

DFG/RFBR project  
“Tensor methods in multi-dimensional spectral problems  
with particular applications in electronic structure calculations”

## FAST AND ACCURATE TENSOR RECOMPRESSION USING WEDDERBURN RANK REDUCTION AND KRYLOV SUBSPACES

DMITRY SAVOSTYANOV  
*joint work with S. GOREINOV, I. OSELEDETS, N. ZAMARASHKIN and EUGENE TYRTYSHNIKOV*  
*Institute of Numerical Mathematics, Russian Academy of Sciences*



Leipzig, 04/05/2010

## WEDDERBURN RANK REDUCTION [1934]

$$B = A - \frac{Ayx^TA}{x^TAy}.$$



Joseph Henry Maclagan Wedderburn (1882–1948)

## WEDDERBURN RANK REDUCTION

$$B = A - \frac{Ayx^TA}{x^TAy}, \quad x^TAy \neq 0$$

## WEDDERBURN RANK REDUCTION

$$B = A - \frac{Ayx^T A}{x^T A y}, \quad x^T A y \neq 0$$

$$By = Ay - \frac{Ayx^T A}{x^T A y}y = Ay - \frac{Ayx^T A y}{x^T A y} = 0$$

$$x^T B = x^T A - x^T \frac{Ayx^T A}{x^T A y} = x^T A - \frac{x^T A y x^T A}{x^T A y} = 0$$

$$\text{rank } B = \text{rank } A - 1$$

WEDDERBURN SEQUENCE [Chu, Funderlic, Golub, 1995]

$$A_k = A_{k-1} - \frac{A_{k-1}y_kx_k^T A_{k-1}}{x_k^T A_{k-1}y_k}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1}y_k \neq 0$$

## WEDDERBURN SEQUENCE

$$A_k = A_{k-1} - \frac{A_{k-1}y_k x_k^T A_{k-1}}{x_k^T A_{k-1} y_k}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

STATEMENT 1

$$A_k = P_k^T A = A Q_k$$

$$P_0 = I, \quad Q_0 = I$$

$$P_1^T = I - \omega_1^{-1} A y_1 x_1^T$$

$$Q_1 = I - \omega_1^{-1} y_1 x_1^T A$$

## WEDDERBURN SEQUENCE

$$A_k = A_{k-1} - \frac{A_{k-1}y_k x_k^T A_{k-1}}{x_k^T A_{k-1} y_k}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

## STATEMENT 1

$$A_k = P_k^T A = A Q_k$$

$$P_0 = I, \quad Q_0 = I$$

$$P_k = P_{k-1} - \omega_k^{-1} P_{k-1} x_k y_k^T A^T P_{k-1}$$

$$Q_k = Q_{k-1} - \omega_k^{-1} Q_{k-1} y_k x_k^T A Q_{k-1}$$

## COROLLARY

$$P_k x_k = 0, \quad Q_k y_k = 0$$

## WEDDERBURN SEQUENCE

$$A_k = A_{k-1} - \frac{A_{k-1}y_k x_k^T A_{k-1}}{x_k^T A_{k-1} y_k}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

$$A_k = P_k^T A = A Q_k$$

$$P_0 = I, \quad P_k = P_{k-1} - \omega_k^{-1} P_{k-1} x_k y_k^T A^T P_{k-1}$$

$$Q_0 = I, \quad Q_k = Q_{k-1} - \omega_k^{-1} Q_{k-1} y_k x_k^T A Q_{k-1}$$

**STATEMENT 2:**  $P_k, Q_k$  are projectors

$$P_k^2 = P_k \quad Q_k^2 = Q_k$$

## WEDDERBURN SEQUENCE

$$A_k = A_{k-1} - \frac{A_{k-1}y_k x_k^T A_{k-1}}{x_k^T A_{k-1} y_k}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

$$A_k = P_k^T A = A Q_k$$

$$P_0 = I, \quad P_k = P_{k-1} - \omega_k^{-1} P_{k-1} x_k y_k^T A^T P_{k-1}$$

$$Q_0 = I, \quad Q_k = Q_{k-1} - \omega_k^{-1} Q_{k-1} y_k x_k^T A Q_{k-1}$$

**STATEMENT 2:**  $P_k, Q_k$  are projectors

$$P_k^2 = P_k \quad Q_k^2 = Q_k$$

$$\begin{aligned} P_k^2 &= P_{k-1}^2 - \omega_k^{-1} P_{k-1}^2 y_s x_s^T A^T P_{k-1} - \omega_k^{-1} P_{k-1} x_s y_s^T A^T P_{k-1}^2 + \\ &\quad + \omega_k^{-2} P_{k-1} x_k \underbrace{y_k^T A^T P_{k-1}^2 x_k}_{\omega_k} y_k^T A^T P_{k-1} \end{aligned}$$

## WEDDERBURN SEQUENCE

$$A_k = A_{k-1} - \frac{A_{k-1}y_k x_k^T A_{k-1}}{x_k^T A_{k-1} y_k}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

$$A_k = P_k^T A = A Q_k$$

$$P_0 = I, \quad P_k = P_{k-1} - \omega_k^{-1} P_{k-1} x_k y_k^T A^T P_{k-1}; \quad P_k^2 = P_k$$

$$Q_0 = I, \quad Q_k = Q_{k-1} - \omega_k^{-1} Q_{k-1} y_k x_k^T A Q_{k-1}; \quad Q_k^2 = Q_k$$

**STATEMENT 3:**  $P_k P_m = P_{\max(k,m)}$ ,  $Q_k Q_m = Q_{\max(k,m)}$

$$Q_k Q_{k-1} = Q_{k-1}^2 - \omega_k^{-1} Q_{k-1} y_k x_k^T A Q_{k-1}^2 = Q_{k-1} - \omega_k^{-1} Q_{k-1} y_k x_k^T A Q_{k-1} = Q_k$$

## WEDDERBURN SEQUENCE

$$A_k = A_{k-1} - \frac{A_{k-1}y_k x_k^T A_{k-1}}{x_k^T A_{k-1} y_k}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

$$A_k = P_k^T A = A Q_k$$

$$P_0 = I, \quad P_k = P_{k-1} - \omega_k^{-1} P_{k-1} x_k y_k^T A^T P_{k-1}; \quad P_k^2 = P_k$$

$$Q_0 = I, \quad Q_k = Q_{k-1} - \omega_k^{-1} Q_{k-1} y_k x_k^T A Q_{k-1}; \quad Q_k^2 = Q_k$$

**STATEMENT 3:**  $P_k P_m = P_{\max(k,m)}$ ,  $Q_k Q_m = Q_{\max(k,m)}$

$$Q_k Q_{k-1} = Q_{k-1}^2 - \omega_k^{-1} Q_{k-1} y_k x_k^T A Q_{k-1}^2 = Q_{k-1} - \omega_k^{-1} Q_{k-1} y_k x_k^T A Q_{k-1} = Q_k$$

## COROLLARY

$$P_k[x_1 \ x_2 \ \dots \ x_k] = 0, \quad Q_k[y_1 \ y_2 \ \dots \ y_k] = 0$$

## WEDDERBURN SEQUENCE

$$A_k = A_{k-1} - \frac{A_{k-1}y_k x_k^T A_{k-1}}{x_k^T A_{k-1} y_k}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

$$A_k = P_k^T A = A Q_k$$

$$P_0 = I, \quad P_k = P_{k-1} - \omega_k^{-1} P_{k-1} x_k y_k^T A^T P_{k-1}; \quad P_k P_m = P_{\max(k,m)}$$

$$Q_0 = I, \quad Q_k = Q_{k-1} - \omega_k^{-1} Q_{k-1} y_k x_k^T A Q_{k-1}; \quad Q_k Q_m = Q_{\max(k,m)}$$

**STATEMENT 4:**  $u_k \stackrel{\text{def}}{=} P_{k-1} x_k$  and  $v_k \stackrel{\text{def}}{=} Q_{k-1} y_k$  provide  $U_k^T A V_k = \Omega_k$

$$u_k^T A v_k = x_k^T \underbrace{P_{k-1}^T A Q_{k-1}}_{=A_{k-1}Q_{k-1}=AQ_{k-1}^2=AQ_{k-1}=A_{k-1}} y_k = x_k^T A_{k-1} y_k \stackrel{\text{def}}{=} \omega_k$$

## WEDDERBURN SEQUENCE

$$A_k = A_{k-1} - \frac{A_{k-1}y_k x_k^T A_{k-1}}{x_k^T A_{k-1} y_k}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

$$A_k = P_k^T A = A Q_k$$

$$P_0 = I, \quad P_k = P_{k-1} - \omega_k^{-1} P_{k-1} x_k y_k^T A^T P_{k-1}; \quad P_k P_m = P_{\max(k,m)}$$

$$Q_0 = I, \quad Q_k = Q_{k-1} - \omega_k^{-1} Q_{k-1} y_k x_k^T A Q_{k-1}; \quad Q_k Q_m = Q_{\max(k,m)}$$

**STATEMENT 4:**  $u_k \stackrel{\text{def}}{=} P_{k-1} x_k$  and  $v_k \stackrel{\text{def}}{=} Q_{k-1} y_k$  provide  $U_k^T A V_k = \Omega_k$

$$u_k^T A v_k = x_k^T P_{k-1}^T A Q_{k-1} y_k = x_k^T A_{k-1} y_k \stackrel{\text{def}}{=} \omega_k$$

$$u_m^T A v_k = x_m^T \underbrace{P_{m-1}^T A Q_{k-1}}_{= A Q_{m-1} Q_{k-1} = A Q_{\max(m-1, k-1)}} y_k = 0 \quad \text{for } m \neq k.$$

## WEDDERBURN SEQUENCE

$$A_k = A_{k-1} - \frac{A_{k-1}y_k x_k^T A_{k-1}}{x_k^T A_{k-1} y_k}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

$$A_k = P_k^T A = A Q_k$$

$$P_0 = I, \quad P_k = P_{k-1} - \omega_k^{-1} P_{k-1} x_k y_k^T A^T P_{k-1}; \quad P_k P_m = P_{\max(k,m)}$$

$$Q_0 = I, \quad Q_k = Q_{k-1} - \omega_k^{-1} Q_{k-1} y_k x_k^T A Q_{k-1}; \quad Q_k Q_m = Q_{\max(k,m)}$$

**STATEMENT 4:**  $u_k \stackrel{\text{def}}{=} P_{k-1} x_k$  and  $v_k \stackrel{\text{def}}{=} Q_{k-1} y_k$  provide  $U_k^T A V_k = \Omega_k$

**COROLLARY:**  $U_k$  and  $V_k$  have full rank

## WEDDERBURN SEQUENCE

$$A_k = A_{k-1} - \frac{A_{k-1}y_k x_k^T A_{k-1}}{x_k^T A_{k-1} y_k}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

$$A_k = P_k^T A = A Q_k$$

$$P_0 = I, \quad P_k = P_{k-1} - \omega_k^{-1} P_{k-1} x_k y_k^T A^T P_{k-1}; \quad P_k P_m = P_{\max(k,m)}$$

$$Q_0 = I, \quad Q_k = Q_{k-1} - \omega_k^{-1} Q_{k-1} y_k x_k^T A Q_{k-1}; \quad Q_k Q_m = Q_{\max(k,m)}$$

**STATEMENT 4:**  $u_k \stackrel{\text{def}}{=} P_{k-1} x_k$  and  $v_k \stackrel{\text{def}}{=} Q_{k-1} y_k$  provide  $U_k^T A V_k = \Omega_k$

**COROLLARY:**  $U_k$  and  $V_k$  have full rank

**COROLLARY:**

$$P_k[u_1 \ u_2 \ \dots \ u_k] = 0, \quad Q_k[v_1 \ v_2 \ \dots \ v_k] = 0$$

## WEDDERBURN SEQUENCE

$$A_k = A_{k-1} - A_{k-1}y_k\omega_{k-1}x_k^T A_{k-1}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

## WEDDERBURN SEQUENCE

$$A_k = A_{k-1} - A_{k-1}y_k\omega_{k-1}x_k^T A_{k-1}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

STATEMENT 5: explicit form

$$A_k = A_{k-1} - Av_k\omega_{k-1}u_k^T A, \quad \omega_k = u_k^T A v_k \neq 0$$

## WEDDERBURN SEQUENCE

$$A_k = A_{k-1} - A_{k-1}y_k\omega_{k-1}x_k^T A_{k-1}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

STATEMENT 5: explicit form

$$A_k = A_{k-1} - Av_k\omega_{k-1}u_k^T A, \quad \omega_k = u_k^T A v_k \neq 0$$

$$\tilde{A}_k = \sum_{p=1}^k Av_p\omega_p^{-1}u_p^T A = AV_k\Omega_k^{-1}U_k^T A$$

$$A = A_k + \tilde{A}_k$$

## WEDDERBURN SEQUENCE

$$A_k = A_{k-1} - A_{k-1}y_k\omega_{k-1}x_k^T A_{k-1}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

