# On the Dubrulle variant of the block CG method


Petr Tichý

Charles University, Prague

joint work with

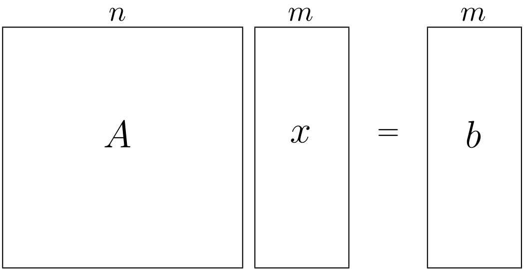
Gérard Meurant


HPCSE 2026

May 18-21, 2026, Soláň

# Problem formulation

$A$ symmetric and positive definite




- Arises naturally in many practical problems [Frommer et al.].
- Accelerate the solution of a single right-hand side system.

# Why blocks?

- Block operations are fast (hardware, libraries).
- A richer search subspace (convergence).

# Block CG algorithm

Introduced by [Dianne O'Leary, 1980].

**input** $A$, $b$, $x_0$
$r_0 = b - Ax_0$
$p_0 = r_0 \, \phi_0$
**for** $k = 1, 2, \dots$ **do**
    $\gamma_{k-1} = \left(p_{k-1}^T A p_{k-1}\right)^{-1} \phi_{k-1}^T r_{k-1}^T r_{k-1}$
    $x_k = x_{k-1} + p_{k-1}\gamma_{k-1}$
    $r_k = r_{k-1} - A p_{k-1}\gamma_{k-1}$
    $\delta_k = \phi_{k-1}^{-1} \left(r_{k-1}^T r_{k-1}\right)^{-1} r_k^T r_k$
    $p_k = (r_k + p_{k-1}\delta_k)\, \phi_k$
**end for**

**nonsingular** $\phi_k \in \mathbb{R}^{m \times m}$ are free to choose.

# Block CG properties

- $A \in \mathbb{R}^{n\times n}$, $v \in \mathbb{R}^{n\times m}$, define **generalized** Krylov subspace

$$\mathcal{K}_k(A, v) \equiv \mathcal{K}_k(A, v^{(1)}) + \cdots + \mathcal{K}_k(A, v^{(m)}).$$

# Block CG properties

- $A \in \mathbb{R}^{n \times n}$, $v \in \mathbb{R}^{n \times m}$, define **generalized** Krylov subspace

$$\mathcal{K}_k(A, v) \equiv \mathcal{K}_k(A, v^{(1)}) + \cdots + \mathcal{K}_k(A, v^{(m)}).$$

- For block CG to work we need the **full rank assumption**

$$\dim \mathcal{K}_k(A, r_0) \; = \; km \, .$$

# Block CG properties

- $A \in \mathbb{R}^{n\times n}$, $v \in \mathbb{R}^{n\times m}$, define **generalized** Krylov subspace

$$\mathcal{K}_k(A, v) \equiv \mathcal{K}_k(A, v^{(1)}) + \cdots + \mathcal{K}_k(A, v^{(m)}).$$

- For block CG to work we need the **full rank assumption**

$$\dim \mathcal{K}_k(A, r_0) \;=\; km\,.$$

- Then it holds that

$$r_k^{(i)} \perp \mathcal{K}_k(A, r_0), \quad p_k^{(i)} \perp_A \mathcal{K}_k(A, r_0)$$

and $x_k^{(i)}$ minimizes

$$\| \, y - x^{(i)} \, \|_A$$

over $y \in x_0^{(i)} + \mathcal{K}_k(A, r_0)$.

# How to deal with rank deficiency?

What to do if the blocks are **singular**?

- deflation [Birk, Frommer, 2014], [Li, Ji, 2017]
- variable block size [Nikishin, Yeremin, 1995, 2003]
- algorithms are complicated

# How to deal with rank deficiency?

What to do if the blocks are **singular**?

