

# Block cross-interactive residual smoothing for Lanczos-type solvers for linear systems with multiple right-hand sides

*Kensuke Aihara, Akira Imakura, Keiichi Morikuni*

Abstract

Block Lanczos-type solvers, such as the block BiCGSTAB method [3], for large sparse linear systems

$$AX = B, \quad A \in \mathbb{R}^{n \times n}, \quad B \in \mathbb{R}^{n \times s}, \quad s \ll n$$

often exhibit large oscillations in the residual norms. In finite precision arithmetic, the large oscillations lead to a large residual gap (the difference  $G_{R_k}$  between the recursively updated residual  $R_k$  and the explicitly computed residual  $B - AX_k$ ) and a loss of attainable accuracy of the approximations, as observed in

$$\|G_{R_k}\| - \|R_k\| \leq \|B - AX_k\| \leq \|G_{R_k}\| + \|R_k\|, \quad G_{R_k} := (B - AX_k) - R_k.$$

This problem is addressed by using cross-interactive residual smoothing (CIRS). Just as the standard Lanczos-type solvers for a single linear system have been extended to their global and block versions for solving systems with multiple right-hand sides, similar extensions of CIRS are naturally considered. While we have developed the global CIRS scheme (Gl-CIRS) in our previous study [1], we propose a block version (Bl-CIRS) in this study. Then, we demonstrate the effectiveness of Bl-CIRS from various perspectives, such as theoretical insights into the convergence behaviors of the residual and approximation norms, numerical experiments on model problems, and a detailed rounding error analysis for the residual gap. In particular, we show for the case of Bl-CIRS that orthonormalizing the columns of direction matrices plays an important role in reducing the residual gap. The presented analysis also complements our previous study above that includes an evaluation for the residual gap of the block Lanczos-type solvers.

**Advances in residual smoothing** Block Lanczos-type solvers typically update the  $k$ th approximation  $X_k$  and residual  $R_k$  by using the recursion formulas

$$X_{k+1} = X_k + P_k \alpha_k^\square, \quad R_{k+1} = R_k - (AP_k) \alpha_k^\square, \quad k = 0, 1, \dots,$$

respectively, where  $P_k \in \mathbb{R}^{n \times s}$  is a direction matrix and  $\alpha_k^\square \in \mathbb{R}^{s \times s}$  is determined under a certain condition. Residual smoothing was introduced by Schönauer [7] to Lanczos-type solvers for a single linear system to get a non-increasing sequence of residual norms [8]. A block version of the simple residual smoothing (Bl-SRS) was presented by Jbilou [5]. Let  $X_k$  and  $R_k$  be the  $k$ th primary approximation and residual, respectively. Then, new sequences of approximations  $Y_k$  and the corresponding smoothed residuals  $S_k (:= B - AY_k)$  are generated by the recursion formulas

$$Y_{k+1} = Y_k + (X_{k+1} - Y_k) \eta_{k+1}^\square, \quad S_{k+1} = S_k + (R_{k+1} - S_k) \eta_{k+1}^\square, \quad k = 0, 1, \dots,$$

respectively, where  $Y_0 = X_0$ ,  $S_0 = R_0$ , and  $\eta_k^\square \in \mathbb{R}^{s \times s}$  is a parameter matrix. With a local minimization of the smoothed residual norm in choosing  $\eta_k^\square$ , a monotonically decreasing sequence of  $\|S_k\|$  is obtained.

Studies on the relationship between residual smoothing and the residual gap have an interesting history. For a single right-hand side case, Gutknecht and Rozložník [4] clarified that the conventional

Table 1: Difference in the recursion formulas for updating smoothed residuals.

Type	SRS scheme	CIRS scheme
Standard $A\mathbf{x} = \mathbf{b}$	[7, 8] $\mathbf{s}_k = \mathbf{s}_{k-1} + \eta_k(\mathbf{r}_k - \mathbf{s}_{k-1}),$ $\mathbf{r}_k, \mathbf{s}_{k-1} \in \mathbb{R}^n, \quad \eta_k \in \mathbb{R}$	[2, 6] $\mathbf{s}_k = \mathbf{s}_{k-1} - \eta_k A\mathbf{v}_k,$ $\mathbf{v}_k, \mathbf{s}_{k-1} \in \mathbb{R}^n, \quad \eta_k \in \mathbb{R}$
Global $AX = B$	[9] $S_k = S_{k-1} + \eta_k(R_k - S_{k-1}),$ $R_k, S_{k-1} \in \mathbb{R}^{n \times s}, \quad \eta_k \in \mathbb{R}$	[1] $S_k = S_{k-1} - \eta_k AV_k,$ $V_k, S_{k-1} \in \mathbb{R}^{n \times s}, \quad \eta_k \in \mathbb{R}$
Block $AX = B$	[5] $S_k = S_{k-1} + (R_k - S_{k-1})\eta_k^\square,$ $R_k, S_{k-1} \in \mathbb{R}^{n \times s}, \quad \eta_k^\square \in \mathbb{R}^{s \times s}$	Present study $S_k = S_{k-1} - (A\tilde{Q}_k)\tilde{\eta}_k^\square,$ $\tilde{Q}_k, S_{k-1} \in \mathbb{R}^{n \times s}, \quad \tilde{\eta}_k^\square \in \mathbb{R}^{s \times s}$

