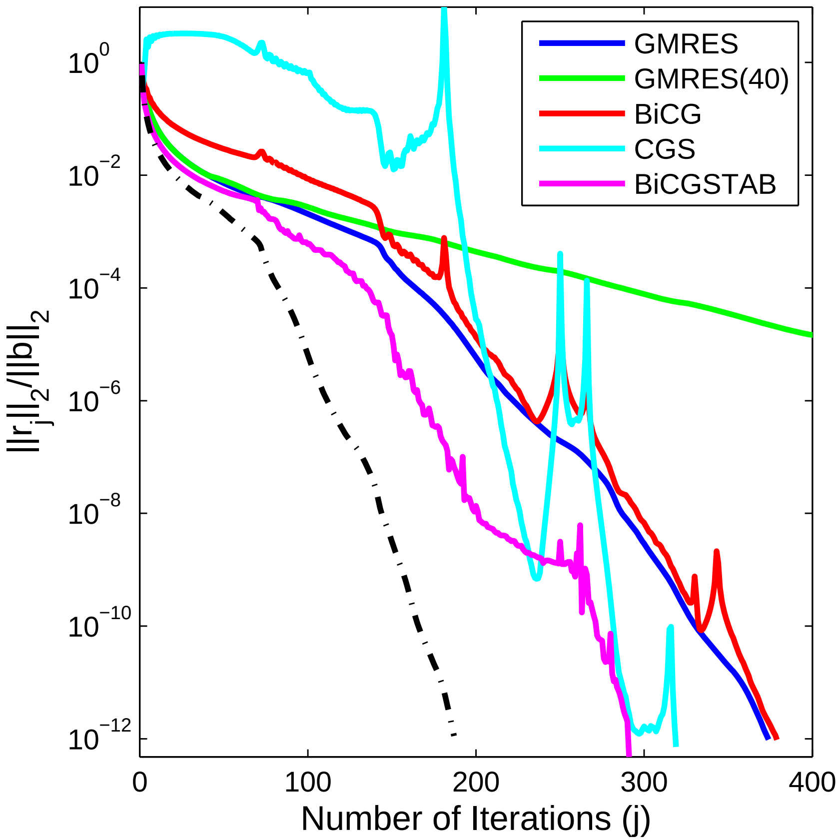
Test 1: Pure Diffusion ($\alpha = 0$, $\epsilon = 1$)