- deflation [Birk, Frommer, 2014], [Li, Ji, 2017]
- variable block size [Nikishin, Yeremin, 1995, 2003]
- algorithms are complicated

Problems:

- Which matrix should be considered **rank deficient**?
- Influence of **finite precision** arithmetic?
- Are deflation ideas still applicable?

# How to deal with rank deficiency?

What to do if the blocks are **singular**?

- deflation [Birk, Frommer, 2014], [Li, Ji, 2017]
- variable block size [Nikishin, Yeremin, 1995, 2003]
- algorithms are complicated

Problems:

- Which matrix should be considered **rank deficient**?
- Influence of **finite precision** arithmetic?
- Are deflation ideas still applicable?

It there a light at the end of the tunnel?

# Recall $\rightarrow$ Block CG algorithm

**input** $A$, $b$, $x_0$

$r_0 = b - Ax_0$

$p_0 = r_0 \, \phi_0$

**for** $k = 1, 2, \ldots$ **do**

$\quad \gamma_{k-1} = \left(p_{k-1}^T A p_{k-1}\right)^{-1} \phi_{k-1}^T r_{k-1}^T r_{k-1}$

$\quad x_k = x_{k-1} + p_{k-1}\gamma_{k-1}$

$\quad r_k = r_{k-1} - Ap_{k-1}\gamma_{k-1}$

$\quad \delta_k = \phi_{k-1}^{-1} \left(r_{k-1}^T r_{k-1}\right)^{-1} r_k^T r_k$

$\quad p_k = (r_k + p_{k-1}\delta_k) \, \phi_k$

**end for**

**Interesting idea:** A. A. Dubrulle, *Retooling the method of block conjugate gradients*, Electron. Trans. Numer. Anal. 12, 2001.

# Idea: Change of variables

under the full rank assumption

- Consider a QR factorization of $r_k$ in the form,

$$[w_k, \sigma_k] = \texttt{qr}(r_k),$$

and define formally

$$\phi_k \equiv \sigma_k^{-1}\sigma_{k-1}^{-1}, \quad s_k \equiv p_k\,\sigma_{k-1}\,.$$

# Idea: Change of variables

under the full rank assumption

- Consider a QR factorization of $r_k$ in the form,

$$[w_k, \sigma_k] = \texttt{qr}(r_k),$$

and define formally

$$\phi_k \equiv \sigma_k^{-1}\sigma_{k-1}^{-1}, \quad s_k \equiv p_k\,\sigma_{k-1}\,.$$

- Then $x_k$, $w_k$, and $s_k$ satisfy the recurrences

$$\begin{aligned} \xi_{k-1} &= \left(s_{k-1}^T A s_{k-1}\right)^{-1}, \\ x_k &= x_{k-1} + s_{k-1}\xi_{k-1}\sigma_{k-1}, \\ w_k\zeta_k &= w_{k-1} - A s_{k-1}\xi_{k-1}, \\ s_k &= w_k + s_{k-1}\zeta_k^T, \end{aligned}$$

where

$$\zeta_k = \sigma_k\sigma_{k-1}^{-1} \;\Rightarrow\; \sigma_k = \zeta_k\sigma_{k-1}\,.$$

# DR-BCG

$[w_0, \sigma_0] = \mathtt{qr}(r_0)$
$s_0 = w_0$
**for** $k = 1, 2, \ldots$ **do**
    $\xi_{k-1} = \left(s_{k-1}^T A s_{k-1}\right)^{-1}$
    $x_k = x_{k-1} + s_{k-1}\xi_{k-1}\sigma_{k-1}$
    $[w_k, \zeta_k] = \mathtt{qr}(w_{k-1} - A s_{k-1}\xi_{k-1})$
    $s_k = w_k + s_{k-1}\zeta_k^T$
    $\sigma_k = \zeta_k \sigma_{k-1}$    % recall $r_k = w_k \sigma_k$
**end for**

