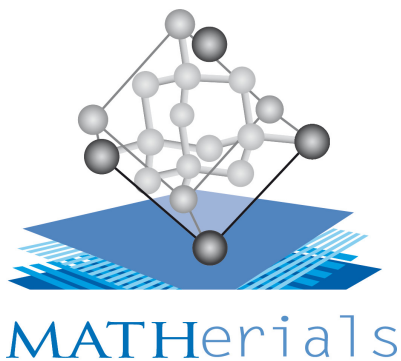
Reduced Basis method for non-symmetric eigenvalue problems: application to neutronics

Yonah Conjungo-Taumhas[1], Geneviève Dusson[3], Virginie Ehrlacher[2], Tony Lelièvre[2] & François Madiot[1]

[1] Université Paris-Saclay, CEA, Service d'Études des Réacteurs et de Mathématiques Appliquées, 91191 Gif-sur-Yvette, France

[2] CERMICS, Ecole des Ponts ParisTech & MATHERIALS team-project, INRIA

[3] Université de Franche-Comté, Besançon



Journées scientifiques du GdR MANU, Le Croisic, 24th October 2023

Outline

Outline

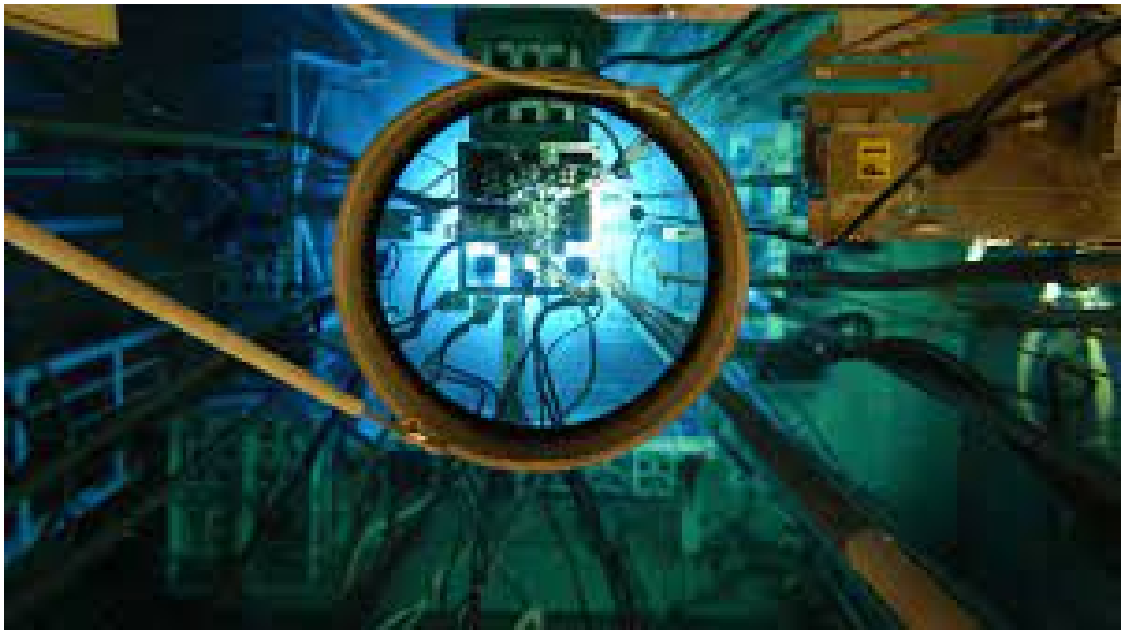
Parametrized PDEs

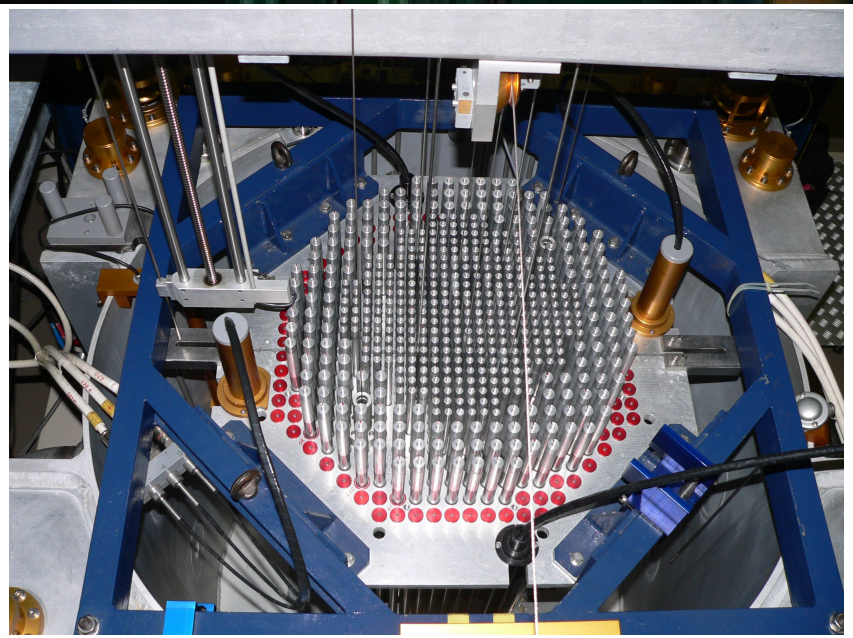
- The behaviour of many systems can be described by the solutions of a system of Partial Differential Equations.
- These equations can depend on one or several parameters $\mu = (\mu_1, \cdots, \mu_p)$ with$p \in \mathbb{N}^*$ which can take values in a set denoted by $\mathcal{P} \subset \mathbb{R}^p$.
  In this case, for one particular value $\mu \in \mathcal{P}$ of this vector of parameters, the associated solution to the PDE system is a function u_μ solution of

$$\mathcal{A}(u_\mu; \mu) = 0,$$

where $\mathcal{A}(\cdot; \mu)$ is some differential operator depending on the parameter vector μ.

Research nuclear core reactor





Simple example: Two-group diffusion model

[Coste-Delclaux, Diop, Nicolas, Bonin, 2013], [Mula, 2014], [Giret, 2018], [Allaire, Blanc, Desprès, Golse, 2019]