STATEMENT 5: explicit form

$$A_k = A_{k-1} - Av_k\omega_{k-1}u_k^TA, \quad \omega_k = u^TAv_k \neq 0$$

$$\tilde{A}_k = \sum_{p=1}^k Av_p\omega_p^{-1}u_p^TA = AV_k\Omega_k^{-1}U_k^TA$$

$$A = A_k + \tilde{A}_k$$

COROLLARY: interpolation

$$X_k^T \tilde{A}_k = X_k^T A, \quad \tilde{A}_k Y_k = A Y_k, \quad U_k^T \tilde{A}_k = U_k^T A, \quad \tilde{A}_k V_k = A V_k.$$

## WEDDERBURN SEQUENCE

$$A_k = A_{k-1} - A_{k-1}y_k\omega_{k-1}x_k^T A_{k-1}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

**STATEMENT 5:** explicit form

$$A_k = A_{k-1} - Av_k\omega_{k-1}u_k^TA, \quad \omega_k = u^TAv_k \neq 0$$

$$\tilde{A}_k = \sum_{p=1}^k Av_p\omega_p^{-1}u_p^TA = AV_k\Omega_k^{-1}U_k^TA$$

$$A = A_k + \tilde{A}_k$$

**COROLLARY:** interpolation

$$X_k^T \tilde{A}_k = X_k^T A, \quad \tilde{A}_k Y_k = A Y_k, \quad U_k^T \tilde{A}_k = U_k^T A, \quad \tilde{A}_k V_k = A V_k.$$

**REMARK:** Gaußian elimination:  $X_k = [e_{i_1}, \dots, e_{i_k}]$  and  $Y_k = [e_{j_1}, \dots, e_{j_k}]$ .

## WEDDERBURN ELIMINATION

$$X_k^T \tilde{A}_k = X_k^T A, \quad \tilde{A}_k Y_k = A Y_k, \quad U_k^T \tilde{A}_k = U_k^T A, \quad \tilde{A}_k V_k = A V_k.$$

## GAUSSIAN ELIMINATION

$$\tilde{A}_k[\mathcal{I},:] = A[\mathcal{I},:], \quad \tilde{A}_k[:,\mathcal{J}] = A[:,\mathcal{J}]$$



Carl Friedrich Gauß (1777–1855)

## WEDDERBURN SEQUENCE: summary

$$A_k = A_{k-1} - \frac{A_{k-1}y_k x_k^T A_{k-1}}{x_k^T A_{k-1} y_k}, \quad \omega_k \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

$$A_k = P_k^T A = A Q_k$$

$$\begin{aligned} P_k^2 &= P_k; & P_k P_m &= P_{\max(k,m)}; & P_k X_k &= 0 \\ Q_k^2 &= Q_k; & Q_k Q_m &= Q_{\max(k,m)}; & Q_k Y_k &= 0 \end{aligned}$$

$$u_k \stackrel{\text{def}}{=} P_{k-1} x_k \quad v_k \stackrel{\text{def}}{=} Q_{k-1} y_k$$

$$A_k = A_{k-1} - A v_k \omega_{k-1} u_k^T A, \quad \omega_k = u^T A v_k \neq 0$$

$$U_k^T A V_k = \Omega_k$$

$$\tilde{A}_k = A V_k \Omega_k^{-1} U_k^T A$$

$$A = \tilde{A}_k + A_k$$

## PIVOTING

$$B = A - Ay\omega^{-1}x^TA, \quad \omega \stackrel{\text{def}}{=} x^TAy \neq 0$$

## PIVOTING

$$B = A - Ay\omega^{-1}x^T A, \quad \omega \stackrel{\text{def}}{=} x^T A y \neq 0$$

$\arg \min_{x,y} \|B\|_F$  = dominant singular vector of  $A$

## PIVOTING

$$B = A - Ay\omega^{-1}x^T A, \quad \omega \stackrel{\text{def}}{=} x^T A y \neq 0$$

$$\arg \min_x \|B\|_F =$$

## PIVOTING

$$B = A - Ay\omega^{-1}x^T A, \quad \omega \stackrel{\text{def}}{=} x^T A y \neq 0$$

$$\arg \min_x \|B\|_F = \arg \min_x \left\| A - \frac{Ayx^T A}{x^T A y} \right\|_F = \frac{Ay}{\|Ay\|}$$

$$\arg \min_y \|B\|_F = \arg \min_y \left\| A - \frac{Ayx^T A}{x^T A y} \right\|_F = \frac{A^T x}{\|A^T x\|}$$

## PIVOTING

$$B = A - Ay\omega^{-1}x^T A, \quad \omega \stackrel{\text{def}}{=} x^T A y \neq 0$$

$$\arg \min_x \|B\|_F = \arg \min_x \left\| A - \frac{Ayx^T A}{x^T A y} \right\|_F = \frac{Ay}{\|Ay\|}$$

$$\arg \min_y \|B\|_F = \arg \min_y \left\| A - \frac{Ayx^T A}{x^T A y} \right\|_F = \frac{A^T x}{\|A^T x\|}$$

choose  $y_k$ , set

$$x_k = \frac{A_{k-1}y_k}{\|A_{k-1}y_k\|}, \quad A_k = (I - x_k x_k^T) A_{k-1}; \quad (\text{WCP})$$

choose  $x_k$ , set

$$y_k = \frac{A_{k-1}^T x_k}{\|A_{k-1}^T x_k\|}, \quad A_k = A_{k-1} (I - y_k y_k^T) \quad (\text{WRP})$$

## PIVOTING

$$A_k = A_{k-1} - A_{k-1}y_k\omega_k^{-1}x_k^T A_{k-1}, \quad \omega \stackrel{\text{def}}{=} x_k^T A_{k-1} y_k \neq 0$$

choose  $y_k$ , set  $x_k = \frac{A_{k-1}y_k}{\|A_{k-1}y_k\|}$ ,  $A_k = (I - x_k x_k^T) A_{k-1}$ ; (WCP)

choose  $x_k$ , set  $y_k = \frac{A_{k-1}^T x_k}{\|A_{k-1}^T x_k\|}$ ,  $A_k = A_{k-1} (I - y_k y_k^T)$  (WRP)

## WCP

1.  $X_k^T X_k = I$ ;
2.  $P_k = P_k^T = I - X_k X_k^T$
3.  $u_k \stackrel{\text{def}}{=} P_{k-1} x_k = x_k$ .

## WRP

1.  $Y_k^T Y_k = I$ ;
2.  $Q_k = Q_k^T = I - Y_k Y_k^T$ ;
3.  $v_k \stackrel{\text{def}}{=} Q_{k-1} y_k = y_k$ .

## ALGORITHM (Wedderburn elimination with column pivoting)

```

1: repeat
2:    $k := k + 1$ , choose unit vector  $y_k$ 
3:    $x := Ay_k$ ;  $x' := (I - X_{k-1}X_{k-1}^T)x$ 
4:   if  $\|x'\| < \text{tol}\|x\|$  then {Breakdown}
      return  $\tilde{A} := X_{k-1}B_{k-1}^T$  {or try another  $y_k$ }
6:   else
7:      $x_k := x'/\|x'\|$ 
8:   end if
9:    $b_k = A^T x_k$ ,  $\text{err} = \|b_k\|$ ,  $\text{nrm}^2 := \text{nrm}^2 + \|b_k\|^2$ 
10:   $X_k = [X_{k-1} \ x_k]$ ,  $B_k = [B_{k-1} \ b_k]$ 
11: until  $\text{err} \leq \varepsilon \text{nrm}$ 
12: return  $\tilde{A} := X_k B_k^T$ 

```

## ALGORITHM (Wedderburn elimination with column pivoting)

```

1: repeat
2:    $k := k + 1$ , choose unit vector  $y_k$  (How?)
3:    $x := Ay_k$ ;  $x' := (I - X_{k-1}X_{k-1}^T)x$ 
4:   if  $\|x'\| < \text{tol}\|x\|$  then {Breakdown}
      return  $\tilde{A} := X_{k-1}B_{k-1}^T$  {or try another  $y_k$ }
6:   else
7:      $x_k := x'/\|x'\|$ 
8:   end if
9:    $b_k = A^T x_k$ ,  $\text{err} = \|b_k\|$ ,  $\text{nrm}^2 := \text{nrm}^2 + \|b_k\|^2$ 
10:   $X_k = [X_{k-1} \ x_k]$ ,  $B_k = [B_{k-1} \ b_k]$ 
11: until  $\text{err} \leq \varepsilon \text{nrm}$ 
12: return  $\tilde{A} := X_k B_k^T$ 

```

## LANCZOS BIDIAGONALIZATION

Initialization:  $y_0 = 0, \beta_0 = 0$ , unit vector  $x_0$

```
1: for k = 1, 2, ... do
2:    $y := A^T x_{k-1}; \quad y' := y - \beta_{k-1} y_{k-1}, \quad \alpha_k := \|y'\|, \quad y_k := y'/\|y'\|$ 
3:    $x := Ay_k; \quad x' := x - \alpha_k x_k, \quad \beta_k := \|x'\|, \quad x_k := x'/\|x'\|$ 
4: end for
```

## LANCZOS BIDIAGONALIZATION

Initialization:  $y_0 = 0, \beta_0 = 0$ , unit vector  $x_0$

```

1: for k = 1, 2, ... do
2:    $y := A^T x_{k-1}; \quad y' := y - \beta_{k-1} y_{k-1}, \quad \alpha_k := \|y'\|, \quad y_k := y'/\|y'\|$ 
3:    $x := A y_k; \quad x' := x - \alpha_k x_k, \quad \beta_k := \|x'\|, \quad x_k := x'/\|x'\|$ 
4: end for

```

## LANCZOS ‘PIVOTING’

$$y_{k+1} = \frac{A_{k-1}^T x_k}{\|A_{k-1}^T x_k\|} = \frac{A^T P_{k-1} x_k}{\|A^T P_{k-1} x_k\|} = \frac{A^T (I - X_{k-1} X_{k-1}^T) x_k}{\|A^T (I - X_{k-1} X_{k-1}^T) x_k\|} = \frac{A^T x_k}{\|A^T x_k\|}$$

## LANCZOS BIDIAGONALIZATION

Initialization:  $y_0 = 0, \beta_0 = 0$ , unit vector  $x_0$

```

1: for k = 1, 2, ... do
2:    $y := A^T x_{k-1}; \quad y' := y - \beta_{k-1} y_{k-1}, \quad \alpha_k := \|y'\|, \quad y_k := y'/\|y'\|$ 
3:    $x := Ay_k; \quad x' := x - \alpha_k x_k, \quad \beta_k := \|x'\|, \quad x_k := x'/\|x'\|$ 
4: end for

```

## LANCZOS ‘PIVOTING’

$$y_{k+1} = \frac{A_{k-1}^T x_k}{\|A_{k-1}^T x_k\|} = \frac{A^T P_{k-1} x_k}{\|A^T P_{k-1} x_k\|} = \frac{A^T(I - X_{k-1} X_{k-1}^T) x_k}{\|A^T(I - X_{k-1} X_{k-1}^T) x_k\|} = \frac{A^T x_k}{\|A^T x_k\|}$$

## KRYLOV SUBSPACES

$$\begin{aligned} \text{span } X_k^{\{\text{wcp}\}} &= \text{span}\{Ay_1, (AA^T)Ay_1, \dots, (AA^T)^{k-1}Ay_1\}, \\ \text{span } X_k^{\{\text{lnc}\}} &= \text{span}\{(AA^T)x_0, (AA^T)^2x_0, \dots, (AA^T)^kx_0\}. \end{aligned}$$

## KRYLOV SUBSPACES

$$\text{span } X_k^{\{\text{wcp}\}} = \text{span } X_k^{\{\text{lnc}\}} = \text{span}\{(A A^T)x_0, (A A^T)^2 x_0, \dots, (A A^T)^k x_0\}$$



Alexei Nikolaevich Krylov /Крылов/ (1863–1945)

**TENSOR**-by-vector-by-vector

$$u = A \times_2 v^T \times_3 w^T, \quad u_i = \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} a_{ijk} v_j w_k$$

$$A \times_2 v^T \times_3 w^T \stackrel{\text{def}}{=} Avw, \quad A \times_3 w^T \times_1 u^T \stackrel{\text{def}}{=} Awu, \quad A \times_1 u^T \times_2 v^T \stackrel{\text{def}}{=} Auv$$

**TENSOR**-by-vector-by-vector

$$u = A \times_2 v^T \times_3 w^T, \quad u_i = \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} a_{ijk} v_j w_k$$

$$A \times_2 v^T \times_3 w^T \stackrel{\text{def}}{=} Avw, \quad A \times_3 w^T \times_1 u^T \stackrel{\text{def}}{=} Awu, \quad A \times_1 u^T \times_2 v^T \stackrel{\text{def}}{=} Auv$$