# DR-BCG

1: $[w_0, \sigma_0] = \texttt{qr}(r_0)$
2: $s_0 = w_0$
3: **for** $k = 1, 2, \ldots$ **do**
4: $\quad \xi_{k-1} = \left(s_{k-1}^T A s_{k-1}\right)^{-1}$
5: $\quad x_k = x_{k-1} + s_{k-1}\xi_{k-1}\sigma_{k-1}$
6: $\quad [w_k, \zeta_k] = \texttt{qr}(w_{k-1} - A s_{k-1}\xi_{k-1})$
7: $\quad s_k = w_k + s_{k-1}\zeta_k^T$
8: $\quad \sigma_k = \zeta_k \sigma_{k-1}$ % recall $r_k = w_k \sigma_k$
9: **end for**

On line 6, we use the MATLAB command

$$[\texttt{w}, \texttt{zeta}] = \texttt{qr}(\texttt{v}, \texttt{"econ"}),$$

$w_k \in \mathbb{R}^{n \times m}$ has orthonormal columns and $\zeta_k$ **can be singular!**

# Properties of DR-BCG

## without the full rank assumption

In [Meurant, T., 2026] we show that

- $\left(s_k^T A s_k\right)$ is always **nonsingular**!

# Properties of DR-BCG

without the full rank assumption

In [Meurant, T., 2026] we show that

- $\left(s_k^T A s_k\right)$ is always **nonsingular**!
- **Local** orthogonality is always **preserved**,

$$r_k^T s_{k-1} = 0, \quad \text{where} \quad r_k = w_k \sigma_k \,.$$

# Properties of DR-BCG

## without the full rank assumption

In [Meurant, T., 2026] we show that

- $\left(s_k^T A s_k\right)$ is always **nonsingular**!
- **Local** orthogonality is always **preserved**,

$$r_k^T s_{k-1} = 0, \quad \text{where} \quad r_k = w_k \sigma_k \,.$$

- DR-BCG **converges** for $k \to \infty$.

# Properties of DR-BCG

without the full rank assumption

In [Meurant, T., 2026] we show that

- $\left(s_k^T A s_k\right)$ is always **nonsingular**!
- **Local** orthogonality is always **preserved**,

$$r_k^T s_{k-1} = 0, \quad \text{where} \quad r_k = w_k \sigma_k \, .$$

- DR-BCG **converges** for $k \to \infty$ .
- **Global** orthogonality is **lost** in the rank-deficient case.

# Properties of DR-BCG

without the full rank assumption

In [Meurant, T., 2026] we show that

- $\left(s_k^T A s_k\right)$ is always **nonsingular**!
- **Local** orthogonality is always **preserved**,

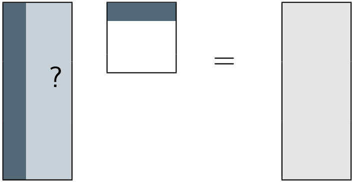
$$r_k^T s_{k-1} = 0, \quad \text{where} \quad r_k = w_k \sigma_k \, .$$

- DR-BCG **converges** for $k \to \infty$ .
- **Global** orthogonality is **lost** in the rank-deficient case.
- We show, how to incorporate **preconditioning** .

# Mystery of DR-BCG

without the full rank assumption

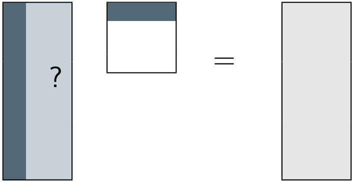
$$[w, \zeta] = \texttt{qr}(v).$$



# Mystery of DR-BCG

## without the full rank assumption

$$[w, \zeta] = \mathtt{qr}(v).$$




- We do not have to use `qr`.
- Columns of $w$ from $\mathcal{N}(v^T)$ can be **arbitrary**.
- $\mathcal{K}_k(A, r_0) \subseteq \operatorname{colspan}\{w_0, \dots, w_{k-1}\}$.
- **Observation**: The solution is found if $\mathcal{K}_k(A, r_0)$ $A$-invariant!

