Enlarged Krylov methods and 2-level preconditioner for the map-making problem in CMB data analysis

UC Berkeley

Thibault Cimic, Laura Grigori

March 5, 2020

Laboratoire INRIA - Equipe ALPINES

The Cosmic Microwave Background

CMB : Cosmic Microwave Background

The Cosmic Microwave Background

CMB : Cosmic Microwave Background

Other names ?

The Cosmic Microwave Background

CMB : Cosmic Microwave Background

Other names ?

Relic radiation

The Cosmic Microwave Background

CMB : Cosmic Microwave Background

Other names ?

Relic radiation $\rightarrow$ First photons that started to travel in the very early hot and dense universe (379,000 years old out of 13.8 billions)

The Cosmic Microwave Background

CMB : Cosmic Microwave Background

Cosmic : come from far away, outside our galaxy

Other names ?

Relic radiation $\rightarrow$ First photons that started to travel in the very early hot and dense universe (379,000 years old out of 13.8 billions)

The Cosmic Microwave Background

CMB : Cosmic Microwave Background

Cosmic : come from far away, outside our galaxy

Microwave : photons as observed today, they lost energy therefore increase wavelengths

Other names ?

Relic radiation $\rightarrow$ First photons that started to travel in the very early hot and dense universe (379,000 years old out of 13.8 billions)

The Cosmic Microwave Background

CMB : Cosmic Microwave Background

Cosmic : come from far away, outside our galaxy

Microwave : photons as observed today, they lost energy therefore increase wavelengths

Background : emitted the same way in every direction from anywhere

Other names ?

Relic radiation $\rightarrow$ First photons that started to travel in the very early hot and dense universe (379,000 years old out of 13.8 billions)

The Cosmic Microwave Background

CMB : Cosmic Microwave Background

Cosmic : come from far away, outside our galaxy

Microwave : photons as observed today, they lost energy therefore increase wavelengths

Background : emitted the same way in every direction from anywhere

Other names ?

Relic radiation $\rightarrow$ First photons that started to travel in the very early hot and dense universe (379,000 years old out of 13.8 billions)

The goal : Reconstruct a map of the physical properties of these early photons

The Cosmic Microwave Background

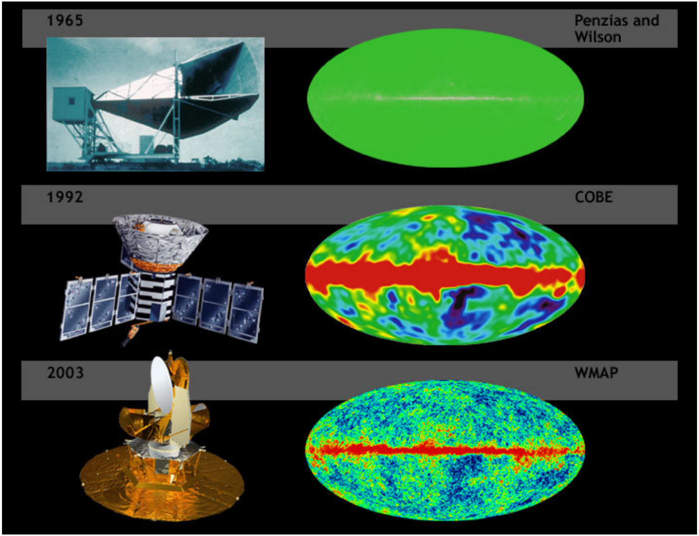


Figure 1: Evolution of the CMB map of temperature

The Cosmic Microwave Background

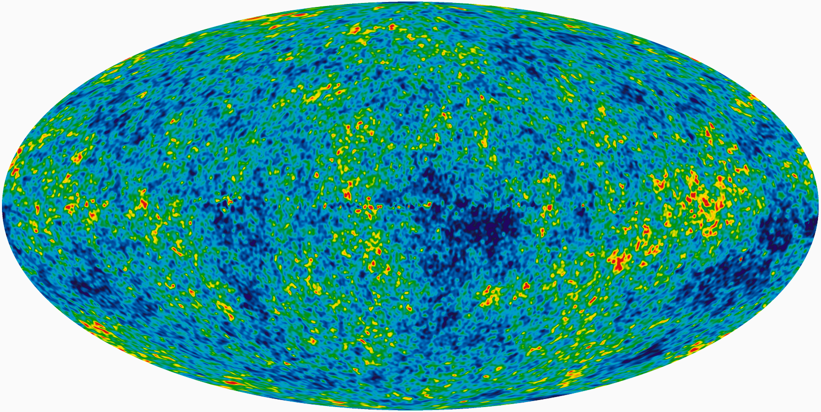


Figure 2: Map of temperature reconstructed from nine years of WMAP data satellite (2003-2012)