- **Spatial domain** $\Omega \subset \mathbb{R}^d$ occupied by the nuclear core reactor
- Neutrons are assumed to be separated into **2 groups** according to their energy: $E = \{E_1, E_2\}$ $(E_1 > E_2)$
- $\mu \in \mathcal{P}$: vector of **parameters** of the problem, which encodes the values of the **physical properties** of the nuclear core

Problem of interest: Find

- $u_\mu = (u_{1,\mu}, u_{2,\mu}) : \Omega \to \mathbb{R}^2$: **neutron scalar fluxes**;
- $\lambda_\mu > 0$ eigenvalue with **smallest modulus**;

solution to the non-symmetric eigenvalue problem

$$\mathcal{A}_\mu u_\mu = \lambda_\mu \mathcal{B}_\mu u_\mu$$

where $\mathcal{A}_\mu$ and $\mathcal{B}_\mu$ are linear operators such that $\mathcal{A}_\mu^{-1}\mathcal{B}_\mu$ satisfies the assumptions of the **Krein-Rutman theorem**.

Krein-Rutman theorem

The **Krein-Rutman theorem** is a generalisation of the **Perron-Frobenius theorem** to operators defined on infinite-dimensional Banach spaces.

Consequence of the Krein-Rutman theorem: there exists a unique eigenvalue with largest modulus $k_{\text{eff},\mu}$, which is **simple, positive** and such that

$$\mathcal{A}_\mu^{-1}\mathcal{B}_\mu u_\mu = k_{\text{eff},\mu} u_\mu$$

- u_μ is then uniquely defined (up to a sign factor);
- $k_{\text{eff},\mu} = \frac{1}{\lambda_\mu}$: **effective multiplication factor**