# Test problems

## SuiteSparse Matrix collection

| problem | $n$ | $m$ | $\kappa(A)$ | $b$ | precond |
|---|---|---|---|---|---|
| bcsstk03 | 112 | 4 | $10^6$ | rand | no |
| s3dkt3m2 | 90 449 | 1…128 | $10^{11}$ | rand | ✓ |

# Test problems

## SuiteSparse Matrix collection

| problem | $n$ | $m$ | $\kappa(A)$ | $b$ | precond |
|---|---|---|---|---|---|
| bcsstk03 | 112 | 4 | $10^6$ | rand | no |
| s3dkt3m2 | 90 449 | 1...128 | $10^{11}$ | rand | ✓ |

- We plot an analogue of the relative $A$-norm of the error

$$\left[\frac{\mathrm{trace}((x - x_k)^T\, A\, (x - x_k))}{\mathrm{trace}(x^T A x)}\right]^{1/2}$$

# Test problems

## SuiteSparse Matrix collection

| problem | $n$ | $m$ | $\kappa(A)$ | $b$ | precond |
|---|---|---|---|---|---|
| bcsstk03 | 112 | 4 | $10^6$ | rand | no |
| s3dkt3m2 | 90 449 | $1 \dots 128$ | $10^{11}$ | rand | ✓ |

- We plot an analogue of the relative $A$-norm of the error

$$\left[\frac{\mathrm{trace}((x - x_k)^T A (x - x_k))}{\mathrm{trace}(x^T A x)}\right]^{1/2}$$

- Methods:
  - **HS-BCG** [O'Leary, 1980]
  - **DR-BCG** [Dubrulle, 2001], [Meurant, T., 2026]
  - **BF-BCG** [Ji, Li, 2017]