**TENVEC**-based approximation (Krylov recursion) [Savas, Eldén, 2009]

```

1: for k = 1, 2, ... do
2:    $u := Av_k w_k;$     $u' := (I - U_k U_k^T)u;$     $u_{k+1} := u'/\|u'\|;$     $U_{k+1} := [U_k \ u_{k+1}]$ 
3:    $v := Aw_k u_{k+1};$     $v' := (I - V_k V_k^T)v;$     $v_{k+1} := v'/\|v'\|;$     $V_{k+1} := [V_k \ v_{k+1}]$ 
4:    $w := Au_{k+1} v_{k+1};$     $w' := (I - W_k W_k^T)w;$     $w_{k+1} := w'/\|w'\|;$     $W_{k+1} := [W_k \ w_{k+1}]$ 
5: end for

```

**TENSOR**-by-vector-by-vector

$$u = A \times_2 v^T \times_3 w^T, \quad u_i = \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} a_{ijk} v_j w_k$$

$$A \times_2 v^T \times_3 w^T \stackrel{\text{def}}{=} Avw, \quad A \times_3 w^T \times_1 u^T \stackrel{\text{def}}{=} Awu, \quad A \times_1 u^T \times_2 v^T \stackrel{\text{def}}{=} Auv$$

**TENVEC**-based approximation (Krylov recursion) [Savas, Eldén, 2009]

```

1: for k = 1, 2, ... do
2:    $u := Av_k w_k$ ;  $u' := (I - U_k U_k^T)u$ ;  $u_{k+1} := u'/\|u'\|$ ;  $U_{k+1} := [U_k \ u_{k+1}]$ 
3:    $v := Aw_k u_{k+1}$ ;  $v' := (I - V_k V_k^T)v$ ;  $v_{k+1} := v'/\|v'\|$ ;  $V_{k+1} := [V_k \ v_{k+1}]$ 
4:    $w := Au_{k+1} v_{k+1}$ ;  $w' := (I - W_k W_k^T)w$ ;  $w_{k+1} := w'/\|w'\|$ ;  $W_{k+1} := [W_k \ w_{k+1}]$ 
5: end for

```

CONVERGENCE: no theory in exact-rank case, ‘bad’ examples

## ALGORITHM (Wedderburn for matrices)

```

1: repeat
2:    $k := k + 1$ , choose unit vector  $y_k (= A^T x_{k-1})$ 
3:    $x := Ay_k$ ;  $x' := (I - X_{k-1}X_{k-1}^T)x$ 
4:   if  $\|x'\| < \text{tol} \|x\|$  then {Breakdown}
      return  $\tilde{A} := X_{k-1}B_{k-1}^T$  {or try another  $y_k$ }
6:   else
7:      $x_k := x'/\|x'\|$ ;  $X_k = [X_{k-1} \ x_k]$ 
8:   end if
9:    $b_k = A^T x_k$ ,  $\text{err} = \|b_k\|$ ,  $\text{nrm}^2 := \text{nrm}^2 + \|b_k\|^2$ ,  $B_k = [B_{k-1} \ b_k]$ 
10: until  $\text{err} \leq \varepsilon \text{nrm}$ 

```

**ALGORITHM** (Wedderburn for tensors)

```

1: repeat
2:    $k := k + 1$ , choose unit vector  $y_k \otimes z_k$  ( $= \text{svd}_1(\mathbf{A} \times_1 x_{k-1}^T)$ )
3:    $x := \mathbf{A}y_k z_k$ ;  $x' := (\mathbf{I} - X_{k-1} X_{k-1}^T)x$ 
4:   if  $\|x'\| < \text{tol} \|x\|$  then {Breakdown}
      return  $X := X_{k-1}$  {or try another  $y_k \otimes z_k$ }
6:   else
7:      $x_k := x'/\|x'\|$ ;  $X_k = [X_{k-1} \ x_k]$ 
8:   end if
9:    $\text{err} = \|\mathbf{A} \times_1 x_k^T\|$ ,  $\text{nrm}^2 := \text{nrm}^2 + \text{err}^2$ 
10:  until  $\text{err} \leq \varepsilon \text{nrm}$ 

```

## ALGORITHM (Wedderburn for tensors: Lanczos-like)

```

1: repeat
2:    $k := k + 1$ , choose unit vector  $y_k \otimes z_k = \textcolor{violet}{y} \otimes \textcolor{violet}{z}$ 
3:    $x := \mathbf{A}y_k z_k$ ;  $x' := (\mathbf{I} - X_{k-1} X_{k-1}^T)x$ 
4:   if  $\|x'\| < \text{tol} \|x\|$  then {Breakdown}
      return  $X := X_{k-1}$  {or try another  $y_k \otimes z_k$ }
6:   else
7:      $x_k := x'/\|x'\|$ ;  $X_k = [X_{k-1} \ x_k]$ 
8:   end if
9:   for  $j = 1, \dots, p_{\text{pow}}$  do {Power iterations to approximate  $\mathbf{A} \times_1 x_k^T \approx: \sigma \textcolor{violet}{y} \textcolor{violet}{z}^T$ }
10:     $y := (\mathbf{A} \times_1 x_k^T) z = \mathbf{A} \times_1 x_k^T \times_3 z^T = \mathbf{A} z x_k$ ,  $\sigma := \|y\|$ ,  $\textcolor{violet}{y} := y/\|y\|$ 
11:     $z := (\mathbf{A} \times_1 x_k^T)^T y = \mathbf{A} \times_1 x_k^T \times_2 y^T = \mathbf{A} x_k y$ ,  $\sigma := \|z\|$ ,  $z := z/\|z\|$ 
12:  end for
13:  err :=  $\sigma$ , nrm2 := nrm2 + err2
14: until err  $\leq \varepsilon$  nrm

```

**ALGORITHM** (Wedderburn for tensors: SVD-like)

```

1: repeat
2:    $k := k + 1$ , choose unit vector  $y_k \otimes z_k = \arg \max_{\|y\|=\|z\|=1} \|x'\|$ 
3:    $x := \mathbf{A}y_k z_k$ ;  $x' := (I - X_{k-1}X_{k-1}^T)x = (I - X_{k-1}X_{k-1}^T)(\mathbf{A}y_k z_k)$ 
4:   if  $\|x'\| < \text{tol}\|x\|$  then {Breakdown impossible!}
      return  $X := X_{k-1}$ 
6: else
7:    $x_k := x'/\|x'\|$ ;  $X_k = [X_{k-1} \ x_k]$ 
8: end if
9: for  $j = 1, \dots, p_{\text{pow}}$  do {Power iterations to approximate  $\mathbf{A} \times_1 x_k^T \approx: \sigma y_k z_k^T$ }
10:   $y := (\mathbf{A} \times_1 x_k^T) z = \mathbf{A} \times_1 x_k^T \times_3 z^T = \mathbf{A} z x_k$ ,  $\sigma := \|y\|$ ,  $y := y/\|y\|$ 
11:   $z := (\mathbf{A} \times_1 x_k^T)^T y = \mathbf{A} \times_1 x_k^T \times_2 y^T = \mathbf{A} x_k y$ ,  $\sigma := \|z\|$ ,  $z := z/\|z\|$ 
12: end for
13: err :=  $\sigma$ , nrm2 := nrm2 + err2
14: until err  $\leq \varepsilon$  nrm

```

**ALGORITHM** (Wedderburn for tensors: SVD-like)

```

1: repeat
2:   for  $j = 1, \dots, p_{\text{als}}$  do {ALS iterations to maximize  $\|(I - X_{k-1}X_{k-1}^T)(Ayz)\|$ }
3:      $x := Ayz, \quad x = x/\|x\|, \quad x' = (I - X_{k-1}X_{k-1}^T)x$ 
4:      $y := Azx', \quad \sigma = \|y\|, \quad y = y/\|y\|$ 
5:      $z := Ax'y, \quad \sigma = \|z\|, \quad z = z/\|z\|$ 
6:   end for
7:    $x := Ayz_k; \quad x' := (I - X_{k-1}X_{k-1}^T)x$ 
8:   if  $\|x'\| < \text{tol}\|x\|$  then {Breakdown impossible!}
9:     return  $X := X_{k-1}$ 
10:    else
11:       $x_k := x'/\|x'\|; \quad X_k = [X_{k-1} \ x_k]$ 
12:    end if
13:    err :=  $\sigma, \quad \text{nrm}^2 := \text{nrm}^2 + \text{err}^2$ 
14:  until  $\text{err} \leq \varepsilon \text{nrm}$ 

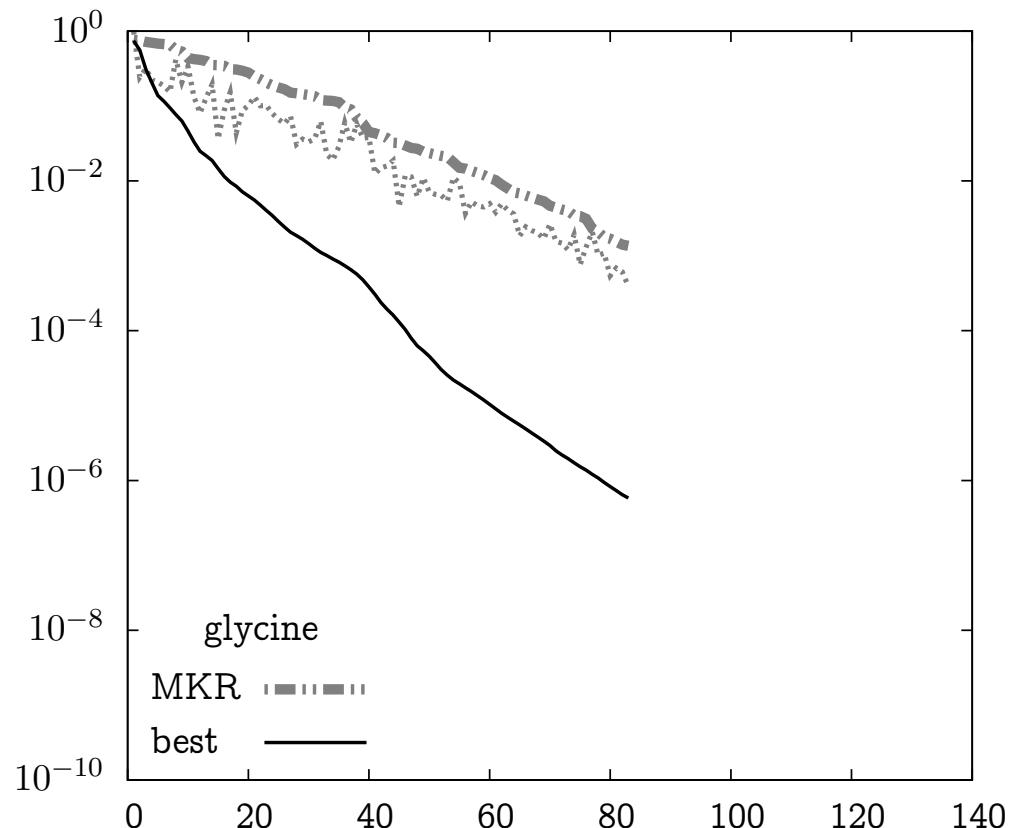
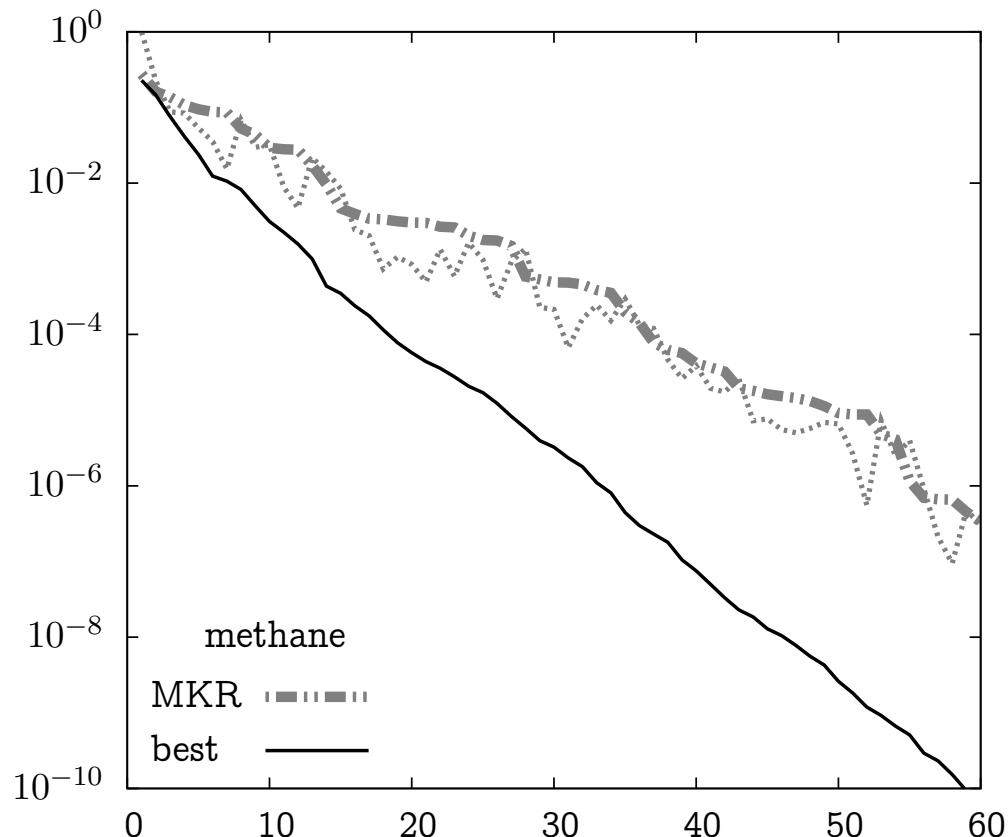
```

