Inria

April 15th, Knoxville, Tennessee

Reducing the memory footprint in Krylov methods through lossy compression

Project talk @ 9th JLESC workshop

Nick Schenkels[1]

Joint work with: Emmanuel Agullo[1] Luc Giraud[1] Sheng Di[2] Franck Cappello[2]

[1]Inria

[2]Argonne National Laboratory

Argonne NATIONAL LABORATORY

Inria inventeurs du monde numérique

Overview

- **Introduction**
  - Krylov methods & inexactness
  - Mixed precision & compression
- **Preliminary results**
  - Some theory
  - Some experiments
- **Outlook**

Introduction

- Krylov methods & inexactness
- Mixed precision & compression

Krylov methods

We want to solve

$$Ax = b$$

with $A \in \mathbb{R}^{n \times n}$ and $x, b \in \mathbb{R}^n$.

Krylov methods:

- Very efficient when A is large and sparse.
- Typically require 1 matrix-vector multiplication with A and/or A^T per iteration.
- Construct a Krylov subspace

$$\mathcal{K}_k(A, r_0) = \left\{ r_0, Ar_0, A^2 r_0, \ldots, A^{k-1} r_0 \right\},$$

with $r_0 = b - Ax_0$ and $x_k \in x_0 + \mathcal{K}_k(A, r_0)$.

Inexact matvec

What happens if the matrix-vector product is not calculated accurately?

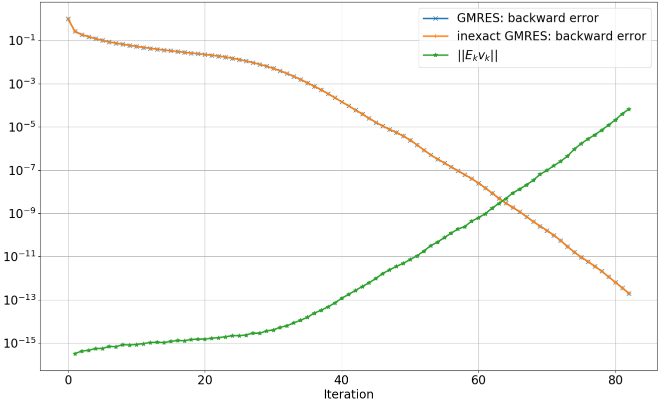
$$z_k = Av_k \qquad \longleftrightarrow \qquad \begin{cases} z_k = Av_k + p_k \\ z_k = (A + E_k)v_k \end{cases}$$

Recent results:

- **A. Bouras and V. Frayssé**, *Inexact matrix-vector products in Krylov methods for solving linear systems: a relaxation strategy*, SIAM journal on matrix analysis and applications **(2005)**.
- **V. Simoncini and D. B. Szyld**, *Theory of inexact Krylov subspace methods and applications to scientific computing*, SIAM journal on scientific computing **(2003)**.
- **L. Giraud, S. Gratton and J. Langou**, *Convergence in backward error of relaxed GMRES*, SIAM journal on scientific computing **(2007)**.

An example ...

- $A \in \mathbb{R}^{225\times 225}$ is PDE225 matrix, $x = (1, \ldots, 1)^T \in \mathbb{R}^{225}$.
- Stopping criterion: $\eta_{A,b}(x_k) = \frac{\|b - Ax_k\|}{\|A\|\|x_k\| + \|b\|} \leq \varepsilon = 1\text{e-}12$.



Mixed precision – method I

Method: perform part of the computation in 32 bit or 16 bit, while still achieving full accuracy for the reconstruction.

- Exploit faster 32 bit & 16 bit calculations on modern hardware.
- Often it is the preconditioner that is calculated in lower precision.

Recent results:

- **A. Buttari et al.**, *Mixed precision iterative refinement techniques for the solution of dense linear systems*, International journal of high performance computing applications **(2007)**.
- **M. Arioli and I. S. Duff**, *Using FGMRES to obtain backward stability in mixed precision*, Electronic transactions on numerical analysis **(2009)**.
- **E. Carson and N. J. Higham**, *Accelerating the solution of linear systems by iterative refinement in three precisions*, MIMS EPrint **(2017)**.