The Cosmic Microwave Background

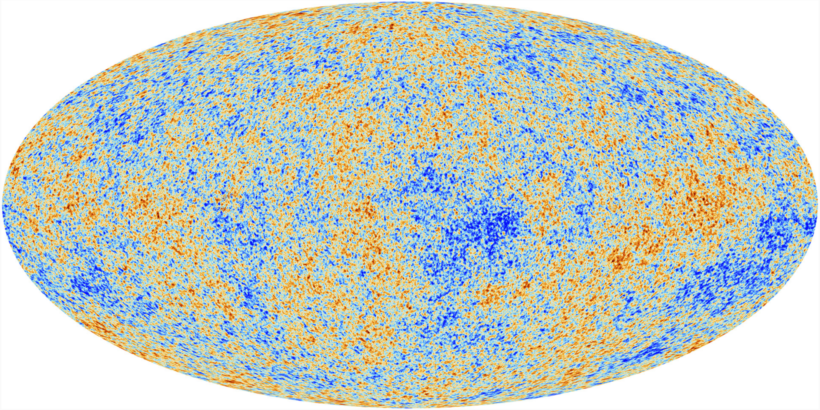


Figure 3: Map of temperature with Planck satellite

The map-making problem

The Map-making problem

Writing the vector of measurement $d \in \mathbb{R}^{n_t}$ as :

$$d = Ps + n$$

The Map-making problem

Writing the vector of measurement $d \in \mathbb{R}^{n_t}$ as :

$$d = Ps + n$$

- $s \in \mathbb{R}^{n_p}$ is the signal

The Map-making problem

Writing the vector of measurement $d \in \mathbb{R}^{n_t}$ as :

$$d = Ps + n$$

- $s \in \mathbb{R}^{n_p}$ is the signal
- $P \in \mathbb{R}^{n_t \times n_p}$ is the depointing matrix, tall and skinny and very sparse

The Map-making problem

Writing the vector of measurement $d \in \mathbb{R}^{n_t}$ as :

$$d = Ps + n$$

- $s \in \mathbb{R}^{n_p}$ is the signal
- $P \in \mathbb{R}^{n_t \times n_p}$ is the depointing matrix, tall and skinny and very sparse
- $n \in \mathbb{R}^{n_t}$ is the noise, modelled as a Gaussian stochastic process

The Map-making problem

Writing the vector of measurement $d \in \mathbb{R}^{n_t}$ as :

$$d = Ps + n$$

- $s \in \mathbb{R}^{n_p}$ is the signal
- $P \in \mathbb{R}^{n_t \times n_p}$ is the depointing matrix, tall and skinny and very sparse
- $n \in \mathbb{R}^{n_t}$ is the noise, modelled as a Gaussian stochastic process

Maximum likelihood estimate, $\hat{s}$, of the signal s given by :

$$\underbrace{\left(P^\top N^{-1} P\right)}_{A} \hat{s} = P^t N^{-1} d$$

Where $N \in \mathbb{R}^{n_t \times n_t}$ is the covariance matrix of the noise.

The Map-making problem

Map-making : scanning strategy & maximum likelihood

What's a scanning strategy ?

The Map-making problem

Map-making : scanning strategy & maximum likelihood

What's a scanning strategy ?

The way we observe the sky, encoded by the depointing matrix P as such : $\forall 1 \leq i \leq n_t$

$$P_{i,.} = (0, ..., 0, t_i, 0, ..., 0) \in \mathbb{R}^{n_p} \tag{1}$$

The Map-making problem

Map-making : scanning strategy & maximum likelihood

What's a scanning strategy ?

The way we observe the sky, encoded by the depointing matrix P as such : $\forall 1 \leq i \leq n_t$

$$P_{i,.} = (0, ..., 0, t_i, 0, ..., 0) \in \mathbb{R}^{n_p} \tag{1}$$

In fact, several parameters per pixel. Typically, t,q and u.

$$P_{i,.} = (0, ..., 0, t_i, q_i, u_i, 0, ..., 0) \in \mathbb{R}^{3n_p} \tag{2}$$

The Map-making problem

What does N^{-1} looks like ?

The Map-making problem

What does N^{-1} looks like ?

N^{-1} is a block diagonal symmetric matrix :

The Map-making problem

What does N^{-1} looks like ?

N^{-1} is a block diagonal symmetric matrix :

$$N^{-1} = \begin{bmatrix} N_1^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & N_{n_{blc}}^{-1} \end{bmatrix} \in \mathbb{R}^{n_t \times n_t}$$

The Map-making problem

What does N^{-1} looks like ?

N^{-1} is a block diagonal symmetric matrix :

$$N^{-1} = \begin{bmatrix} N_1^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & N_{n_{blc}}^{-1} \end{bmatrix} \in \mathbb{R}^{n_t \times n_t}$$

Each block is band-diagonal, toeplitz :

$$N_l^{-1} = \begin{bmatrix} d_l & \cdots & e_{\lambda_l}^l & & \\ \vdots & \ddots & & \ddots & \\ e_{\lambda_l}^l & & \ddots & & e_{\lambda_l}^l \\ & \ddots & & \ddots & \vdots \\ & & e_{\lambda_l}^l & \cdots & d_l \end{bmatrix} \in \mathbb{R}^{n_{t_i} \times n_{t_i}}$$

From CG to Enlarged-CG

The Enlarged-CG

Define T_t, for $t \in \mathbb{N}$ the splitting parameter :

$$T_t : \begin{array}{c} \mathbb{R}^n \rightarrow \mathbb{R}^{n \times t} \\ x \mapsto T_t(x) \end{array} \tag{3}$$

with $T_t(x)$ s.t. $T_t(x) * \mathbf{1}_t = x$ and $T_t(x)$ has t linearly independent columns.