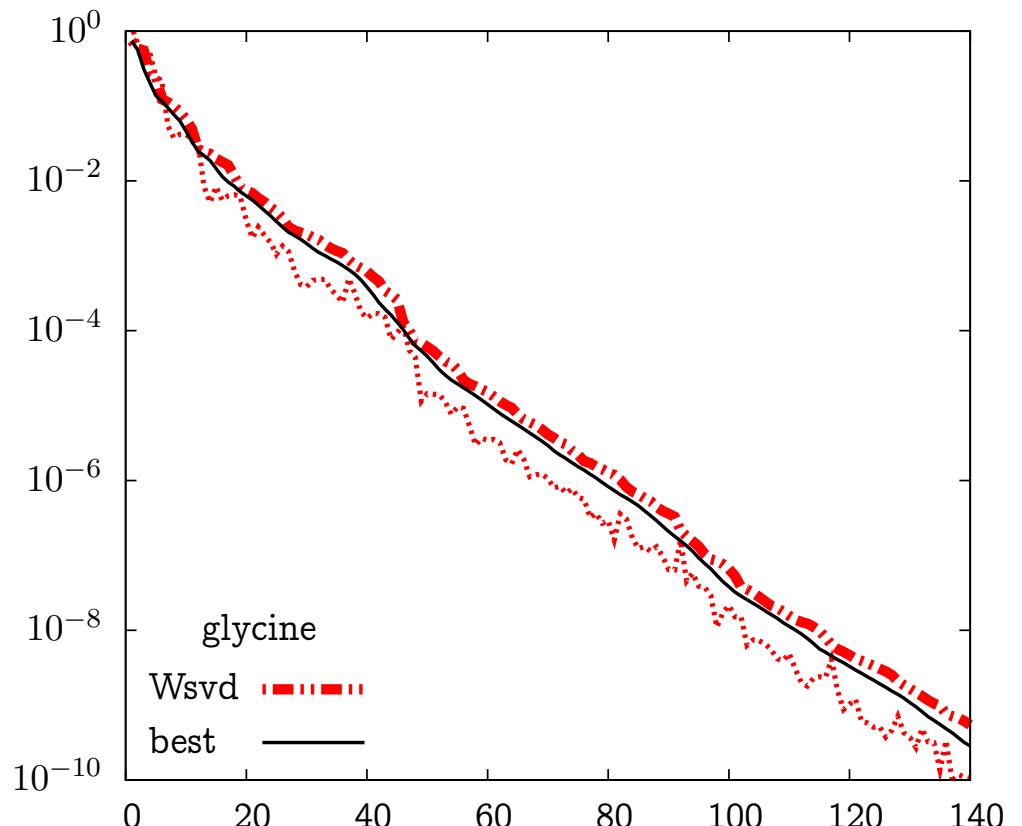
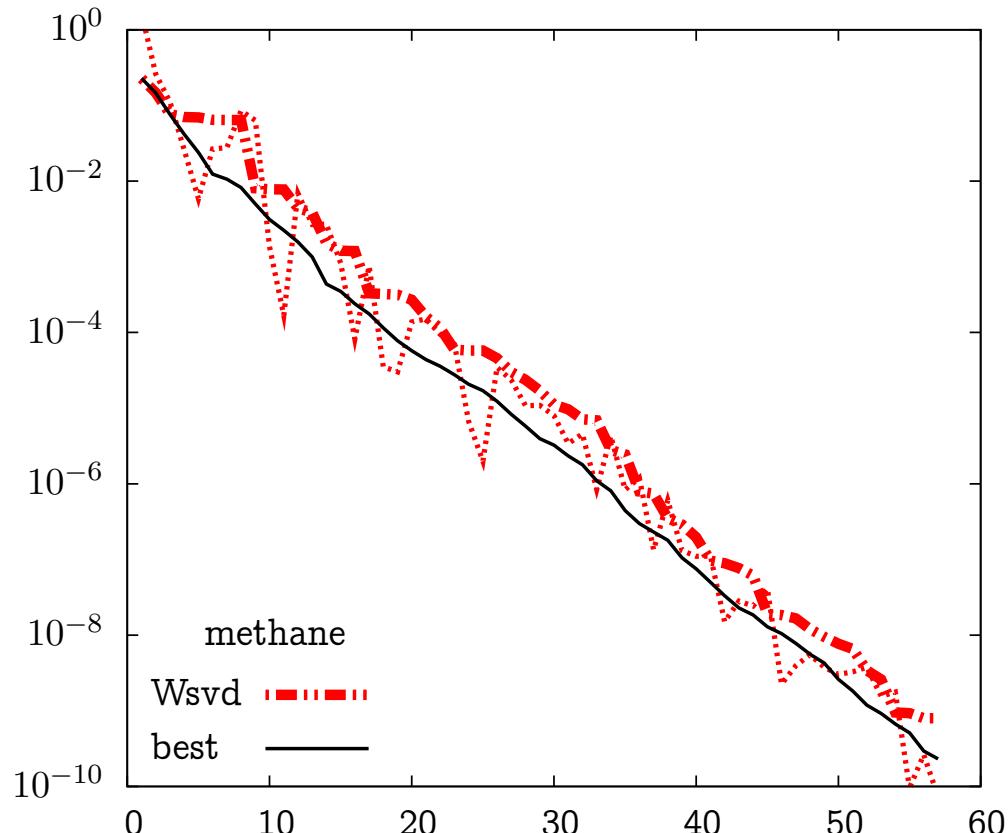
## COMPLEXITY of algorithms for rank- $(r_1, r_2, r_3)$ tensor approximation

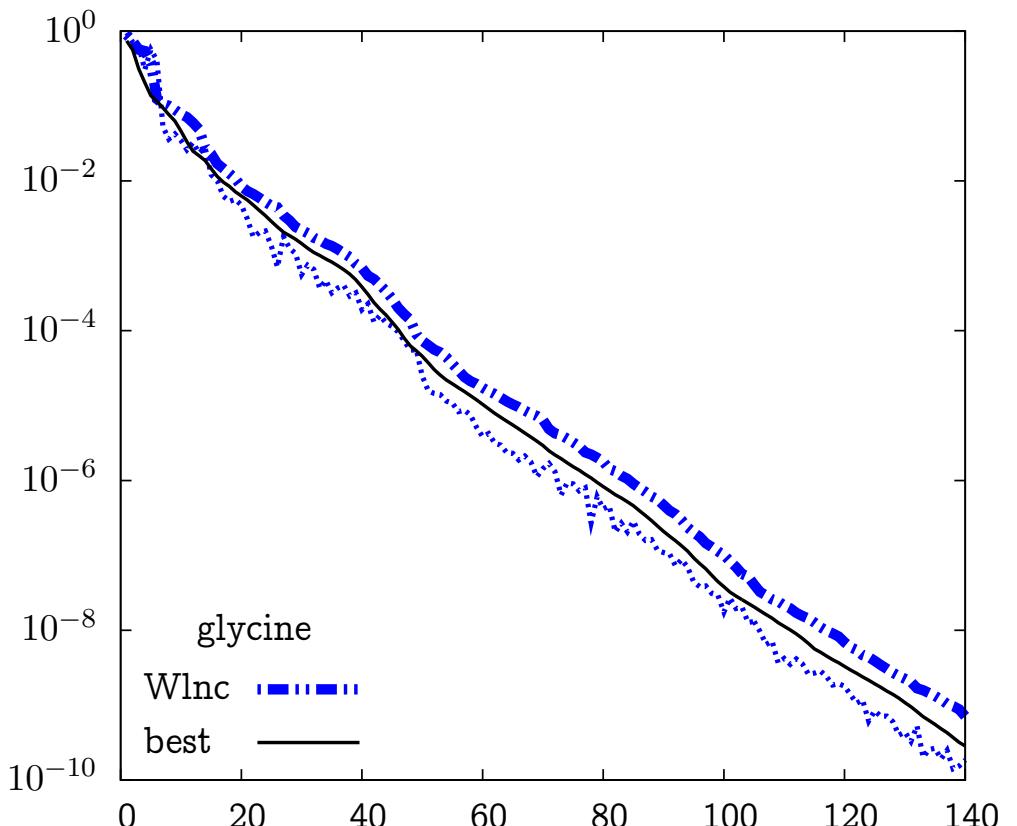
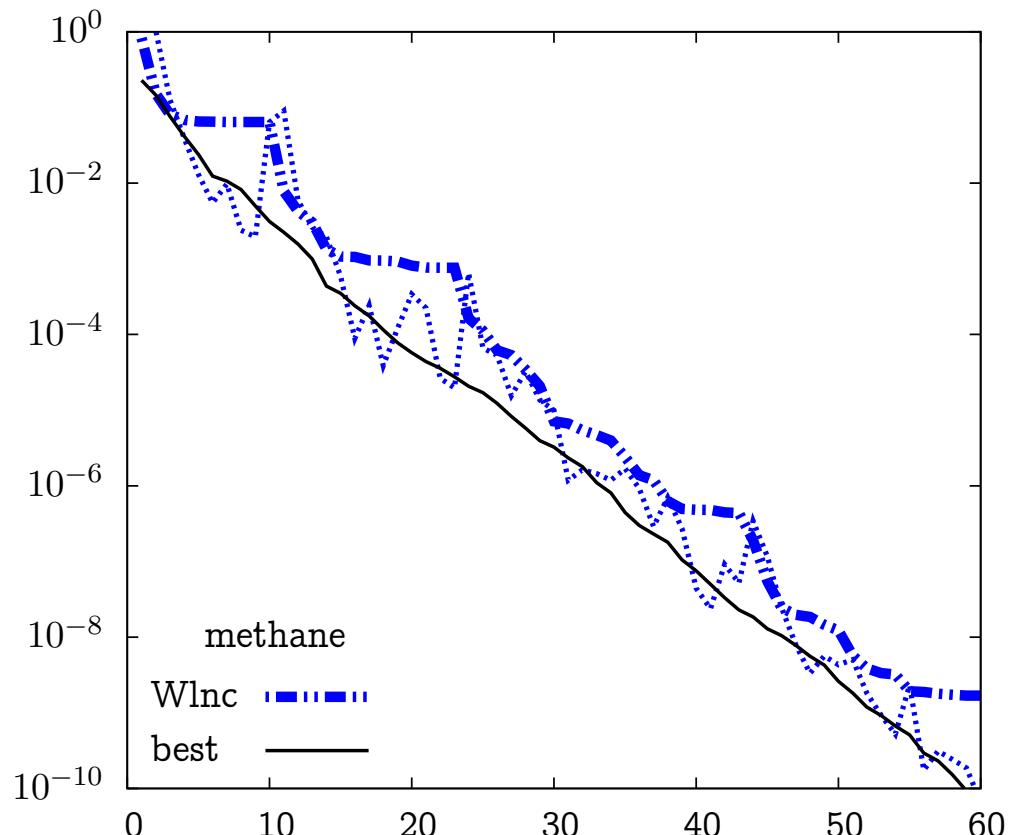
name	output	tenvecs
MKR	$U, V, W$	$3r$
Wsvd	$U, V, W$	$9p_{als}r + 3r$
Wlcz	$U, V, W$	$6p_{pow}r + 3r$
WsvdR	$U, V, W$	$9p_{als}r + 3r$
WlczR	$U, V, W$	$6p_{pow}r + 3r$
WlczR	$U, V, W, G$	$r^2 + 3r$