Effective multiplication factor

- $k_{\text{eff},\mu} < 1$: the fission reaction is not the prevailing phenomenon, then the total mean number of neutrons tends towards zero along time;
  the reactor is said to be **subcritical**
- $k_{\text{eff},\mu} = 1$: both creation and absorption of neutrons take as much place as the other inside the system;
  the reactor is said to be **critical**
- $k_{\text{eff},\mu} > 1$: the fission dominates the absorption phenomenon, therefore a chain reaction phenomenon takes place inside the system, and the total mean number of neutrons increases at an exponential rate, the system then tends to collapse;
  the reactor is said to be **supercritical**

Two-Group Diffusion Equation

$$\mathcal{A}_\mu u_\mu = \lambda_\mu \mathcal{B}_\mu u_\mu$$

Two-group Diffusion Equation

$$\begin{aligned} &-\nabla \cdot (D_{1,\mu} \nabla u_{1,\mu}) + \Sigma_{11,\mu} u_{1,\mu} + \Sigma_{12,\mu} u_{2,\mu} \\ &= \lambda_\mu \left[\chi_{1,\mu} \left((\nu\Sigma_f)_{1,\mu} u_{1,\mu} + (\nu\Sigma_f)_{2,\mu} u_{2,\mu}\right)\right] \\ \\ &-\nabla . (D_{2,\mu} \nabla u_{2,\mu}) + \Sigma_{22,\mu} u_{2,\mu} + \Sigma_{21,\mu} u_{1,\mu} \\ &= \lambda_\mu \left[\chi_{2,\mu} \left((\nu\Sigma_f)_{1,\mu} u_{1,\mu} + (\nu\Sigma_f)_{2,\mu} u_{2,\mu}\right)\right] \end{aligned} \tag{1}$$

- $\Sigma_{ii} = \Sigma_{ti} - \Sigma_{s,ii}$;
- Σ_{ti}: total cross-section of group i;
- $\Sigma_{s,ij}$: scattering cross-section from group i to group j;
- $\Sigma_{ij} = -\Sigma_{s,ij}$;
- $D_i = \frac{1}{3\Sigma_{ti}}$: diffusion coefficient of group i;
- Σ_{fi}: fission cross-section of group i;
- ν_i: average number of neutrons of groupd i emitted per fission;
- χ_i: fission spectrum of group i

Parameters of the problem

- $\mu \in \mathcal{P}$ represents the **physical** properties of the core and its configuration.
- The spatial domain of calculation Ω is split into a structured grid that defines K regions. On each region Ω_k, μ^k represents the set of material parameters inside the domain Ω_k, so that $\mu = (\mu^1, \ldots, \mu^K) \in \mathcal{P}$.