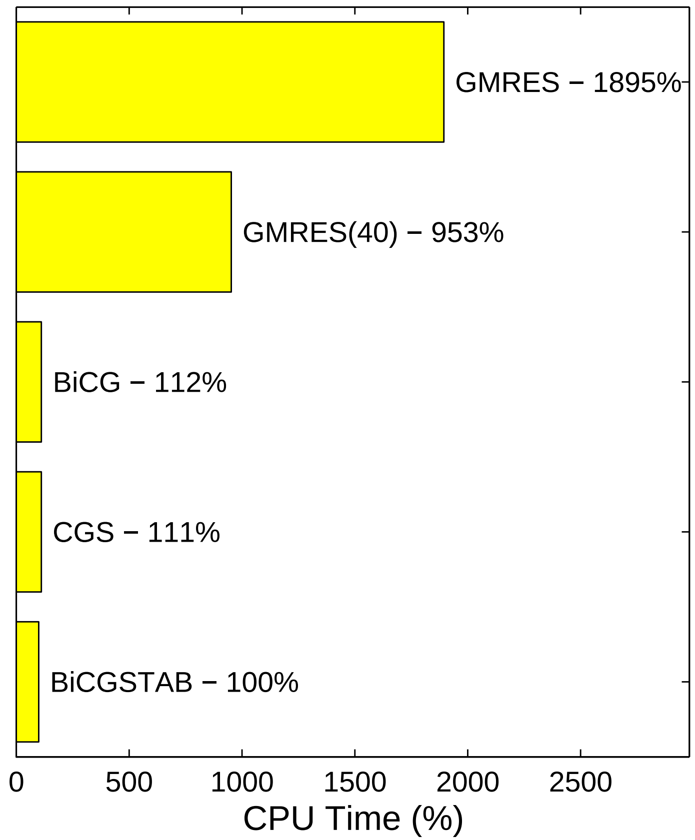




Comparison of GMRES, BiCG, CGS and BiCGSTAB

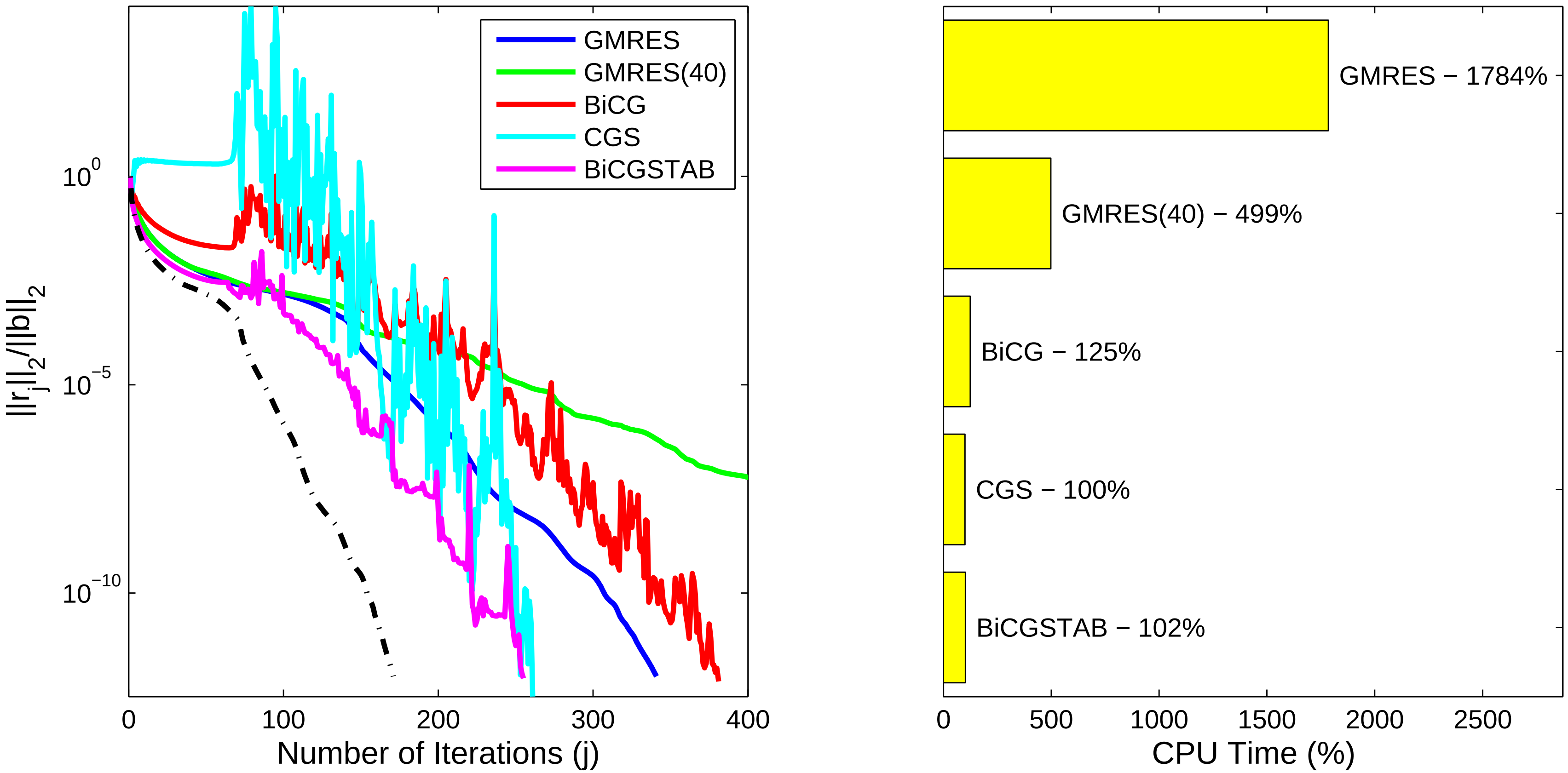
Test 2: Weak Convection-Diffusion ($\alpha = 0.1$, $\epsilon = 1$)





Comparison of GMRES, BiCG, CGS and BiCGSTAB

Test 3: Convection-Diffusion ($\alpha = 1$, $\epsilon = 0.1$)



BiCGSTAB-Algorithm - Summary

Derivation:

- Based on BiCG-Algorithm
- Squaring the polynomial representation
- Residual minimization

Advantages:

- Keenly less storage requirements (compared to GMRES)
- No symmetry constraint on A (compared to CG)
- Requires no multiplications with A^T (compared to BiCG)
- Additional minimization technique (compared to CGS)
  $\rightarrow$ Smooth convergence history

Disadvantages:

- Possible break down due to division by $(A\,p_j, r_0)$