Mixed precision – method II

method: store part of the data in lower precision, while still achieving full accuracy for the reconstruction.

- Calculations are done in **full precision**.
- Less communication and lower memory requirements.
- Read/write operations consume more energy than computations.

Recent results:

H. Antz et al., *Adaptive precision in block-Jacobi preconditioning for iterative sparse linear system solvers*, MIMS EPrint **(2017)**.

Why limit ourselves to 64 bit, 32 bit, 16 bit, etc?

$\Downarrow$

Use data compression techniques

The SZ compressor *developed @ Argonne et al.*

Lossy compression technique:

- Higher compression factors than lossless techniques.
- Designed to deal with irregular data and spiky changes.
- Compression error can controlled:
  absolute error, relative error or point-wise relative error.

D. Tao et al., *Significantly improving lossy compression for scientific data sets based on multidimensional prediction and error-controlled quantization*, IEEE IPDPS **(2017)**.

https://github.com/disheng222/SZ

Project goal:

Incorporate the SZ compressor into a Krylov method in order to reduce the memory and communication costs.

Preliminary results

- Some theory
- Some experiments

GMRES

The GMRES algorithm revolves around the Arnoldi-relation

$$AV_k = V_{k+1} H_k.$$

Here $V_k \in \mathbb{R}^{n\times k}$, $V_k^T V_k = I_k$, span $V_k = \mathcal{K}_k(A, r_0)$, $H_k \in \mathbb{R}^{(k+1)\times k}$ is upper Hessenberg and

$$x_k = x_0 + V_k y_k \quad \text{with} \quad y_k = \arg\min_{y\in\mathbb{R}^k} \Big\| \|r_0\| - H_k y \Big\|.$$

- **inexact GMRES**:
  $\rightarrow z = (A + E_k) v_k$
  $\rightarrow A\widetilde{V}_k = \widetilde{V}_{k+1} \widetilde{H}_k$
  $\rightarrow$ **still GMRES**
- **compressed GMRES**:
  $\rightarrow$ compress v_k at some point
  $\rightarrow z \perp$ the compressed v_k ???
  $\rightarrow$ **no Arnoldi relation ...**

GMRES: Arnoldi part

```
1: v_1 = r_0 / ||r_0||
2: for k = 1, 2, ... do
3:     z = A v_k
4:     z ← Arnoldi(V_k).
5:     v_{k+1} = z / ||z||
6: end for
```

FGMRES = GMRES with varying right preconditioner M_k^{-1}

Different Arnoldi-relation:

$$AZ_k = V_{k+1} H_k.$$

Here $V_k \in \mathbb{R}^{n \times k}$, $V_k^T V_k = I_k$, $Z_k \in \mathbb{R}^{n \times k}$, $H_k \in \mathbb{R}^{(k+1) \times k}$ is upper Hessenberg and

$$x_k = x_0 + Z_k y_k \quad \text{with} \quad y_k = \arg\min_{y \in \mathbb{R}^k} \left\| \|r_0\| - H_k y \right\|.$$

- span $V_k \neq \mathcal{K}_k(A, r_0)$.
- Not a Krylov method.
- z_k can in theory be random.

FGMRES: Arnoldi part

1: $v_1 = r_0 / \|r_0\|$
2: **for** $k = 1, 2, \ldots$ **do**
3: $\quad z_k = M_k^{-1} v_k$
4: $\quad w = A z_k$
5: $\quad w \leftarrow \text{Arnoldi}(V_k)$.
6: $\quad v_{k+1} = z / \|z\|$
7: **end for**

FGMRES

Y. Saad, *A flexible inner-outer preconditioned GMRES algorithm*, SIAM journal on scientific computing **(1993)**.

- Allows iterative methods to be used as preconditioners.
- **Is not guaranteed to converge.**
- **Double the memory:** Z_k and V_{k+1}

z_k are not orthogonal $\Rightarrow$ compress z_k instead of v_k.

If compressed FGMRES converges within the same number of iterations, then we gain in memory.

Assumptions & compression model

- Assume that FGMRES converges with the preconditioners

$$\left\{ M_k^{-1} \right\}_{k=1}^{n}$$

- Let ε_k be the maximum relative point-wise error.