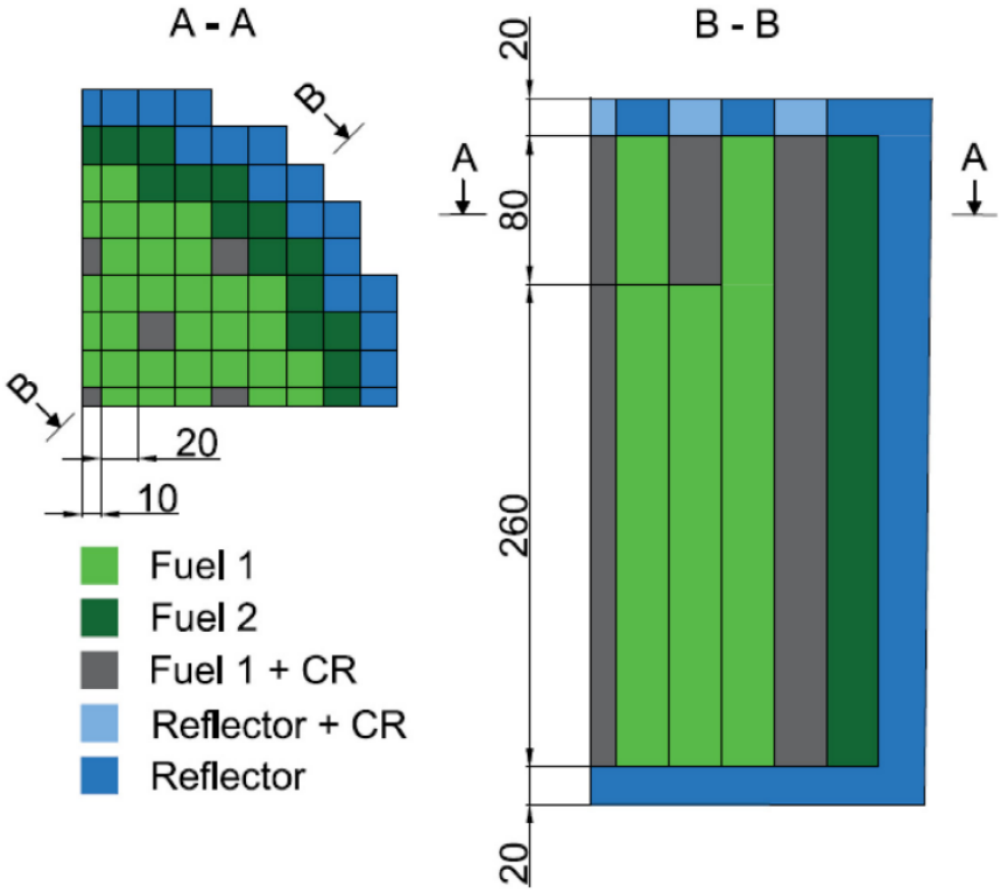


Figure: Cross-sectional view of the BSS-11 nuclear core reactor

Motivation of model-order reduction methods

- For a particular value of $\mu \in \mathcal{P}$, a numerical approximation of the solution u_μ is computed by some numerical scheme (for instance with a finite element code), the resolution of which may be very costly from a computational point of view.
- There exist a wide variety of contexts in which it is necessary to perform **parametric studies** of the problem at hand, i.e. to compute (a numerical approximation of) the solution u_μ for a very large number of values of the parameter vector μ as quickly as possible!

Examples:

- Design optimization
- Inverse problems
- Real-time control
- Uncertainty quantification

In such contexts, naive parametric studies using a standard finite element code may be extremely expensive from a computational point of view and time-consuming!

Principle of model-order reduction

Model-order reduction methods have been developped to circumvent this difficulty. The principle of these methods is the following:

- **Offline stage:** Compute u_μ with a standard numerical scheme (for instance finite elements) for a **small** number of well-chosen values of the parameter vector μ; this stage can be quite **expensive** from a computational point of view.
- Build another model, a **reduced model** from these few (expensive) computations in order to compute numerical approximations of u_μ for many other values of μ, but at a computational cost which is **much cheaper** than the initial (finite element) scheme.
- **Online stage:** Use the reduced model (instead of the original finite element code) in order to compute much faster u_μ for a large number of values of μ.

Outline

Reduced-basis method

There exists a huge number of model-order reduction techniques in the litterature.

In this talk: **Reduced Basis method** for accelerating the resolution of parametrized **generalized non-symmetric eigenvalue problems**, with a view to accelerating parametric studies for criticity calculations.

A few seminal references:

- Cohen, Dahmen, DeVore, Maday, Patera...
- *Reduced Basis Methods for Partial Differential Equations: An Introduction*, Alfio Quarteroni, Andrea Manzoni, Federico Negri
- *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, Jan S Hesthaven, Gianluigi Rozza, Benjamin Stamm

References on reduced basis techniques for symmetric eigenvalue problems:
[Fumagalli, Manzoni, Parolini, Verani, 2016], [Horger, Wohlmuth, Dickopf, 2017]

References on reduced basis techniques for neutronic applications:
[Sartori, Cammi, Luzzi, Rozza, 2016]

Two-Group Diffusion Equation (discrete formulation)

- Discretization of the spatial domain Ω with $P1$ crossed-triangular **finite elements** over a rectangle mesh
- The solution u_μ is approximated by an element $u_{\mu,h}$ belonging to a finite-dimensional subspace V_h of dimension N_h (number of DoFs): $V_h = \mathrm{Span}\{(\varphi_i)_{i=1,N_h}\}$