smoothing schemes (including the Zhou–Walker implementation [10]) do not help to improve the attainable accuracy. To be more specific, rounding errors accumulated in the primary sequences propagate to the smoothed sequences, and the smoothed true residual norms stagnate at the same order of magnitude as the primary ones. In order to remedy this phenomenon, Komeyama et al. [2, 6] modified the Zhou–Walker implementation so that the primary and smoothed sequences influence one another. This modification is referred to as cross-interactive residual smoothing (CIRS) and is indeed effective in reducing the residual gap and increasing attainable accuracy. As SRS has been extended to global and block versions [9, 5], CIRS is also extended. In this perspective, our previous study [1] presented a global version of CIRS (GI-CIRS) for the global- and block-type solvers, and therefore, we propose a block version of CIRS (BI-CIRS) in this study. Table 1 summarizes the recursion formulas for the aforementioned residual smoothing schemes. This table shows that this study fills a gap in the literature of the CIRS schemes.

**Block cross-interactive residual smoothing** This study proposes updating approximations and the corresponding residuals by the recursion formulas

$$\begin{aligned} \text{smoothed} \quad Y_{k+1} &= Y_k + V_{k+1}\eta_{k+1}^\square, & S_{k+1} &= S_k - (AV_{k+1})\eta_{k+1}^\square, \\ \text{primary} \quad X_{k+1} &= Y_{k+1} + V_{k+1}(I_s - \eta_{k+1}^\square), & R_{k+1} &= S_{k+1} - (AV_{k+1})(I_s - \eta_{k+1}^\square), \end{aligned}$$

respectively, for  $k = 0, 1, \dots$  with  $Y_0 = X_0$  and  $S_0 = R_0$  so that the primary and smoothed sequences influence one another, where  $V_{k+1} = V_k(I_s - \eta_k^\square) + \tilde{P}_k$  is an auxiliary matrix for  $\eta_0^\square = O \in \mathbb{R}^{s \times s}$  and  $V_0 = O \in \mathbb{R}^{n \times s}$ . Here,  $\tilde{P}_k \in \mathbb{R}^{n \times s}$  is a direction matrix in the recursion formula  $X_{k+1} = X_k + \tilde{P}_k$ . Again with a local minimization of the smoothed residual norm in choosing  $\eta_k^\square$ , a monotonically decreasing sequence of  $\|S_k\|$  is obtained. Note that the essential difference of BI-CIRS from GI-CIRS [1, Algorithm 3.1] is that the smoothing parameter  $\eta_k^\square$  of BI-CIRS is an  $s$ -by- $s$  matrix instead of a scalar.

For numerical stability, BI-CIRS needs to orthonormalize the columns of the auxiliary matrix  $V_k$  for each iteration. Let  $V_k = \tilde{Q}_k\tilde{\xi}_k$  be the QR decomposition of  $V_k$ , where  $\tilde{Q}_k \in \mathbb{R}^{n \times s}$  and  $\tilde{\xi}_k \in \mathbb{R}^{s \times s}$  are the Q- and R-factors, respectively. With the auxiliary matrix  $\tilde{\eta}_k^\square := \tilde{\xi}_k\eta_k^\square$ , the above formulas are equivalently rewritten as

$$\begin{aligned} \text{smoothed} \quad Y_{k+1} &= Y_k + \tilde{Q}_{k+1}\tilde{\eta}_{k+1}^\square, & S_{k+1} &= S_k - (A\tilde{Q}_{k+1})\tilde{\eta}_{k+1}^\square, \\ \text{primary} \quad X_{k+1} &= Y_{k+1} + \tilde{Q}_{k+1}(\tilde{\xi}_{k+1} - \tilde{\eta}_{k+1}^\square), & R_{k+1} &= S_{k+1} - (A\tilde{Q}_{k+1})(\tilde{\xi}_{k+1} - \tilde{\eta}_{k+1}^\square) \end{aligned}$$

together with  $V_{k+1} = \tilde{Q}_k(\tilde{\xi}_k - \tilde{\eta}_k^\square) + \tilde{P}_k$  for  $k = 0, 1, \dots$ .