We introduce an error after applying the preconditioner M_k^{-1}:

$$z_k = M_k^{-1} v_k \quad \longleftrightarrow \quad \tilde{z}_k = (I + \mathrm{diag}(\delta_k))\, z_k = (I + \mathrm{diag}(\delta_k))\, M_k^{-1} v_k$$

with $|(\delta_k)_i| \leq \varepsilon_k$, for $i = 1, \ldots n$

- Still FGMRES, but with the preconditioners

$$\left\{ (I + \mathrm{diag}(\delta_k))\, M_k^{-1} \right\}_{k=1}^{n}$$

Compressed FGMRES

Compressed Arnoldi-relation:

$$A\tilde{Z}_k = \tilde{V}_{k+1}\tilde{H}_k$$

$\longrightarrow$ optimization in a different subspace of $\mathbb{R}^n$!!!

If

$$\tilde{y}_k = \underset{\tilde{y}\in\mathbb{R}^k}{\arg\min} \left\| r_0 - A\tilde{Z}\tilde{y} \right\| \quad \text{and} \quad y_k = \underset{y\in\mathbb{R}^k}{\arg\min} \| r_0 - AZ_k y \|$$

then

$$\underbrace{\left\| r_0 - A\tilde{Z}_k\tilde{y}_k \right\|}_{\text{compressed residual}} \leq \underbrace{\left\| r_0 - AZ_k y_k \right\|}_{\text{exact residual}} + \underbrace{\left\| A\left[\mathrm{diag}(\delta_1)z_1, \dots, \mathrm{diag}(\delta_k)z_k\right] y_k \right\|}_{\text{residual gap}}$$

How large can ε_k become such that $\eta_{A,b}(\tilde{x}_k) \leq \varepsilon$?

Bounding ε_k

For some targeted tolerance ε and $c \in]0, 1[$ we know that there exists a k such that

$$\eta_{A,b}(x_k) = \frac{\left\| b - Ax_k \right\|}{\|A\| \, \|x_k\| + \|b\|} \leq c\varepsilon$$

Result:

If, for all $i = 1, \ldots, k$,

$$\varepsilon_i \leq (1-c)\frac{\sigma_{min}(H_k)}{k^2 \, \|A\|} \frac{\|b\|}{\|r_{i-1}\|}\varepsilon \qquad (1)$$

then

$$\eta_{A,b}(\tilde{x}_k) = \frac{\left\| b - A\tilde{x}_k \right\|}{\|A\| \, \|\tilde{x}_k\| + \|b\|} \leq \varepsilon$$

Some practical bounds

- (1) contains values that are not available.
- (1) is a very conservative bound.
- Consider the relaxed bounds:

$$\varepsilon_i \leq (1-c)\frac{\sigma_{min}(A)}{k^2\,\|A\|}\frac{\|b\|}{\|\tilde{r}_{i-1}\|}\varepsilon \tag{b1}$$

$$\varepsilon_i \leq (1-c)\frac{1}{\|A\|}\frac{\|b\|}{\|\tilde{r}_{i-1}\|}\varepsilon \tag{b2}$$

$$\varepsilon_i \leq (1-c)\frac{\|b\|}{\|\tilde{r}_{i-1}\|}\varepsilon \tag{b3}$$