Weak formulation of the problem

$$\text{Find } \left(u_{\mu,h}, u^*_{\mu,h}, k_{\mu,h}\right) \in V_h \times V_h \times \mathbb{R}^*_+ \text{ such that}$$

$$\forall v_h \in V_h, \quad a_{\mu,h}(u_{\mu,h}, v_h) = \frac{1}{k_{\mu,h}} b_{\mu,h}(u_{\mu,h}, v_h).$$

Adjoint problem $\quad \forall v_h \in V_h, \quad a_{\mu,h}(v_h, u^*_{\mu,h}) = \frac{1}{k_{\mu,h}} b_{\mu,h}(v_h, u^*_{\mu,h}).$

Two-Group Diffusion Equation (matrix form)

$$u_{\mu,h} = \sum_{i=1}^{N_h} (U_{\mu,h})_i \varphi_i, \quad u^*_{\mu,h} = \sum_{i=1}^{N_h} (U^*_{\mu,h})_i \varphi_i \tag{2}$$

Matrix form of the problem

Find $\left(U_{\mu,h}, U^*_{\mu,h}, k_{\mu,h}\right) \in \mathbb{R}^{N_h} \times \mathbb{R}^{N_h} \times \mathbb{R}^*_+$ such that

$$A_{\mu,h} U_{\mu,h} = \frac{1}{k_{\mu,h}} B_{\mu,h} U_{\mu,h} \tag{3}$$

Adjoint problem

$$A^T_{\mu,h} U^*_{\mu,h} = \frac{1}{k_{\mu,h}} B^T_{\mu,h} U^*_{\mu,h}$$

- **Generalized eigenvalue problem**
- $A_{\mu,h} \in \mathbb{R}^{N_h \times N_h}$ is **non-symmetric**, invertible, with a coercive symmetric part
- $B_{\mu,h} \in \mathbb{R}^{N_h \times N_h}$ is **non-symmetric**, not invertible and positive

$\longrightarrow$ **High-fidelity** problem

Reduced basis method

- The resolution of the high-fidelity problem for a large number of values of the parameter vector $\mu \in \mathcal{P}$ may be very costly from a computational point of view because N_h is large!
- The principle of the reduced basis method is to approximate the solution $\left(u_{\mu,h}, u^*_{\mu,h}, k_{\mu,h}\right)$ by a Galerkin approximation associated to a linear subspace $V_N \subset V_h$ of dimension at most $2N$ with N much smaller than N_h.
- The **reduced space** V_N is chosen such that

$$V_N = \text{Vect}\left\{u_{\mu_1,h}, u^*_{\mu_1,h}, \cdots, u_{\mu_N,h}, u^*_{\mu_N,h}\right\},$$

  where $\mu_1, \cdots, \mu_N$ are N particular well-chosen values of the parameter vector μ.
- In the **offline stage**, the high-fidelity problem is only solved for this N values of the parameter vector.

Online stage of the Reduced Basis method

Galerkin approximation of the eigenvalue problem in the discretization space V_N

Weak formulation of the reduced problem

$$\text{Find } (u_{\mu,N}, u^*_{\mu,N}, k_{\mu,N}) \in V_N \times V_N \times \mathbb{R}^*_+ \text{ such that}$$

$$\forall v_N \in V_N, \quad a_{\mu,h}(u_{\mu,N}, v_N) = \frac{1}{k_{\mu,N}} b_{\mu,h}(u_{\mu,N}, v_N).$$

Adjoint problem $\quad \forall v_N \in V_N, \quad a_{\mu,h}(v_N, u^*_{\mu,N}) = \frac{1}{k_{\mu,N}} b_{\mu,h}(v_N, u^*_{\mu,N}).$

Online stage of the reduced basis method

- In the **online stage**, for each new value of $\mu \in \mathcal{P}$, an atmost $2N$-dimensional matrix eigenvalue problem is solved. When $N \ll N_h$, the resolution of the reduced problem is much cheaper from a computational point of view than the resolution of the original high-fidelity problem!
- **Reduced basis**: Let $n := \dim V_N$ and $(\theta_1, \cdots, \theta_n)$ be an orthonormal basis of V_N. Denoting by

$$\Theta_N := (\theta_1 | \cdots | \theta_n) \in \mathbb{R}^{N_h \times n},$$

We define the $n \times n$ reduced matrices:

$$\begin{cases} A_{\mu,N} = \Theta_N^T A_{\mu,h} \Theta_N \\ B_{\mu,N} = \Theta_N^T B_{\mu,h} \Theta_N \end{cases}.$$