The Enlarged-CG

Define T_t, for $t \in \mathbb{N}$ the splitting parameter :

$$T_t : \begin{array}{c} \mathbb{R}^n \to \mathbb{R}^{n\times t} \\ x \mapsto T_t(x) \end{array} \tag{3}$$

with $T_t(x)$ s.t. $T_t(x) * \mathbf{1}_t = x$ and $T_t(x)$ has t linearly independent columns.

$$x = \begin{pmatrix} * \\ * \\ \vdots \\ * \\ * \\ * \end{pmatrix} \mapsto T_t(x) = \begin{pmatrix} * & & 0 \\ * & & \vdots \\ 0 & \dots & * \\ \vdots & & \vdots \\ 0 & & * \end{pmatrix}$$

The Enlarged-CG

$A \in \mathbb{R}^{n \times n}$ SPD, $b \in \mathbb{R}^n$, solve $A\underline{x} = b$

The Enlarged-CG

$A \in \mathbb{R}^{n \times n}$ SPD, $b \in \mathbb{R}^n$, solve $A\underline{x} = b$

Enlarged the Krylov space with T_t :

$$K_{k,t} = \text{Span}\left(T_t(r_0), AT_t(r_0), ..., A^{k-1}T_t(r_0)\right) \subset \mathbb{R}^n$$

The Enlarged-CG

$A \in \mathbb{R}^{n \times n}$ SPD, $b \in \mathbb{R}^n$, solve $A\underline{x} = b$

Enlarged the Krylov space with T_t :

$$K_{k,t} = \text{Span}\left(T_t(r_0), AT_t(r_0), ..., A^{k-1}T_t(r_0)\right) \subset \mathbb{R}^n$$

For $x_0 \in \mathbb{R}^n$, build the sequence $(x_k)_{k \geq 0}$ s.t. :

$$\begin{cases} x_{k+1} \in x_0 + K_{k,t} \\ r_{k+1} = b - Ax_k \perp K_{k,t} \end{cases} \tag{4}$$

The Enlarged-CG

$A \in \mathbb{R}^{n\times n}$ SPD, $b \in \mathbb{R}^n$, solve $A\underline{x} = b$

Enlarged the Krylov space with T_t :

$$K_{k,t} = \text{Span}\left(T_t(r_0), AT_t(r_0), ..., A^{k-1}T_t(r_0)\right) \subset \mathbb{R}^n$$

For $x_0 \in \mathbb{R}^n$, build the sequence $(x_k)_{k\geq 0}$ s.t. :

$$\begin{cases} x_{k+1} \in x_0 + K_{k,t} \\ r_{k+1} = b - Ax_k \perp K_{k,t} \end{cases} \tag{4}$$

Lemma (L. Grigori, S. Moufawad, F. Nataf)

For x_k the k-th approximation build from (4), x_{k+1} satisfies :

$$K_k \subset K_{k,t}$$

$$||x_{k+1} - \underline{x}||_A = \min_{x \in x_0 + K_{k,t}} ||x - \underline{x}||_A$$

The Enlarged-CG

Theorem (O. Tissot, L. Grigori)

Let x_k be the k-th iterate build with (4), then we have :

$$||x_k - \underline{x}||_A \leq C \left(\frac{\sqrt{\kappa_t} - 1}{\sqrt{\kappa_t} + 1} \right)^k \tag{5}$$

with $\kappa_t = \lambda_n / \lambda_t$ where λ_t is the t^{th} smallest eigenvalue of A and C is a constant independent of k.

The Enlarged-CG

Theorem (O. Tissot, L. Grigori)

Let x_k be the k-th iterate build with (4), then we have :

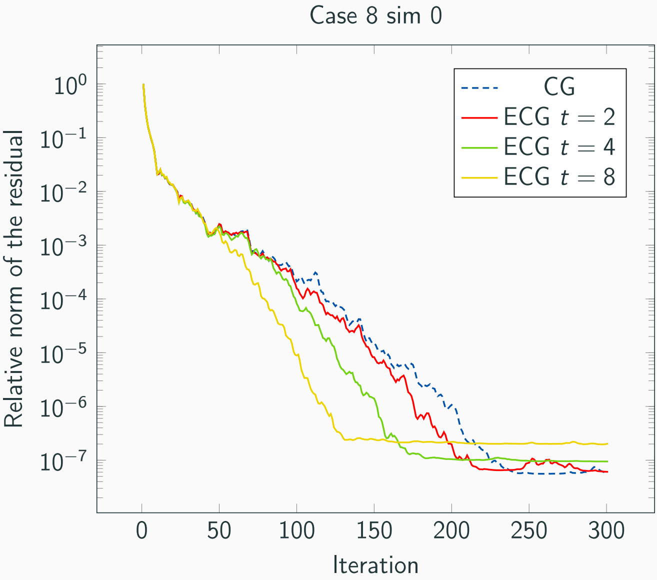
$$||x_k - \underline{x}||_A \leq C \left(\frac{\sqrt{\kappa_t} - 1}{\sqrt{\kappa_t} + 1} \right)^k \tag{5}$$