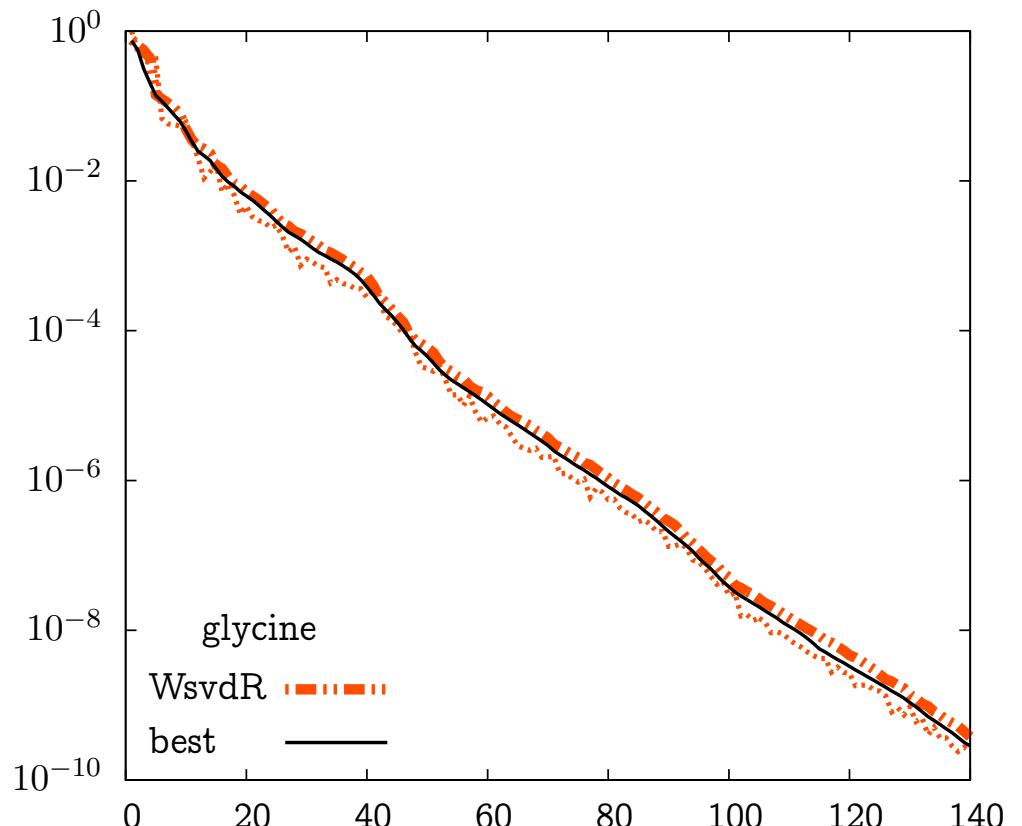
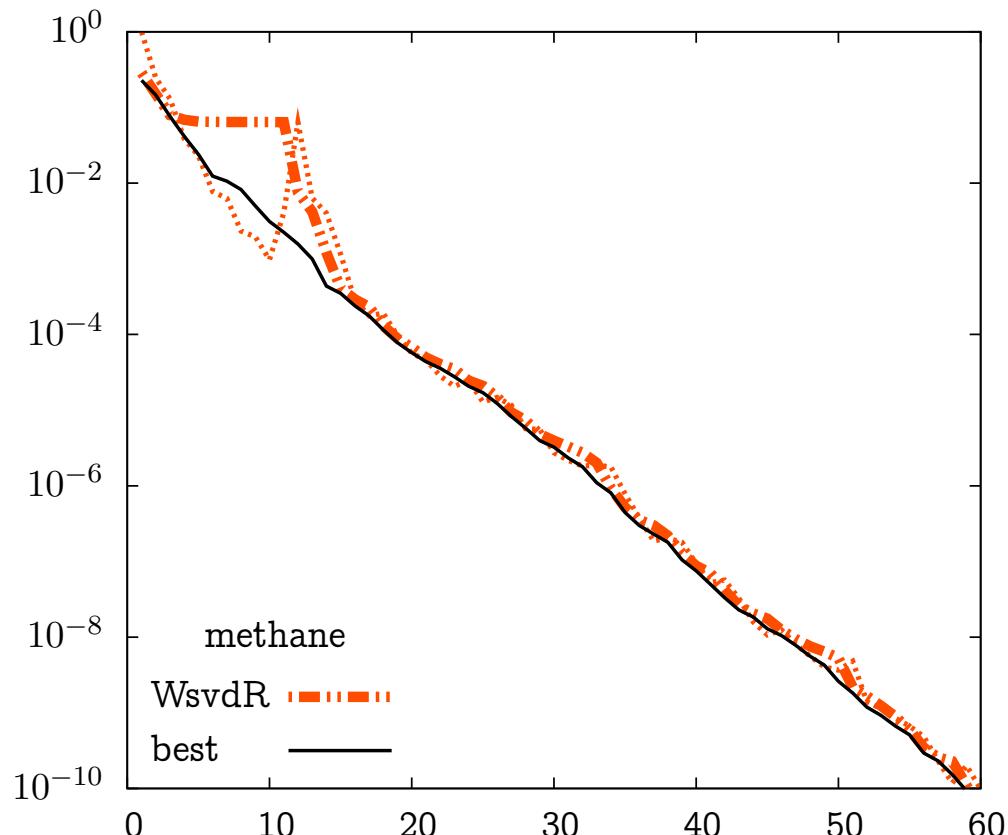
## CANONICAL → TUCKER

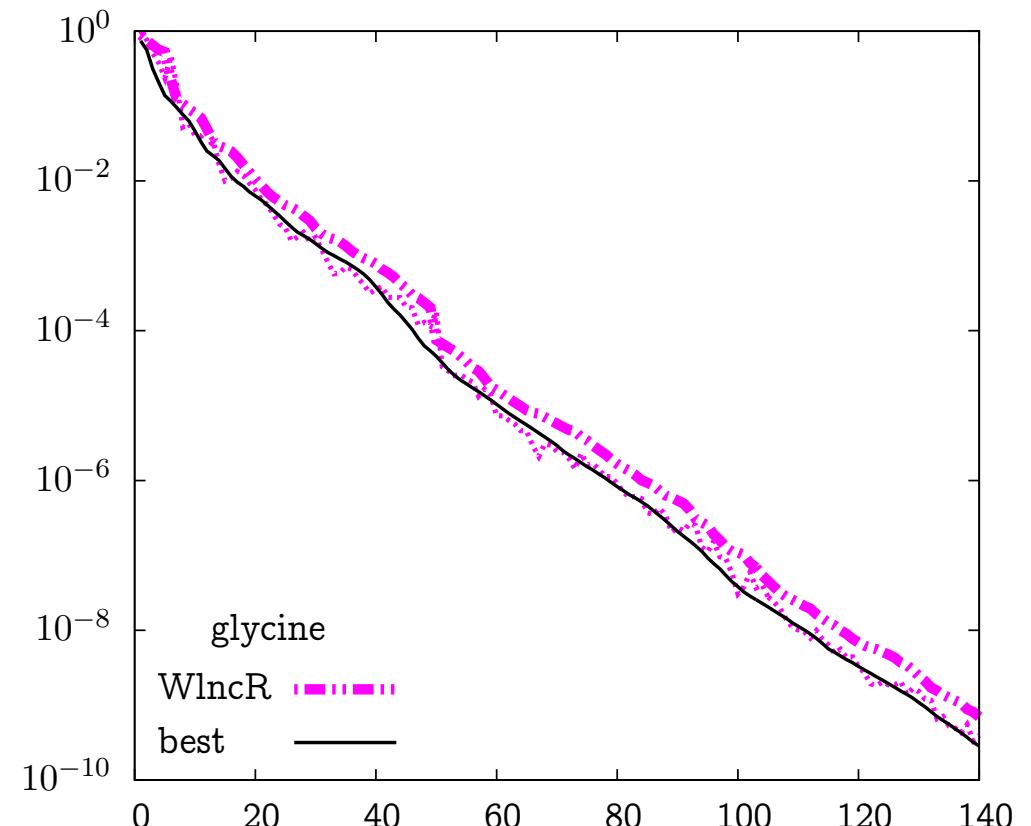
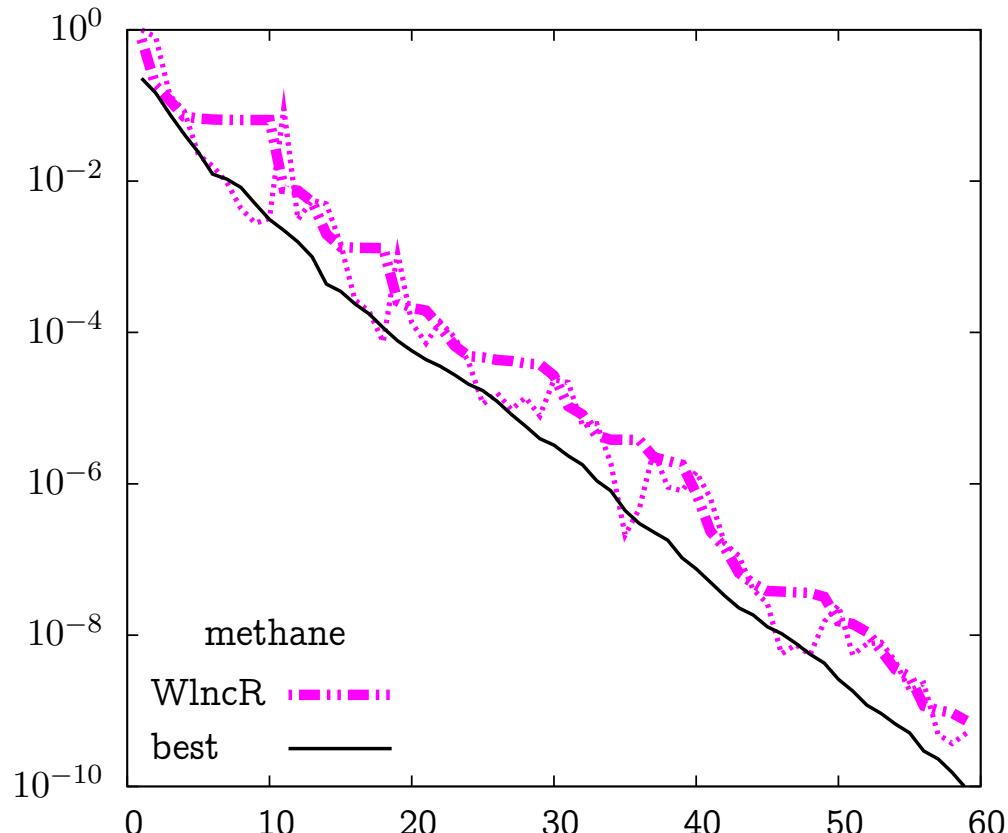
molecule	accuracy	MKR	Wsvd	Wlnc	WsvdR	WlncR	cross	TALS(1)
methane R = 1334	$10^{-4}$	1.0	6.1	4.4	2.0	0.5	1.0	2.6
	$10^{-6}$	1.7	10.7	8.2	3.7	0.8	1.4	9.6
	$10^{-8}$	—	14.5	11.1	5.1	1.4	1.9	30
	$10^{-10}$	—	20.4	16.1	6.9	2.0	2.9	59
ethane R = 3744	$10^{-4}$	2.7	17	15	6.7	1.6	3.0	4.6
	$10^{-6}$	5.2	32	25	12	2.8	4.9	17
	$10^{-8}$	—	45	35	17	3.9	6.3	42
	$10^{-10}$	—	61	50	23	5.3	8.2	83
ethanol R = 6945	$10^{-4}$	8.0	54	45	17	4.7	8.1	22
	$10^{-6}$	14	89	74	30	8.4	13	81
	$10^{-8}$	—	135	108	45	13	17	194
	$10^{-10}$	—	180	145	61	18	22	391
glycine R = 9208	$10^{-4}$	—	85	69	37	7.5	24	32
	$10^{-6}$	—	131	112	57	13	33	96
	$10^{-8}$	—	200	160	80	18	43	211
	$10^{-10}$	—	268	211	114	24	60	412

CANONICAL  $\rightarrow$  TUCKER (MKR)

CANONICAL  $\rightarrow$  TUCKER (Wsvd)

CANONICAL  $\rightarrow$  TUCKER (Wlnc)

CANONICAL  $\rightarrow$  TUCKER (WsvdR)

CANONICAL  $\rightarrow$  TUCKER (WlncR)