# $\texttt{bcsstk03}$, $n = 112$, $m = 4$

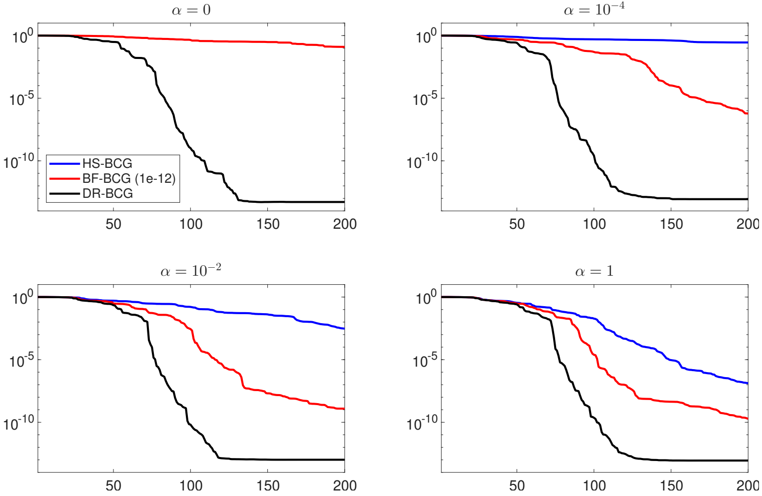
methods and rank-deficiency

$$b \;=\; (1 - \alpha)\,[c, \ldots, c] \;+\; \alpha\,\texttt{rand}(n, m)$$

# $\texttt{bcsstk03}$, $n = 112$, $m = 4$

methods and rank-deficiency

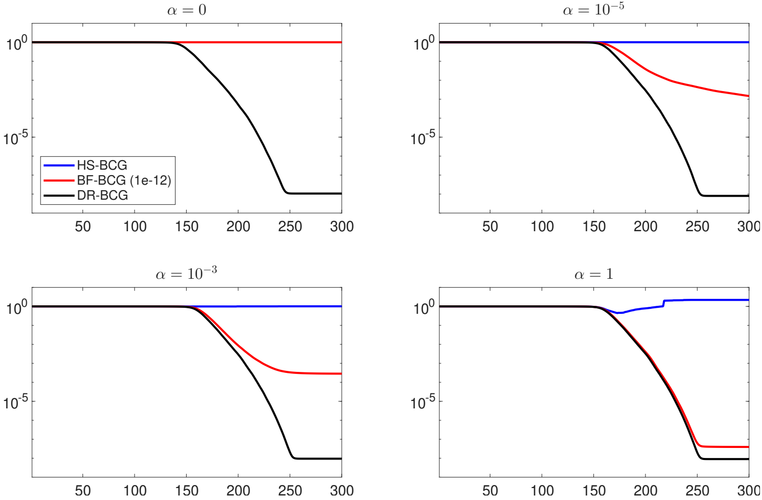
$$b = (1-\alpha)\,[c,\ldots,c] + \alpha\,\texttt{rand}(n,m)$$



# $\texttt{s3dkt3m2}$, $n = 90\,449$, $m = 32$

methods and rank-deficiency

$$b \;=\; (1 - \alpha)\,[c, \ldots, c] \;+\; \alpha\,\texttt{rand}(n, m)$$



# DR-BCG for `s3dkt3m2`, $n = 90\,449$

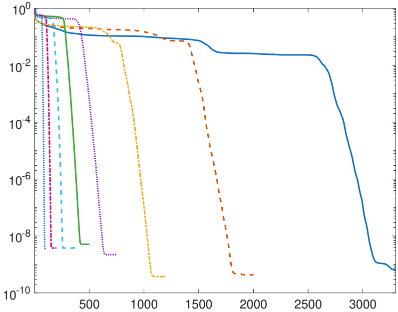
## Convergence and time for various number of right hand sides

$$b \;=\; \texttt{rand}(n, m), \quad m = 1, 2, 4, \ldots, 128$$

# DR-BCG for `s3dkt3m2`, $n = 90\,449$

Convergence and time for various number of right hand sides

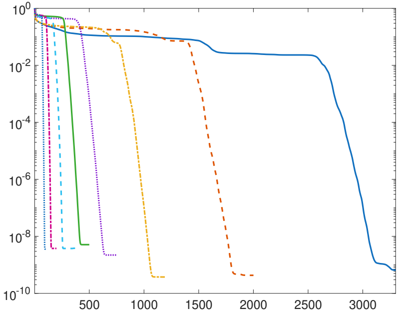
$$b \ = \ \texttt{rand}(n, m), \quad m = 1, 2, 4, \ldots, 128$$

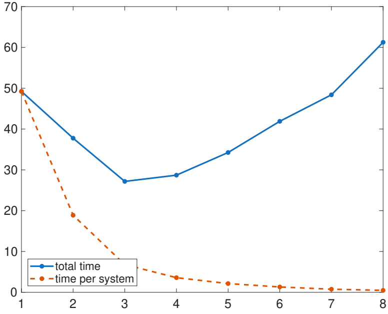


# DR-BCG for `s3dkt3m2`, $n = 90\,449$

Convergence and time for various number of right hand sides

$$b \;=\; \texttt{rand}(n, m), \quad m = 1, 2, 4, \ldots, 128$$




left: convergence curves for $b$ of size $n \times m$ with $m = 1, 2, 4, \ldots, 128$
right: the total time and the time per single system (in seconds)

# Conclusions

- **DR-BCG** is **superior** to other BCG variants.

  It consistently outperforms other block CG algorithms, both in exact arithmetic and in finite precision computations, no problems with rank deficiency.

# Conclusions

- **DR-BCG** is **superior** to other BCG variants.

  It consistently outperforms other block CG algorithms, both in exact arithmetic and in finite precision computations, no problems with rank deficiency.

- **Algorithmically**, it is very **simple**.