with $\kappa_t = \lambda_n / \lambda_t$ where λ_t is the t^{th} smallest eigenvalue of A and C is a constant independent of k.

Conclusion :

- Deflation of t eigenvectors $\sim$ ECG(t) w.r.t. nb of iteration
- Each iteration is now t times more costly

The Enlarged-CG



2-level preconditioner from fictitious space lemma

The fictitious space lemma

Given an SPD operator, $\mathcal{A} : H \longrightarrow H$ ($u \mapsto Au$),

$a : H \times H \longrightarrow \mathbb{R}$.

The fictitious space lemma

Given an SPD operator, $\mathcal{A} : H \longrightarrow H$ ($u \mapsto Au$), $a : H \times H \longrightarrow \mathbb{R}$.

How to build $\mathcal{M}^{-1} : H \longrightarrow H$ ($u \mapsto M^{-1}u$) s.t. $\Lambda(\mathcal{M}^{-1}\mathcal{A}) \subset [c_L, c_U]$?

The fictitious space lemma

Given an SPD operator, $\mathcal{A} : H \longrightarrow H$ $(u \mapsto Au)$,
$a : H \times H \longrightarrow \mathbb{R}$.
How to build $\mathcal{M}^{-1} : H \longrightarrow H$ $(u \mapsto M^{-1}u)$ s.t.
$\Lambda(\mathcal{M}^{-1}\mathcal{A}) \subset [c_L, c_U]$?
$\rightarrow \mathcal{M}^{-1} = \mathcal{R}\mathcal{B}^{-1}\mathcal{R}^*$ avec $\mathcal{R} : H_D \longrightarrow H$ et $\mathcal{B} : H_D \longrightarrow H_D$
$(b : H_D \times H_D \longrightarrow \mathbb{R})$, s.t. :

The fictitious space lemma

Given an SPD operator, $\mathcal{A} : H \longrightarrow H$ ($u \mapsto Au$),
$a : H \times H \longrightarrow \mathbb{R}$.
How to build $\mathcal{M}^{-1} : H \longrightarrow H$ ($u \mapsto M^{-1}u$) s.t.
$\Lambda(\mathcal{M}^{-1}\mathcal{A}) \subset [c_L, c_U]$?
$\rightarrow \mathcal{M}^{-1} = \mathcal{R}\mathcal{B}^{-1}\mathcal{R}^*$ avec $\mathcal{R} : H_D \longrightarrow H$ et $\mathcal{B} : H_D \longrightarrow H_D$
($b : H_D \times H_D \longrightarrow \mathbb{R}$), s.t. :

1. $\mathcal{R}$ surjective

The fictitious space lemma

Given an SPD operator, $\mathcal{A} : H \longrightarrow H$ $(u \mapsto Au)$,
$a : H \times H \longrightarrow \mathbb{R}$.
How to build $\mathcal{M}^{-1} : H \longrightarrow H$ $(u \mapsto M^{-1}u)$ s.t.
$\Lambda(\mathcal{M}^{-1}\mathcal{A}) \subset [c_L, c_U]$?
$\rightarrow \mathcal{M}^{-1} = \mathcal{R}\mathcal{B}^{-1}\mathcal{R}^*$ avec $\mathcal{R} : H_D \longrightarrow H$ et $\mathcal{B} : H_D \longrightarrow H_D$
$(b : H_D \times H_D \longrightarrow \mathbb{R})$, s.t. :

1. $\mathcal{R}$ surjective
2. $\exists c_U > 0, \forall u_D \in H_D, a(\mathcal{R}u_D, \mathcal{R}u_D) \leq c_U b(u_D, u_D)$

The fictitious space lemma

Given an SPD operator, $\mathcal{A} : H \longrightarrow H$ $(u \mapsto Au)$,
$a : H \times H \longrightarrow \mathbb{R}$.
How to build $\mathcal{M}^{-1} : H \longrightarrow H$ $(u \mapsto M^{-1}u)$ s.t.
$\Lambda(\mathcal{M}^{-1}\mathcal{A}) \subset [c_L, c_U]$?
$\rightarrow \mathcal{M}^{-1} = \mathcal{R}\mathcal{B}^{-1}\mathcal{R}^*$ avec $\mathcal{R} : H_D \longrightarrow H$ et $\mathcal{B} : H_D \longrightarrow H_D$
$(b : H_D \times H_D \longrightarrow \mathbb{R})$, s.t. :