TUCKER-BY-TUCKER → TUCKER

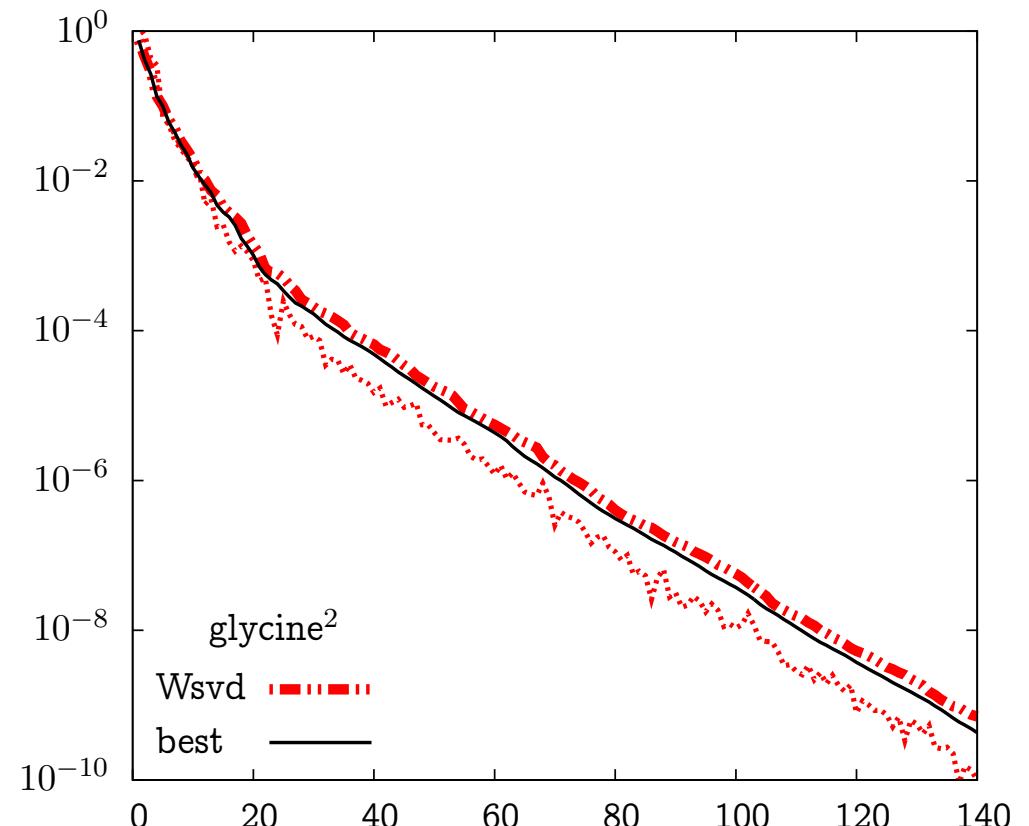
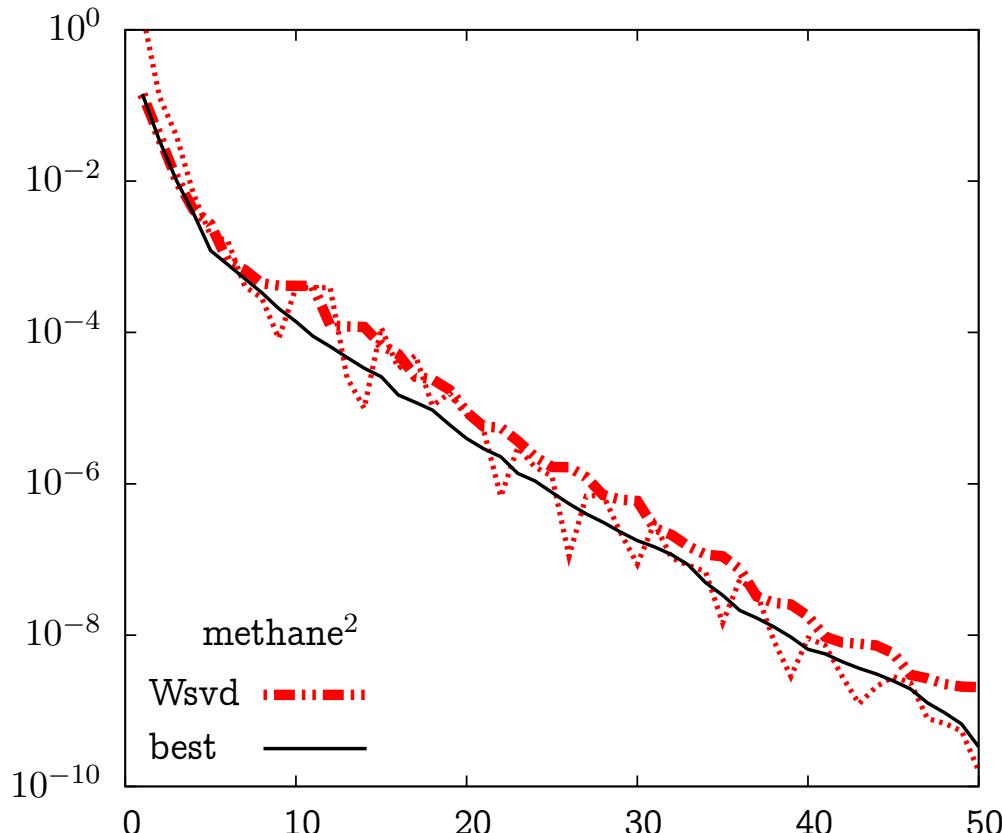
$$\begin{aligned} \mathbf{A}[i, j, k] &= \mathbf{G}[p, q, s] \times_1 \mathbf{U}^{(A)}[i, p] \times_2 \mathbf{V}^{(A)}[j, q] \times_3 \mathbf{W}^{(A)}[k, s], \\ \mathbf{B}[i, j, k] &= \mathbf{H}[a, b, c] \times_1 \mathbf{U}^{(B)}[i, a] \times_2 \mathbf{V}^{(B)}[j, b] \times_3 \mathbf{W}^{(B)}[k, c] \end{aligned}$$

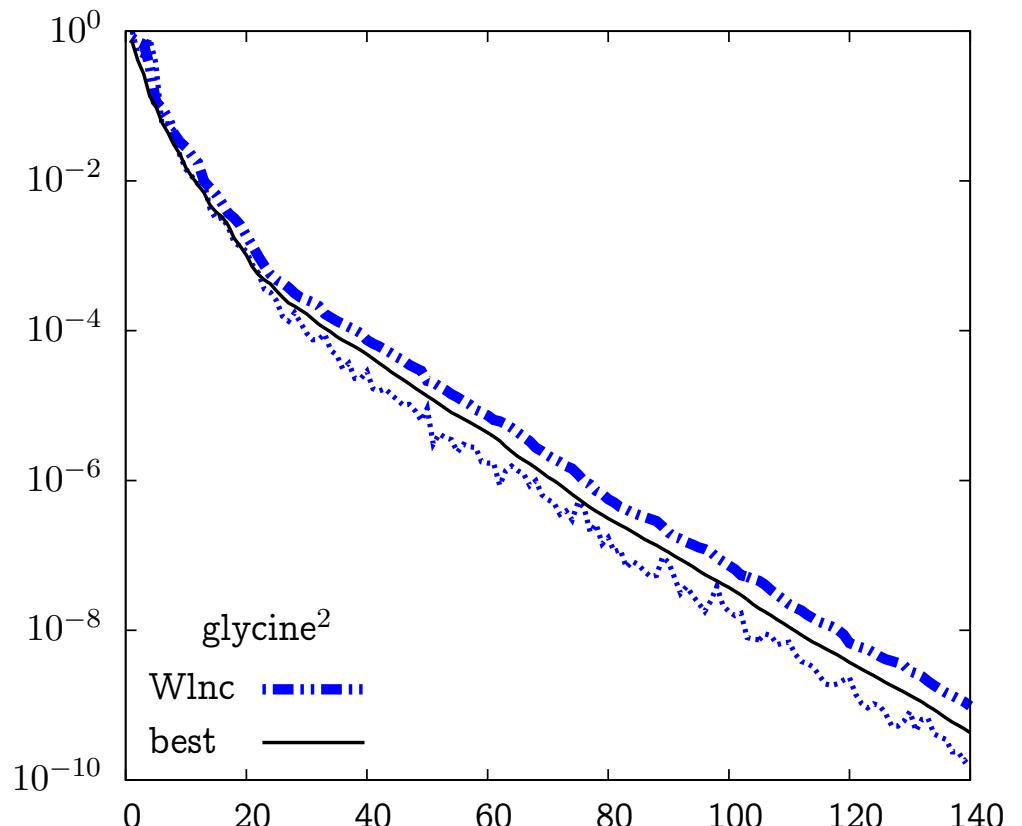
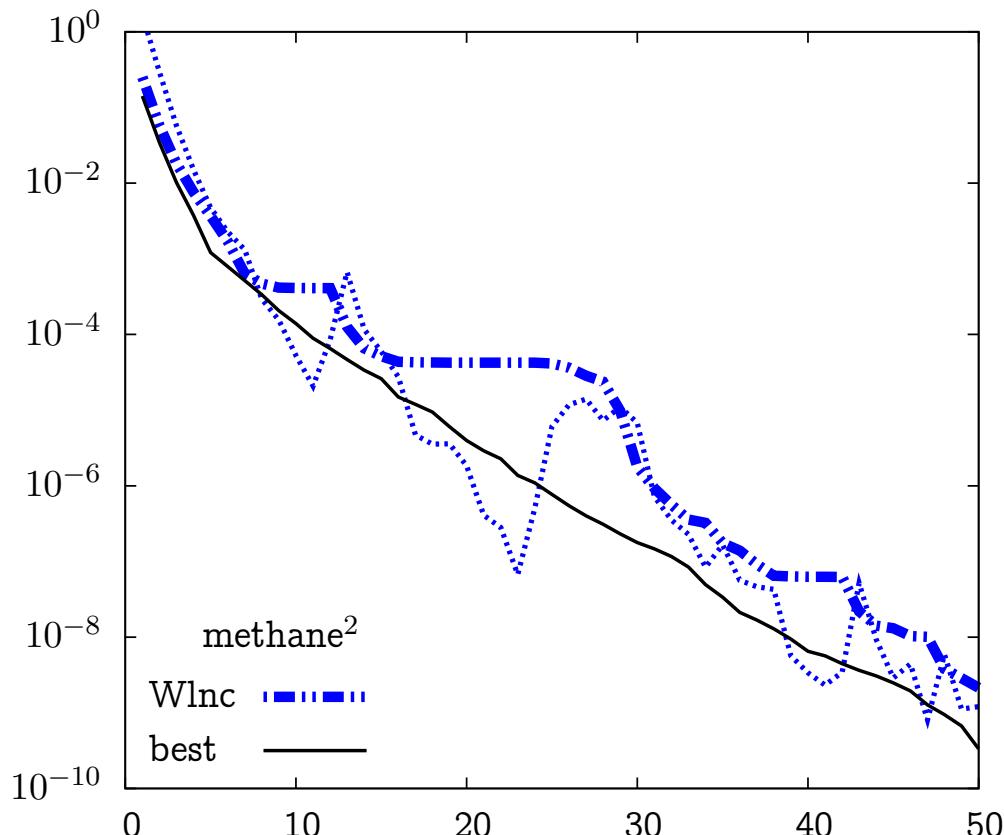
$$\mathbf{F}[i, j, k] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 \mathbf{U}[i, ap] \times_2 \mathbf{V}[j, bq] \times_3 \mathbf{W}[k, cs],$$

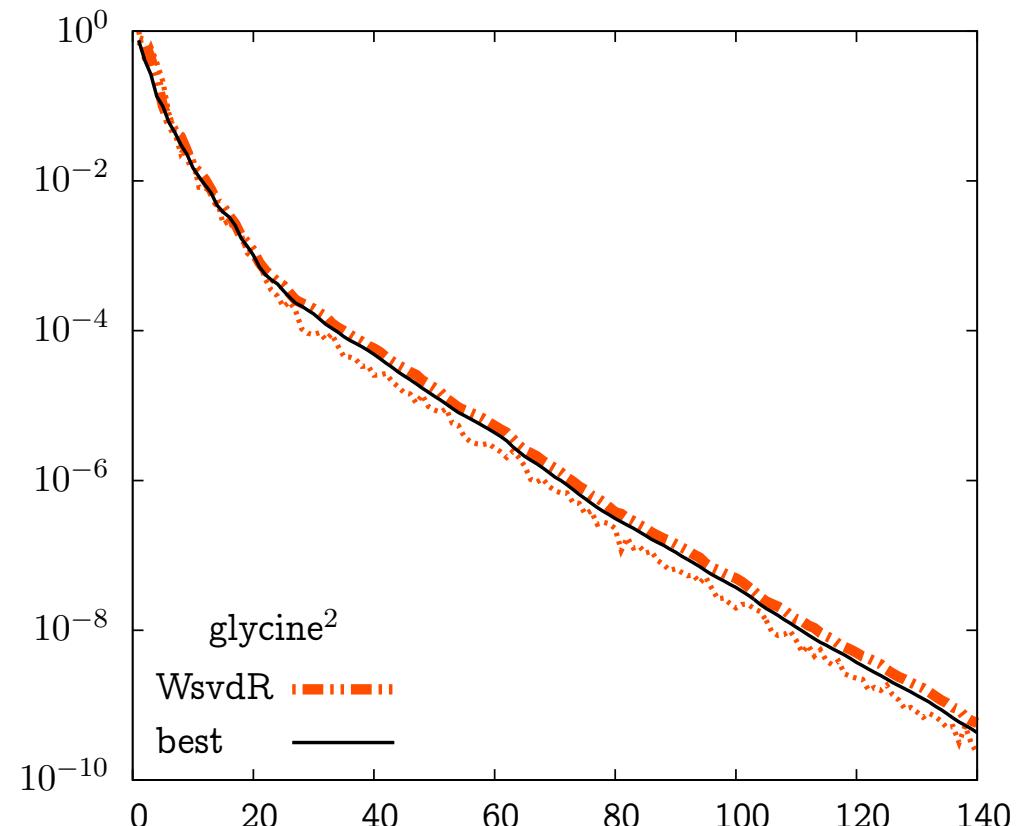
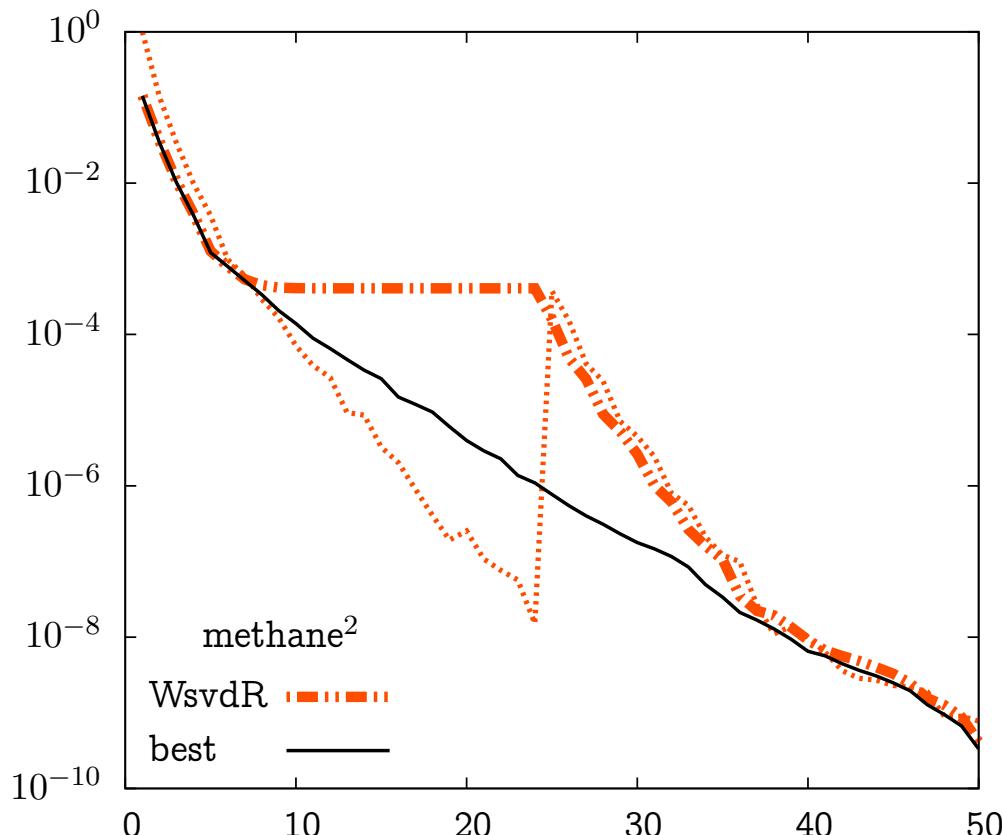
$$f(i, j, k) = \sum_{pqs} \sum_{abc} g(p, q, s) h(a, b, c) u(i, ap) v(j, bq) w(k, cs)$$