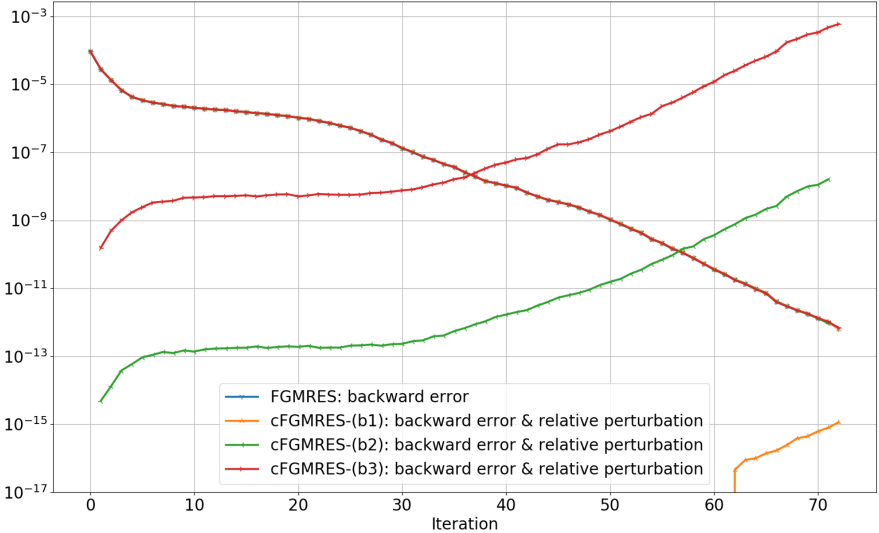
Numerical experiments

- FMGRES-GMRES(25)
- Outer stopping criterium: $\eta_{A,b}(x_k) \leq$ 1e-12 and maxit ≤ 256.
- Inner stopping criterium: $\eta_{A,v_k}(z_k) \leq$ 1e-4 and maxit ≤ 50.

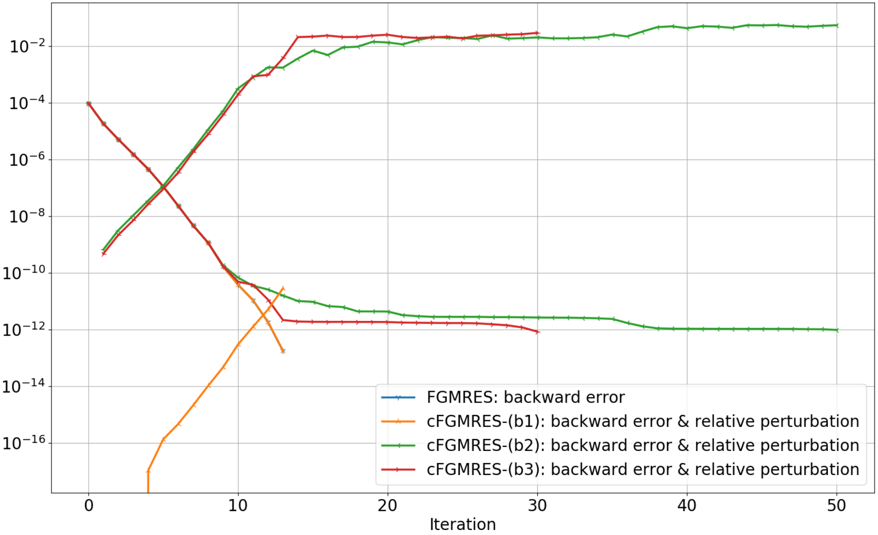
Matrix	n	FGMRES	cFGMRES-(b1)		cFGMRES-(b2)		cFGMRES-(b3)	
pde225	225×225	3	3	5.5e-11	3	1.4e-5	3	1.6e-4
e05r0000	236×236	11	11	2.1e-11	11	3.2e-3	11	3.6e-2
orsirr_1	1030×1030	33	33	2.3e-15	33	2.4e-11	33	1.2e-5
1138_bus	1138×1138	72	72	1.1e-15	71	1.6e-8	72	5.9e-4
cavity05	1182×1182	22	22	3.8e-13	22	1.6e-3	24	3.0e-2
fidap003	1821×1821	136	136	0.0	136	4.0e-10	194	9.8e-3
watt__1	1856×1856	138	138	2.9e-16	138	8.8e-2	138	8.7e-2
bwm2000	2000×2000	31	31	6.6e-14	31	8.1e-9	32	2.5e-3
olm2000	2000×2000	113	113	6.8e-15	115	4.1e-9	190	3.0e-3
add20	2395×2395	13	13	2.9e-11	50	5.5e-2	19	2.9e-2

Matrices taken from matrix market and $x = (1, \ldots, 1)^T \in \mathbb{R}^n$.

Results for 1138_bus



Results for add20



Outlook

Outlook

Done:

- Theoretical framework for the size of the compression.
- Initial numerical experiments.
- A compressor.

To be done:

- Combine the theoretical results with the compressor.
- Study the gain in memory & speedup.
- Empirical generalizations:
    - Compression of blocks.
    - Restarting strategies.
    - ...

Is this the end ...

Thank you for your attention.

Any questions?