1. $\mathcal{R}$ surjective
2. $\exists c_U > 0, \forall u_D \in H_D, a(\mathcal{R}u_D, \mathcal{R}u_D) \leq c_U b(u_D, u_D)$
3. $\exists c_L > 0, \forall u \in H, \exists u_D \in H_D, \mathcal{R}u_D = u,$

$$c_L b(u_D, u_D) \leq a(\mathcal{R}u_D, \mathcal{R}u_D) = a(u, u)$$

The fictitious space lemma

Given an SPD operator, $\mathcal{A} : H \longrightarrow H$ ($u \mapsto Au$),
$a : H \times H \longrightarrow \mathbb{R}$.
How to build $\mathcal{M}^{-1} : H \longrightarrow H$ ($u \mapsto M^{-1}u$) s.t.
$\Lambda(\mathcal{M}^{-1}\mathcal{A}) \subset [c_L, c_U]$?
$\rightarrow \mathcal{M}^{-1} = \mathcal{R}\mathcal{B}^{-1}\mathcal{R}^*$ avec $\mathcal{R} : H_D \longrightarrow H$ et $\mathcal{B} : H_D \longrightarrow H_D$
($b : H_D \times H_D \longrightarrow \mathbb{R}$), s.t. :

1. $\mathcal{R}$ surjective
2. $\exists c_U > 0, \forall u_D \in H_D, a(\mathcal{R}u_D, \mathcal{R}u_D) \leq c_U b(u_D, u_D)$
3. $\exists c_L > 0, \forall u \in H, \exists u_D \in H_D, \mathcal{R}u_D = u,$

$$c_L b(u_D, u_D) \leq a(\mathcal{R}u_D, \mathcal{R}u_D) = a(u, u)$$

$\Rightarrow \Lambda(\mathcal{R}\mathcal{B}^{-1}\mathcal{R}^*\mathcal{A}) \subset [c_L, c_U]$

Map-making context

- n_{blc} blocks of N^{-1} splits $\{1, ..., n_t\}$ in n_{blc} domains :

$$\mathbb{R}^{n_t} \cong \prod_{l=0}^{n_{blc}} \mathbb{R}^{n_l} =: H_D$$

Map-making context

- n_{blc} blocks of N^{-1} splits $\{1, ..., n_t\}$ in n_{blc} domains :

$$\mathbb{R}^{n_t} \cong \prod_{l=0}^{n_{blc}} \mathbb{R}^{n_l} =: H_D$$

$P^\top = (P_1^\top | ... | P_{n_{blc}}^\top)$ and $P^\top N^{-1} P = \sum_{l=1}^{n_{blc}} P_l^\top N_l^{-1} P_l$

Map-making context

- n_{blc} blocks of N^{-1} splits $\{1, ..., n_t\}$ in n_{blc} domains :

$$\mathbb{R}^{n_t} \cong \prod_{l=0}^{n_{blc}} \mathbb{R}^{n_l} =: H_D$$

$P^\top = (P_1^\top|...|P_{n_{blc}}^\top)$ and $P^\top N^{-1} P = \sum_{l=1}^{n_{blc}} P_l^\top N_l^{-1} P_l$

- Define :

$$\mathcal{R} : \begin{array}{c} H_D \longrightarrow H \\ (u_l)_{l=0...n_{blc}} \mapsto Z_0^\top u_0 + \sum_{l=1}^{n_{blc}} P_l^\top u_l \end{array}$$

Map-making context

- n_{blc} blocks of N^{-1} splits $\{1, ..., n_t\}$ in n_{blc} domains :

$$\mathbb{R}^{n_t} \cong \prod_{l=0}^{n_{blc}} \mathbb{R}^{n_l} =: H_D$$

$P^\top = (P_1^\top | ... | P_{n_{blc}}^\top)$ and $P^\top N^{-1} P = \sum_{l=1}^{n_{blc}} P_l^\top N_l^{-1} P_l$

- Define :

$$\mathcal{R} : \begin{array}{c} H_D \longrightarrow H \\ (u_l)_{l=0...n_{blc}} \mapsto Z_0^\top u_0 + \sum_{l=1}^{n_{blc}} P_l^\top u_l \end{array}$$

Lemma (Surjectivity of $\mathcal{R}$)

$\mathcal{R}$ define as above is surjective.

Map-making context

- n_{blc} blocks of N^{-1} splits $\{1, ..., n_t\}$ in n_{blc} domains :

$$\mathbb{R}^{n_t} \cong \prod_{l=0}^{n_{blc}} \mathbb{R}^{n_l} =: H_D$$

$P^\top = (P_1^\top | ... | P_{n_{blc}}^\top)$ and $P^\top N^{-1} P = \sum_{l=1}^{n_{blc}} P_l^\top N_l^{-1} P_l$

- Define :

$$\mathcal{R} : \begin{array}{c} H_D \longrightarrow H \\ (u_l)_{l=0 \dots n_{blc}} \mapsto Z_0^\top u_0 + \sum_{l=1}^{n_{blc}} P_l^\top u_l \end{array}$$

Lemma (Surjectivity of $\mathcal{R}$)

$\mathcal{R}$ define as above is surjective.