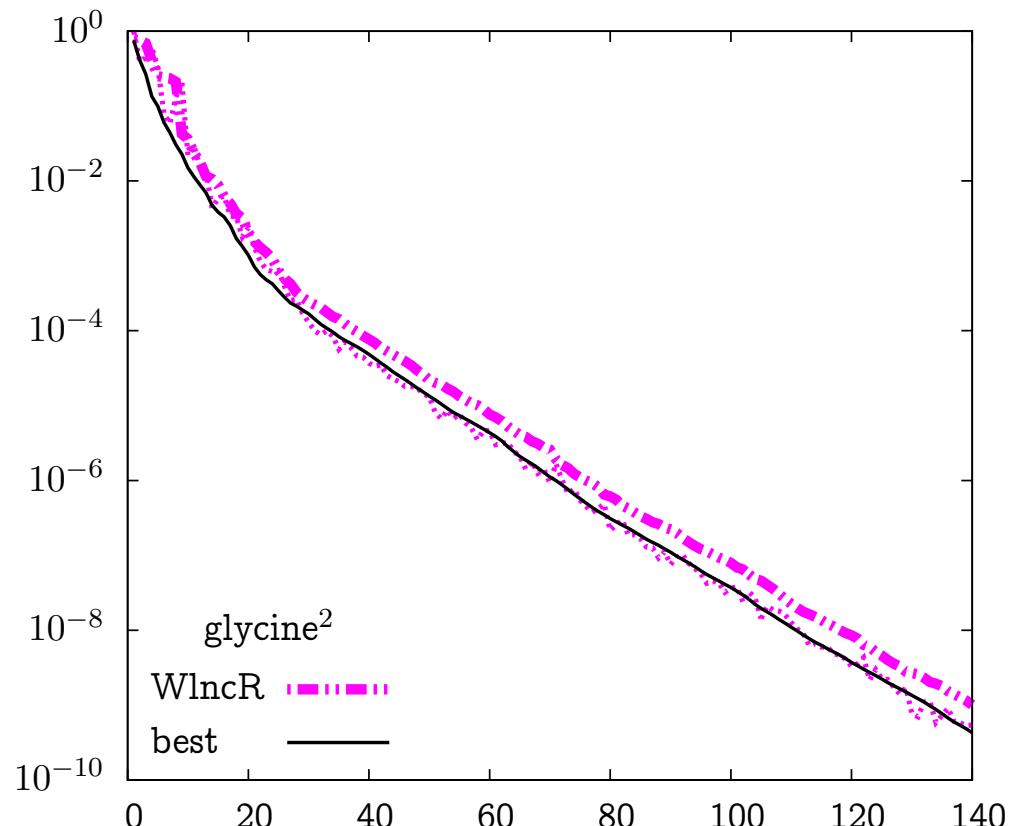
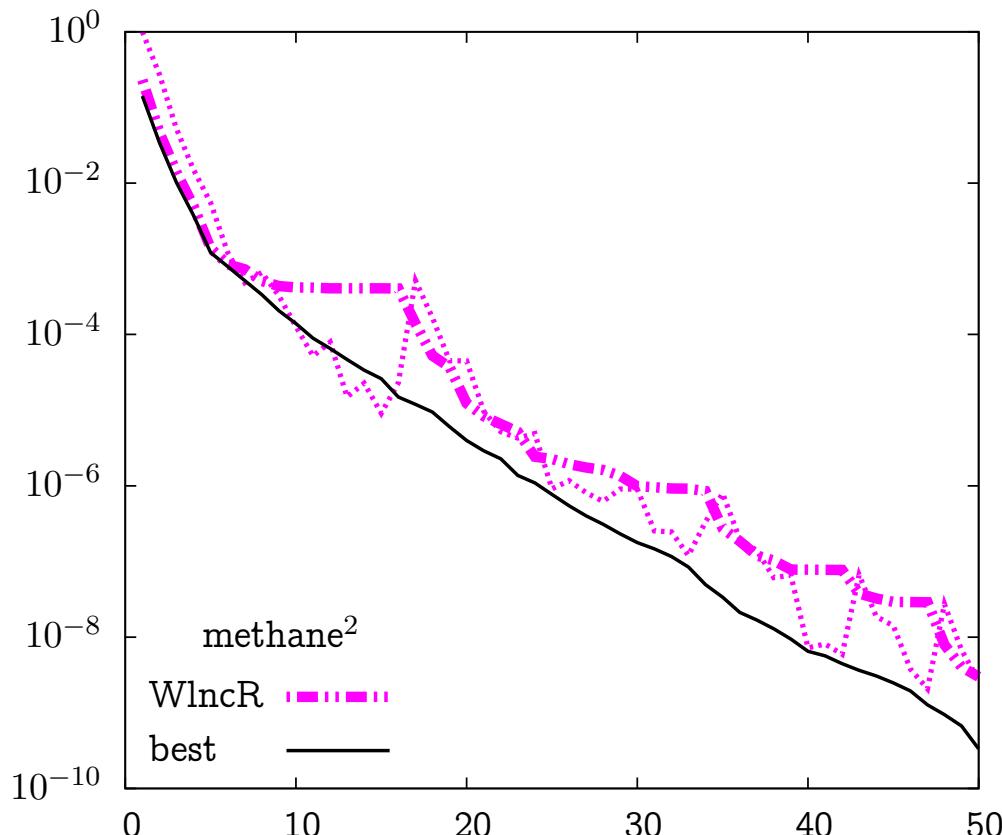
## TUCKER-BY-TUCKER → TUCKER

molecule	accuracy	Wsvd	Wlnc	WsvdR	WlncR	TALS(1)
methane (74, 74, 74)	$10^{-4}$	15.5	13.8	8.1	2.8	2.4
	$10^{-6}$	34	33	14.8	9.7	14.3
	$10^{-8}$	63	46	43	18.6	42
	$10^{-10}$	93	74	68	40	97
ethane (67, 94, 83)	$10^{-4}$	22	20	11.8	4.2	4.2
	$10^{-6}$	46	41	33	14.0	16.7
	$10^{-8}$	82	68	72	27	45
	$10^{-10}$	125	105	127	56	117
ethanol (128, 127, 134)	$10^{-4}$	120	101	106	45	52
	$10^{-6}$	281	228	293	176	257
	$10^{-8}$	493	419	635	441	678
	$10^{-10}$	736	653	1100	808	1370
glycine (62, 176, 186)	$10^{-4}$	179	170	177	60	64
	$10^{-6}$	442	380	600	217	270
	$10^{-8}$	732	600	1033	500	646
	$10^{-10}$	1010	850	1530	888	1223

TUCKER-BY-TUCKER  $\rightarrow$  TUCKER (Wsvd)

TUCKER-BY-TUCKER  $\rightarrow$  TUCKER (Wlnc)

TUCKER-BY-TUCKER  $\rightarrow$  TUCKER (WsvdR)

TUCKER-BY-TUCKER  $\rightarrow$  TUCKER (WlncR)

AND A LITTLE MORE

TUCKER-BY-TUCKER → TUCKER

$$\mathbf{F}[i, j, k] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 U[i, ap] \times_2 V[j, bq] \times_3 W[k, cs],$$

$$f(i, j, k) = \sum_{pqs} \sum_{abc} g(p, q, s) h(a, b, c) u(i, ap) v(j, bq) w(k, cs)$$

TUCKER-BY-TUCKER → TUCKER

$$\mathbf{F}[i, j, k] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 U[i, ap] \times_2 V[j, bq] \times_3 W[k, cs],$$

$$f(i, j, k) = \sum_{pqs} \sum_{abc} g(p, q, s) h(a, b, c) u(i, ap) v(j, bq) w(k, cs)$$

$$\mathbf{F}[i, j, k] = \mathbf{U}'[i, bq, cs] \times_2 V[j, bq] \times_3 W[k, cs],$$

$$\mathbf{U}'[i, bq, cs] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 U(i, ap).$$

TUCKER-BY-TUCKER → TUCKER

$$\mathbf{F}[i, j, k] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 U[i, ap] \times_2 V[j, bq] \times_3 W[k, cs],$$

$$f(i, j, k) = \sum_{pqs} \sum_{abc} g(p, q, s) h(a, b, c) u(i, ap) v(j, bq) w(k, cs)$$

$$\mathbf{F}[i, j, k] = \mathbf{U}'[i, bq, cs] \times_2 V[j, bq] \times_3 W[k, cs],$$

$$\mathbf{U}'[i, bq, cs] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 U(i, ap).$$

Compute mode-1 dominant subspace of  $\mathbf{U}'[i, bq, cs]$  instead of  $\mathbf{F}[i, j, k]$

TUCKER-BY-TUCKER → TUCKER

$$\mathbf{F}[i, j, k] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 \mathbf{U}[i, ap] \times_2 \mathbf{V}[j, bq] \times_3 \mathbf{W}[k, cs],$$

$$f(i, j, k) = \sum_{pqs} \sum_{abc} g(p, q, s) h(a, b, c) u(i, ap) v(j, bq) w(k, cs)$$

$$\mathbf{F}[i, j, k] = \mathbf{U}'[i, bq, cs] \times_2 \mathbf{V}[j, bq] \times_3 \mathbf{W}[k, cs],$$

$$\mathbf{U}'[i, bq, cs] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 \mathbf{U}(i, ap).$$

## ACCURACY

$$\frac{\|\Delta \mathbf{F}\|_F}{\|\mathbf{F}\|_F} \leq c_F \frac{\|\Delta \mathbf{U}'\|_F}{\|\mathbf{U}'\|_F},$$

$$\frac{\|\Delta \mathbf{F}\|_2}{\|\mathbf{F}\|_2} \leq c_2 \frac{\|\Delta \mathbf{U}'\|_2}{\|\mathbf{U}'\|_2},$$

$$\frac{\|\Delta \mathbf{F}\|_2}{\|\mathbf{F}\|_2} \leq c_2 \frac{\|\Delta \mathbf{U}'\|_2}{\|\mathbf{U}'\|_2},$$

$$c_F = \frac{\|\mathbf{U}'\|_F \|\mathbf{V}\|_2 \|\mathbf{W}\|_2}{\|\mathbf{F}\|_F};$$

$$c_2 = \frac{\|\mathbf{U}'\|_2 \|\mathbf{V}\|_2 \|\mathbf{W}\|_2}{\|\mathbf{F}\|_2};$$

$$c_2 = \frac{\|\mathbf{U}'\|_2 \|\mathbf{V}\|_2 \|\mathbf{W}\|_2}{\|\mathbf{F}\|_2}.$$