Our main results via a rounding error analysis shows that BI-CIRS with the orthonormalization strategy suppresses the residual gap.

**Theorem 1.** *In finite precision arithmetic, let  $X_k \in \mathbb{F}^{n \times s}$  and  $R_k \in \mathbb{F}^{n \times s}$  be the  $k$ th approximation and residual generated by the recursion formulas*

$$X_k = X_{k-1} + \widehat{Q}_{k-1} \alpha_{k-1}^\square, \quad R_k = R_{k-1} - (A \widehat{Q}_{k-1}) \alpha_{k-1}^\square, \quad k = 0, 1, \dots, \quad (1)$$

respectively, where  $\mathbb{F} \subset \mathbb{R}$  is a set of floating point numbers and  $\widehat{Q}_{k-1}$  is a  $Q$ -factor of the direction matrix  $P_{k-1}$  obtained from the QR decomposition with Givens rotations or Householder transformations. Then, with  $X_0 = O$ , the residual gap  $G_{R_k}$  satisfies the bound

$$\|G_{R_k}\| < (8\sqrt{s}\gamma_{m+3s} + \gamma_1) k \|A\| \max_{0 < i \leq k} \|X_i\| + k\gamma_1 \max_{0 < i \leq k} \|R_i\|,$$

where  $\gamma_k := k\mathbf{u}/(1 - k\mathbf{u})$  with a unit roundoff  $\mathbf{u}$  and  $m$  is the maximum number of nonzero entries per row of  $A$ .

In the case of BI-CIRS, replacing  $X_k$  and  $R_k$  by  $Y_k$  and  $S_k$ , respectively, in (1), this theorem holds for the residual gap  $G_{S_k} = (B - AY_k) - S_k$ . Therefore, even when using an inexact orthonormalization for the columns of iteration matrices, BI-CIRS in which the residual and approximation norms converge smoothly is indeed useful to reduce the residual gap. This theoretical result is consistent with our numerical results. Numerical experiments demonstrate that BI-CIRS is effective for suppressing the residual gap and improving the attainable accuracy of the approximations.

## References

- [1] K. AIHARA, A. IMAKURA, AND K. MORIKIUNI, *Cross-interactive residual smoothing for global and block Lanczos-type solvers for linear systems with multiple right-hand sides*, SIAM J. Matrix Anal. Appl., 43 (2022), pp. 1308–1330.
- [2] K. AIHARA, R. KOMEYAMA, AND E. ISHIWATA, *Variants of residual smoothing with a small residual gap*, BIT, 59 (2019), pp. 565–584.
- [3] A. EL GUENNOUNI, K. JBILOU, AND H. SADOK, *A block version of BiCGSTAB for linear systems with multiple right-hand sides*, Electron. Trans. Numer. Anal., 16 (2003), pp. 129–142.
- [4] M. H. GUTKNECHT AND M. ROZLOŽNÍK, *Residual smoothing techniques: Do they improve the limiting accuracy of iterative solvers?*, BIT, 41 (2001), pp. 86–114.
- [5] K. JBILOU, *Smoothing iterative block methods for linear systems with multiple right-hand sides*, J. Comput. Appl. Math., 107 (1999), pp. 97–109.
- [6] R. KOMEYAMA, K. AIHARA, AND E. ISHIWATA, *Reconsideration of residual smoothing technique for improving the accuracy of approximate solutions of CGS-type iterative methods*, Trans. Japan Soc. Ind. Appl. Math., 28 (2018), pp. 18–38.
- [7] W. SCHÖNAUER, *Scientific Computing on Vector Computers*, Elsevier, Amsterdam, 1987.
- [8] R. WEISS, *Parameter-free iterative linear solvers*, Akademie Verlag, Berlin, 1996.

- [9] J. ZHANG AND F. DAI, *Global CGS algorithm for linear systems with multiple right-hand sides*, Numer. Math. J. Chinese Univ., 30 (2008), pp. 390–399.
- [10] L. ZHOU AND H. F. WALKER, *Residual smoothing techniques for iterative methods*, SIAM J. Sci. Comput., 15 (1994), pp. 297–312.