Uses :

- The structure of P
- The fact that the sets of $(P_l)_{l=1 \dots n_{blc}}$ observe all the pixels.

AS2

- Define operator B :

AS2

- Define operator B :
  $B(\mathcal{U}) := (\underbrace{Z_0^\top A Z_0}_{E} u_0, diag(\mathcal{N}_1^{-1}) u_1, ..., diag(\mathcal{N}_{n_{blc}}^{-1}) u_{n_{blc}})$

AS2

- Define operator B :
  $B(\mathcal{U}) := (\underbrace{Z_0^\top A Z_0}_{E} u_0, diag(\mathcal{N}_1^{-1}) u_1, ..., diag(\mathcal{N}_{n_{blc}}^{-1}) u_{n_{blc}})$
- Set the preconditioner M_2^{-1} :

$$M_2^{-1} = \mathcal{R} B^{-1} \mathcal{R}^* = \underbrace{Z_0 E^{-1} Z_0^\top}_{Q} + \underbrace{P^\top diag(\mathcal{N}^{-1})^{-1} P}_{M_1^{-1}}$$

AS2

- Define operator B :
  $B(\mathcal{U}) := (\underbrace{Z_0^\top A Z_0}_{E} u_0, diag(\mathcal{N}_1^{-1})u_1, ..., diag(\mathcal{N}_{n_{blc}}^{-1})u_{n_{blc}})$
- Set the preconditioner M_2^{-1} :

$$M_2^{-1} = \mathcal{R}B^{-1}\mathcal{R}^* = \underbrace{Z_0 E^{-1} Z_0^\top}_{Q} + \underbrace{P^\top diag(\mathcal{N}^{-1})^{-1} P}_{M_1^{-1}}$$

Main question : How to efficiently compute a good coarse space Z_0 in this framework ?

AS2

Lemma (Continuity of $\mathcal{R}$)

As such : $\exists c_U$ *s.t.* $\forall u_D \in H_D, a(\mathcal{R}u_D, \mathcal{R}u_D) \leq c_U b(u_D, u_D)$

AS2

Lemma (Continuity of $\mathcal{R}$)

As such : $\exists c_U$ *s.t.* $\forall u_D \in H_D, a(\mathcal{R}u_D, \mathcal{R}u_D) \leq c_U b(u_D, u_D)$

- Not new in cosmology : see J. Papež, L. Grigori and R. Stompor, Astronomy & Astrophysics, 2018
- In the experiment : $c_u = 2$

AS2

Lemma (Continuity of $\mathcal{R}$)

As such : $\exists c_U$ *s.t.* $\forall u_D \in H_D, a(\mathcal{R}u_D, \mathcal{R}u_D) \leq c_U b(u_D, u_D)$

- Not new in cosmology : see J. Papež, L. Grigori and R. Stompor, Astronomy & Astrophysics, 2018
- In the experiment : $c_u = 2$

Lemma (Stable decomposition)

As such : $\exists c_L > 0, \forall u \in H, \exists u_D \in H_D, \mathcal{R}u_D = u$

$$c_L b(u_D, u_D) \leq a(\mathcal{R}u_D, \mathcal{R}u_D) = a(u, u)$$

AS2

1. Start from :

$$b(u_D, u_D) \leq 2a(u, u) + K \sum_{l=1}^{n_{blc}} u_l^\top diag(N_l^{-1}) u_l$$

AS2

1. Start from :

$$b(u_D, u_D) \leq 2a(u, u) + K \sum_{l=1}^{n_{blc}} u_l^\top diag(N_l^{-1}) u_l$$

2. For $l = 1 ... n_{blc}$, solve generalized eigen problem :

$$\begin{cases} \text{Find } (v_{l,i}, \lambda_i) \text{ s.t. :} \\ N_l^{-1} v_{l,i} = \lambda_i diag(N_l^{-1}) v_{l,i} \end{cases}$$

AS2

1. Start from :

$$b(u_D, u_D) \leq 2a(u,u) + K \sum_{l=1}^{n_{blc}} u_l^\top diag(N_l^{-1}) u_l$$

2. For $l = 1...n_{blc}$, solve generalized eigen problem :

$$\begin{cases} \text{Find } (v_{l,i}, \lambda_i) \text{ s.t. :} \\ N_l^{-1} v_{l,i} = \lambda_i diag(N_l^{-1}) v_{l,i} \end{cases}$$

3. Construct $\Pi_l \perp$ projection on $Z_l = Span(v_{l,i} | \lambda_i \leq \frac{1}{\tau})$:

$$u_l := M_l(I - \Pi_l)P_l u \quad (u_0 = \ ?)$$

AS2

1. Start from :

$$b(u_D, u_D) \leq 2a(u,u) + K \sum_{l=1}^{n_{blc}} u_l^\top diag(N_l^{-1}) u_l$$