  It does not require complicated deflation procedures or detection of numerically linearly dependent vectors. A natural generalization of CG to the block case.

# Conclusions

- **DR-BCG** is **superior** to other BCG variants.

  It consistently outperforms other block CG algorithms, both in exact arithmetic and in finite precision computations, no problems with rank deficiency.

- **Algorithmically**, it is very **simple**.

  It does not require complicated deflation procedures or detection of numerically linearly dependent vectors. A natural generalization of CG to the block case.

- A comprehensive **theory** is still **missing**.

  We can only prove asymptotic convergence, not the convergence in a finite number of iterations. How the added vectors influence convergence?

# Conclusions

- **DR-BCG** is **superior** to other BCG variants.

  It consistently outperforms other block CG algorithms, both in exact arithmetic and in finite precision computations, no problems with rank deficiency.

- **Algorithmically**, it is very **simple**.

  It does not require complicated deflation procedures or detection of numerically linearly dependent vectors. A natural generalization of CG to the block case.

- A comprehensive **theory** is still **missing**.

  We can only prove asymptotic convergence, not the convergence in a finite number of iterations. How the added vectors influence convergence?

- "**Dubrullization**" of block methods deserves **more attention**.

  Dubrulle's approach yields algorithms whose mathematical properties have not yet been systematically studied or described.

# Related papers

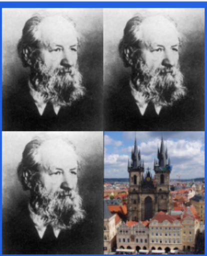

- G. Meurant and P. Tichý, *Dubrulle's variant of the block conjugate gradient algorithm*, submitted, January, 2026.
- G. Meurant, J. Papež, and P. Tichý, *Block conjugate gradient methods with error norm estimates for least squares problems*, BIT 66(2), 2026.
- G. Meurant and P. Tichý, *Error norm estimates for the block conjugate gradient algorithm*, SIMAX 46(4), 2025.
- P. Tichý, G. Meurant, D. Šimonová, *Block CG algorithms revisited*, Numer. Algorithms 100, 2025.

- D. P. O'Leary, *The block conjugate gradient algorithm and related methods*, Linear Algebra Appl. 29, 1980.
- A. A. Dubrulle, *Retooling the method of block conjugate gradients*, Electron. Trans. Numer. Anal. 12, 2001.
- S. Birk and A. Frommer, *A deflated CG method for multiple right hand sides and multiple shifts*, Numer. Algorithms 67, 2014.
- H. Ji and Y. Li, *A breakdown-free block CG method*, BIT 57, 2017.

# Recent Developments in Block Krylov Methods

Theory, Applications, and Implementation

📅 September 7-8, 2026 📍 Prague, Czech Republic



## About

This workshop focuses on recent developments in block Krylov subspace methods, covering theoretical advances, applications, and practical implementation aspects. It aims to bring together researchers and practitioners and to provide a platform for discussing emerging challenges and opportunities.

`https://workshop.math.cas.cz/BlockKrylov/`

organized by Jan Papež and Petr Tichý

# Breakdown-free version of Ji and Li

## A version with deflation

[Ji & Li 2017]

$p_0 = \texttt{orthog}(r_0)$
**for** $k = 1, 2, \ldots$ **do**
  $\gamma_{k-1} = \left(p_{k-1}^T A p_{k-1}\right)^{-1} p_{k-1}^T r_{k-1}$
  $x_k = x_{k-1} + p_{k-1}\gamma_{k-1}$
  $r_k = r_{k-1} - A p_{k-1}\gamma_{k-1}$
  $\delta_k = -\left(p_{k-1}^T A p_{k-1}\right)^{-1} p_{k-1}^T A r_k$
  $p_k = \texttt{orthog}\left(r_k + p_{k-1}\delta_k\right)$
**end for**

We determine the orthonormal basis using `orthog` as the left singular vectors corresponding to singular values whose relative size is greater than a given tolerance.