TUCKER-BY-TUCKER → TUCKER

$$\mathbf{F}[i, j, k] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 \mathbf{U}[i, ap] \times_2 \mathbf{V}[j, bq] \times_3 \mathbf{W}[k, cs],$$

$$f(i, j, k) = \sum_{pqs} \sum_{abc} g(p, q, s) h(a, b, c) u(i, ap) v(j, bq) w(k, cs)$$

$$\mathbf{F}[i, j, k] = \mathbf{U}'[i, bq, cs] \times_2 \mathbf{V}[j, bq] \times_3 \mathbf{W}[k, cs],$$

$$\mathbf{U}'[i, bq, cs] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 \mathbf{U}(i, ap).$$

## ACCURACY

$$\begin{aligned} \frac{\|\Delta \mathbf{F}\|_F}{\|\mathbf{F}\|_F} &\leq c_F \frac{\|\Delta \mathbf{U}'\|_F}{\|\mathbf{U}'\|_F}, & c_F &= \frac{\|\mathbf{U}'\|_F \|\mathbf{V}\|_2 \|\mathbf{W}\|_2}{\|\mathbf{F}\|_F}; \\ \frac{\|\Delta \mathbf{F}\|_2}{\|\mathbf{F}\|_2} &\leq c_2 \frac{\|\Delta \mathbf{U}'\|_2}{\|\mathbf{U}'\|_2}, & c_2 &= \frac{\|\mathbf{U}'\|_2 \|\mathbf{V}\|_2 \|\mathbf{W}\|_2}{\|\mathbf{F}\|_2}; \\ \frac{\|\Delta \mathbf{F}\|_2}{\|\mathbf{F}\|_2} &\leq c_2 \frac{\|\Delta \mathbf{U}'\|_2}{\|\mathbf{U}'\|_2}, & c_2 &= \frac{\|\mathbf{U}'\|_2 \|\mathbf{V}\|_2 \|\mathbf{W}\|_2}{\|\mathbf{F}\|_2}. \end{aligned}$$

$$\|\mathbf{A}\|_2 \stackrel{\text{def}}{=} \max_{\|\mathbf{u}\|=\|\mathbf{v}\|=\|\mathbf{w}\|=1} \mathbf{A} \times_1 \mathbf{u}^\top \times_2 \mathbf{v}^\top \times_3 \mathbf{w}^\top = \max_{\|\mathbf{u}\|=\|\mathbf{v}\|=\|\mathbf{w}\|=1} \langle \mathbf{A}, \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w} \rangle,$$

TUCKER-BY-TUCKER → TUCKER

$$\mathbf{F}[i, j, k] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 \mathbf{U}[i, ap] \times_2 \mathbf{V}[j, bq] \times_3 \mathbf{W}[k, cs],$$

$$f(i, j, k) = \sum_{pqs} \sum_{abc} g(p, q, s) h(a, b, c) u(i, ap) v(j, bq) w(k, cs)$$

$$\mathbf{F}[i, j, k] = \mathbf{U}'[i, bq, cs] \times_2 \mathbf{V}[j, bq] \times_3 \mathbf{W}[k, cs],$$

$$\mathbf{U}'[i, bq, cs] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 \mathbf{U}(i, ap).$$

## ACCURACY

$$\frac{\|\Delta \mathbf{F}\|_F}{\|\mathbf{F}\|_F} \leq c_F \frac{\|\Delta \mathbf{U}'\|_F}{\|\mathbf{U}'\|_F},$$

$$\frac{\|\Delta \mathbf{F}\|_2}{\|\mathbf{F}\|_2} \leq c_2 \frac{\|\Delta \mathbf{U}'\|_2}{\|\mathbf{U}'\|_2},$$

$$\frac{\|\Delta \mathbf{F}\|_2}{\|\mathbf{F}\|_2} \leq c_2 \frac{\|\Delta \mathbf{U}'\|_2}{\|\mathbf{U}'\|_2},$$

$$c_F = \frac{\|\mathbf{U}'\|_F \|\mathbf{V}\|_2 \|\mathbf{W}\|_2}{\|\mathbf{F}\|_F};$$

$$c_2 = \frac{\|\mathbf{U}'\|_2 \|\mathbf{V}\|_2 \|\mathbf{W}\|_2}{\|\mathbf{F}\|_2};$$

$$c_2 = \frac{\|\mathbf{U}'\|_2 \|\mathbf{V}\|_2 \|\mathbf{W}\|_2}{\|\mathbf{F}\|_2}.$$

**REMARK:** For any tensor,  $\|\mathbf{A}[i, j, k]\|_2 \leq \|\mathbf{A}[i, jk]\|_2 \leq \|\mathbf{A}[i, j, k]\|_F$ .

TUCKER-BY-TUCKER → TUCKER

$$\mathbf{F}[i, j, k] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 U[i, ap] \times_2 V[j, bq] \times_3 W[k, cs],$$

$$f(i, j, k) = \sum_{pqs} \sum_{abc} g(p, q, s) h(a, b, c) u(i, ap) v(j, bq) w(k, cs)$$

$$\mathbf{F}[i, j, k] = \mathbf{U}'[i, bq, cs] \times_2 V[j, bq] \times_3 W[k, cs],$$

$$\mathbf{U}'[i, bq, cs] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 U(i, ap).$$

GRAM matrix

$$A[i, i'] = (\mathbf{U}'\mathbf{U}'^T)[i, i'] = U[i, ap] \left( \hat{G}[p, p'] \otimes \hat{H}[a, a'] \right) U[a'p', i'],$$

$$\hat{G}[p, p'] = G[p, qs]G[qs, p'], \quad \hat{H}[a, a'] = H[a, bc]H[bc, a'].$$

TUCKER-BY-TUCKER → TUCKER

$$\mathbf{F}[i, j, k] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 U[i, ap] \times_2 V[j, bq] \times_3 W[k, cs],$$

$$f(i, j, k) = \sum_{pqs} \sum_{abc} g(p, q, s) h(a, b, c) u(i, ap) v(j, bq) w(k, cs)$$

$$\mathbf{F}[i, j, k] = \mathbf{U}'[i, bq, cs] \times_2 V[j, bq] \times_3 W[k, cs],$$

$$\mathbf{U}'[i, bq, cs] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 U(i, ap).$$

GRAM matrix

$$\begin{aligned} A[i, i'] &= (\mathbf{U}'\mathbf{U}'^T)[i, i'] = U[i, ap] \left( \hat{G}[p, p'] \otimes \hat{H}[a, a'] \right) U[a'p', i'], \\ \hat{G}[p, p'] &= G[p, qs]G[qs, p'], \quad \hat{H}[a, a'] = H[a, bc]H[bc, a']. \end{aligned}$$

Cost (d-dimensional case)

$\hat{G}$ and $\hat{H}$	$\mathcal{O}(r^{d+1})$
$a(i, i')$	$\mathcal{O}(r^3)$
column $a(i, :)$	$\mathcal{O}(nr^2 + r^3)$

TUCKER-BY-TUCKER  $\rightarrow$  TUCKER

$$\mathbf{F}[i, j, k] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 U[i, ap] \times_2 V[j, bq] \times_3 W[k, cs],$$

$$f(i, j, k) = \sum_{pqs} \sum_{abc} g(p, q, s) h(a, b, c) u(i, ap) v(j, bq) w(k, cs)$$

$$\mathbf{F}[i, j, k] = \mathbf{U}'[i, bq, cs] \times_2 V[j, bq] \times_3 W[k, cs],$$

$$\mathbf{U}'[i, bq, cs] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 U(i, ap).$$

GRAM matrix

$$\begin{aligned} A[i, i'] &= (U'U'^T)[i, i'] = U[i, ap] \left( \hat{G}[p, p'] \otimes \hat{H}[a, a'] \right) U[a'p', i'], \\ \hat{G}[p, p'] &= G[p, qs]G[qs, p'], \quad \hat{H}[a, a'] = H[a, bc]H[bc, a']. \end{aligned}$$

COST (d-dimensional case)

$\hat{G}$ and $\hat{H}$	$\mathcal{O}(r^{d+1})$
$a(i, i')$	$\mathcal{O}(r^3)$
column $a(i, :)$	$\mathcal{O}(nr^2 + r^3)$
incomplete Cholesky	$\mathcal{O}(nr^3 + r^4 + r^{d+1})$

## Cost (d-dimensional case)

$\hat{G}$ and $\hat{H}$	$\mathcal{O}(r^{d+1})$
$a(i, i')$	$\mathcal{O}(r^3)$
column $a(i, :)$	$\mathcal{O}(nr^2 + r^3)$
incomplete Cholesky	$\mathcal{O}(nr^3 + r^4 + r^{d+1})$



André-Louis Cholesky (1875–1918)

## TUCKER-BY-TUCKER → TUCKER

$$\mathbf{F}[i, j, k] = \mathbf{Kron}(\mathbf{G}, \mathbf{H})[ap, bq, cs] \times_1 U[i, ap] \times_2 V[j, bq] \times_3 W[k, cs],$$

$$f(i, j, k) = \sum_{pqs} \sum_{abc} g(p, q, s) h(a, b, c) u(i, ap) v(j, bq) w(k, cs)$$

## TIME for subspaces

molecule	$r_1, r_2, r_3$	T	$\varepsilon$	T(c3d)	T(wsvdr)	T(tals <sub>1</sub> )	$\varepsilon(tals_1)$
methane	(74, 74, 74)	4.0	$3 \cdot 10^{-7}$	78.6	12.4	37	$7 \cdot 10^{-13}$
ethane	(67, 94, 83)	5.3	$6 \cdot 10^{-7}$	76.8	15.1	42	$8 \cdot 10^{-13}$
ethanol	(128, 127, 134)	20	$5 \cdot 10^{-7}$	1050	210	473	$9 \cdot 10^{-13}$
glycine	(62, 176, 186)	38	$8 \cdot 10^{-7}$	1260	237	442	$9 \cdot 10^{-13}$

## INVERSE 3D LAPLACIAN: Newton

$$\mathbf{X}_{k+1} = 2\mathbf{X}_k - \mathbf{X}_k \Delta_3^{-1} \mathbf{X}_k, \quad \mathbf{X}_k \rightarrow \Delta_3^{-1}$$

INVERSE 3D LAPLACIAN: modified Newton [OST, Computing 2009]

$$\begin{aligned} H_k &= \mathcal{P}(2I - Y_k) \\ Y_{k+1} &= \mathcal{P}_1(Y_k H_k), \quad Y_{k+1} \rightarrow I \\ X_{k+1} &= \mathcal{P}_2(X_k H_k), \quad X_{k+1} \rightarrow \Delta_3^{-1} \end{aligned}$$

INVERSE 3D LAPLACIAN: modified Newton

$$\begin{aligned} H_k &= \mathcal{P}(2I - Y_k) \\ Y_{k+1} &= \mathcal{P}_1(Y_k H_k), \quad Y_{k+1} \rightarrow I \\ X_{k+1} &= \mathcal{P}_2(X_k H_k), \quad X_{k+1} \rightarrow \Delta_3^{-1} \end{aligned}$$

OLD: fixed-rank  $\mathcal{P} = \mathcal{P}_{r=(2,2,2)}$ ,  $\mathcal{P}_1 = \mathcal{P}_2 = \mathcal{P}_{r=(12,12,12)}$ , full  $r^2 \times r^2 \times r^2$  core

NEW: fixed-accuracy  $\mathcal{P} = \mathcal{P}_1 = \mathcal{P}_2 = \mathcal{P}_{\varepsilon=10^{-9}}$ , Cholesky + Wedderburn recompression

INVERSE 3D LAPLACIAN: modified Newton

$$\begin{aligned} H_k &= \mathcal{P}(2I - Y_k) \\ Y_{k+1} &= \mathcal{P}_1(Y_k H_k), \quad Y_{k+1} \rightarrow I \\ X_{k+1} &= \mathcal{P}_2(X_k H_k), \quad X_{k+1} \rightarrow \Delta_3^{-1} \end{aligned}$$

OLD: fixed-rank  $\mathcal{P} = \mathcal{P}_{r=(2,2,2)}$ ,  $\mathcal{P}_1 = \mathcal{P}_2 = \mathcal{P}_{r=(12,12,12)}$ , full  $r^2 \times r^2 \times r^2$  core

NEW: fixed-accuracy  $\mathcal{P} = \mathcal{P}_1 = \mathcal{P}_2 = \mathcal{P}_{\varepsilon=10^{-9}}$ , Cholesky + Wedderburn recompression

size	$128^3$	$256^3$	$512^3$
$T(\text{old})$	24	227	—
$\varepsilon(\text{old})$	$10^{-5}$	$10^{-4}$	—
$T(\text{new})$	26	182	820
$\varepsilon(\text{new})$	$10^{-12}$	$10^{-12}$	$10^{-12}$

## References

---

### PREPRINTS

- S. A. Goreinov, I. V. Oseledets, D. V. Savostyanov.  
Wedderburn rank reduction and Krylov subspace method for tensor approximation.  
Part 1: Tucker case  
Preprint INM RAS 2010-01, arXiv:1004:1968, submitted to SISC.
- D. V. Savostyanov, E. E. Tyrtyshnikov, N. L. Zamarashkin  
Fast truncation of mode ranks for bilinear tensor operations  
Preprint INM RAS 2010-02, arXiv:1004:4919, submitted to NLAA.

THANK YOU FOR ATTENTION!