2. For $l = 1 \ldots n_{blc}$, solve generalized eigen problem :

$$\begin{cases} \text{Find } (v_{l,i}, \lambda_i) \text{ s.t. :} \\ N_l^{-1} v_{l,i} = \lambda_i diag(N_l^{-1}) v_{l,i} \end{cases}$$

3. Construct $\Pi_l \perp$ projection on $Z_l = Span(v_{l,i} | \lambda_i \leq \frac{1}{\tau})$:

$$u_l := M_l(I - \Pi_l) P_l u \quad (u_0 = \ ?)$$

4. Then you have :

$$u_l^\top diag(N_l^{-1}) u_l \leq \tau u^\top P_l^\top N_l^{-1} P_l u$$

AS2

1. Start from :

$$b(u_D, u_D) \leq 2a(u,u) + K \sum_{l=1}^{n_{blc}} u_l^\top diag(N_l^{-1}) u_l$$

2. For $l = 1...n_{blc}$, solve generalized eigen problem :

$$\left\{ \begin{array}{c} \text{Find } (v_{l,i}, \lambda_i) \text{ s.t. :} \\ N_l^{-1} v_{l,i} = \lambda_i diag(N_l^{-1}) v_{l,i} \end{array} \right.$$

3. Construct $\Pi_l \perp$ projection on $Z_l = Span(v_{l,i} | \lambda_i \leq \frac{1}{\tau})$:

$$u_l := M_l(I - \Pi_l) P_l u \quad (u_0 = \ ?)$$

4. Then you have :

$$u_l^\top diag(N_l^{-1}) u_l \leq \tau u^\top P_l^\top N_l^{-1} P_l u$$

5. Finally, sum :

$$b(u_D, u_D) \leq (2 + K\tau) a(u,u)$$

AS2

By choosing u_0 for having a decomposition :

$$\mathcal{R}(u_D) = u$$
$$\iff u_0 = (Z_0 Z_0^\top)^{-1} Z_0 \sum_{l=1}^{n_{blc}} P_l^\top M_l \Pi_l P_l u$$

AS2

By choosing u_0 for having a decomposition :

$$\mathcal{R}(u_D) = u$$
$$\iff u_0 = (Z_0 Z_0^\top)^{-1} Z_0 \sum_{l=1}^{n_{blc}} P_l^\top M_l \Pi_l P_l u$$

$Z_0 \perp$ basis of $Z = \bigoplus_{l=1}^{n_{blc}} P_l^\top M_l Z_l$

AS2

By choosing u_0 for having a decomposition :

$$\mathcal{R}(u_D) = u$$
$$\iff u_0 = (Z_0 Z_0^\top)^{-1} Z_0 \sum_{l=1}^{n_{blc}} P_l^\top M_l \Pi_l P_l u$$

$Z_0 \perp$ basis of $Z = \bigoplus_{l=1}^{n_{blc}} P_l^\top M_l Z_l$

Without any details :

$$P_l P_l^\top M_l = Id_{n_{t_l}}$$

But never constructed :

$$P_l^\top M_l = (P_l^\top P_l)^\dagger P_l^\top$$

AS2

By choosing u_0 for having a decomposition :

$$\mathcal{R}(u_D) = u$$
$$\iff u_0 = (Z_0 Z_0^\top)^{-1} Z_0 \sum_{l=1}^{n_{blc}} P_l^\top M_l \Pi_l P_l u$$

$Z_0 \perp$ basis of $Z = \bigoplus_{l=1}^{n_{blc}} P_l^\top M_l Z_l$

Without any details :

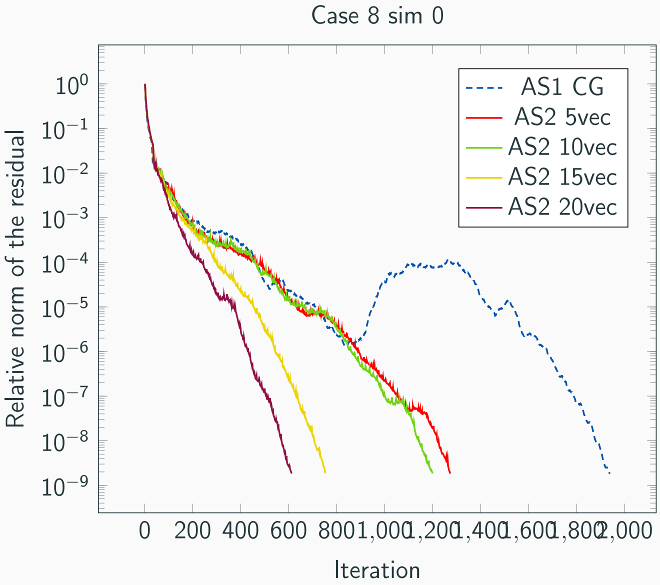
$$P_l P_l^\top M_l = Id_{n_{t_l}}$$

But never constructed :

$$P_l^\top M_l = (P_l^\top P_l)^\dagger P_l^\top$$

And : $Z = \bigoplus_{l=1}^{n_{blc}} (P_l^\top P_l)^\dagger P_l^\top Z_l$

AS2



Case 8 sim 0
Relative norm of the residual
Iteration
AS1 CG
AS2 5vec
AS2 10vec
AS2 15vec
AS2 20vec
10^0
10^−1
10^−2
10^−3
10^−4
10^−5
10^−6
10^−7
10^−8
10^−9
0
200
400
600
800
1,000
1,200
1,400
1,600
1,800
2,000

References

(1) L. Grigori, S. Moufawad, F. Nataf : Enlarged Krylov subspace conjugate gradient methods for reducing communication, SIAM Journal on Matrix Analysis and Applications, 2016

(2) L. Grigori, O. Tissot : Scalable Linear Solvers based on Enlarged Krylov subspaces with Dynamic Reduction of Search Directions, Research Rapport RR-9190, 2018

(3) L. Grigori, H. Al Daas : A class of efficient locally constructed preconditioners based on coarse spaces, SIAM Journal on Matrix Analysis and Applications, 2019

(4) V. Dolean, P. Jolivet, F. Nataf : An introduction to domain decomposition methods, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2015

Thank you for you attention !

Any questions : thibault [dot] cimic [at] inria [dot] fr

The Enlarged-CG

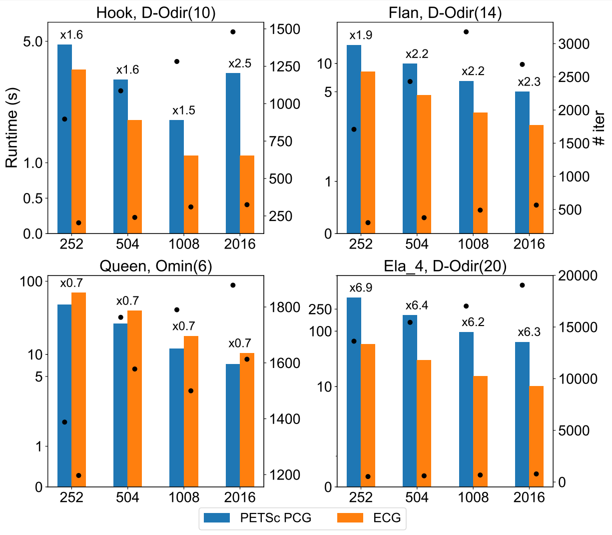


Figure 4: Tim Davis' collection + 3D elasticity problem, O. Tissot

AS2

Lemma

$(H, (.,.))$, $(H_D, (.,.)_D)$ two Hilbert spaces, two symmetric positive bilinear forms $a : H \times H \longrightarrow \mathbb{R}$, $b : H_D \times H_D \longrightarrow \mathbb{R}$, generated by the SPD operators $\mathcal{A} : H \longrightarrow H$ and $B : H_D \longrightarrow H_D$, respectively. Suppose that there exists a linear operator $\mathcal{R} : H_D \longrightarrow H$ such that the following holds :

- *$\mathcal{R}$ is surjective.*
- *$\exists c_U$ s.t. $\forall u_D \in H_D, a(\mathcal{R}u_D, \mathcal{R}u_D) \leq c_U b(u_D, u_D)$*
- *$\exists c_L$ s.t. $\forall u \in H, \exists u_D \in H_D$ s.t. $\mathcal{R}u_D = u$, $c_L b(u_D, u_D) \leq a(\mathcal{R}u_D, \mathcal{R}u_D) = a(u, u)$*

$\mathcal{R}^* : H \longrightarrow H_D$ the adjoint operator of $\mathcal{R}$, then :

$$\Lambda\left(\mathcal{R}B^{-1}\mathcal{R}^*\mathcal{A}\right) \subset [c_L, c_U]$$

Algorithm 1 Enlarged CG

Require: $A \in \mathbb{R}^{n \times n}, b \in \mathbb{R}^n, x_0 \in \mathbb{R}^n, k_{max} \in \mathbb{N}, \varepsilon > 0,\ t \in \mathbb{N}$

Ensure: $||b - Ax_k||_2 < \varepsilon ||b||_2$ or $k = k_{max}$

$k = 0,\ p_0 = r_0 = b - Ax_0$

$X_0 = T_t(x_0),\ P_0 = T_t(p_0),\ R_0 = T_t(r_0)$

while $||r_{k+1}|| > \varepsilon ||b||$ ou $k < k_{max}$ **do**

A-orthonormalize P_k

$\alpha_k = P_k^t R_k$

$X_{k+1} = X_k + P_k \alpha_k$

$R_{k+1} = R_k - AP_k \alpha_k$

$r_{k+1} = R_{k+1} \mathbf{1}_t$

$P_{k+1} = R_{k+1} - P_k P_k^\top A R_{k+1}$

$k = k + 1$

end while

Return $x_{k+1} = X_{k+1} * \mathbf{1}_t$