Approximate solutions given by the reduced basis method

Reduced problem

Find $\left(c_{\mu,N}, c^*_{\mu,N}, k_{\mu,N}\right) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^*_+$ such that

$$A_{\mu,N} c_{\mu,N} = \frac{1}{k_{\mu,N}} B_{\mu,N} c_{\mu,N} \quad \text{and} \quad U_{\mu,N} = \Theta_N c_{\mu,N}$$

$$A^T_{\mu,N} c^*_{\mu,N} = \frac{1}{k_{\mu,N}} B^T_{\mu,N} c^*_{\mu,N} \quad \text{and} \quad U^*_{\mu,N} = \Theta_N c^*_{\mu,N}$$

$$u_{\mu,N} := \sum_{i=1}^{N_h} \left(U_{\mu,N}\right)_i \varphi_i, \quad u^*_{\mu,N} := \sum_{i=1}^{N_h} \left(U^*_{\mu,N}\right)_i \varphi_i.$$

Efficient computation of the reduced matrices

Assumption: Separability of the data of the problem

$$A_{\mu,h} = \sum_{p=1}^{P} g_p(\mu) A_{p,h}$$

Offline phase: Compute for all $1 \leq p \leq P$

$$A_{p,N} = \Theta_N^T A_{p,h} \Theta_N$$

Online phase: For $\mu \in \mathcal{P}$, compute $A_{\mu,N}$ as

$$A_{\mu,N} = \sum_{p=1}^{P} g_p(\mu) A_{p,N}$$

Complexity: $\mathcal{O}(Pn^2)$ (no dependence in N_h)

Similar computation for $B_{\mu,N}$.

How to build V_N ?

This is usually done via an iterative algorithm called a **greedy algorithm**.

Need to choose a finite subset $\mathcal{P}_{\text{train}} \subset \mathcal{P}$, called **training set**.

Naive Greedy algorithm

- Choose randomly $\mu_1 \in \mathcal{P}_{\text{train}}$.

$$V_1 = \text{Vect}\left\{u_{\mu_1,h}, u^*_{\mu_1,h}\right\}$$

- **Iteration** N**:** Choose $\mu_N \in \mathcal{P}_{\text{train}}$ such that

$$\mu_N \in \underset{\mu \in \mathcal{P}_{\text{train}}}{\text{argmax}} \, |k_{\mu,h} - k_{\mu,N-1}|$$

$$V_N = \text{Vect}\left\{u_{\mu_1,h}, u^*_{\mu_1,h}, \cdots, u_{\mu_N,h}, u^*_{\mu_N,h}\right\}$$

A naive version of the Greedy algorithm requires to evaluate $k_{\mu,h}$, for all $\mu \in \Lambda_{\text{train}}$
$\rightarrow$ too expensive...

Practical algorithm:
Replace $e^k_{N-1}(\mu) := |k_{\mu,h} - k_{\mu,N-1}|$ by an easy-to-compute **a posteriori error estimator** $\Delta^k_{N-1}(\mu)$.

Error on the eigenvalue

- **Residuals**:

$$R_{\mu,N} = (B_{\mu,h} - k_{\mu,N} A_{\mu,h})\, u_{\mu,N}$$
$$R^*_{\mu,N} = \left(B^T_{\mu,h} - k_{\mu,N} A^T_{\mu,h}\right) u^*_{\mu,N} \tag{4}$$

Proposition. *A posteriori* error estimator

There exists a positive constant $C^k(\mu) > 0$ (called the **prefactor**) such that for all $\mu \in \mathcal{P}$,

$$e^k_N(\mu) = |k_{\mu,h} - k_{\mu,N}| \leqslant C^k(\mu) \frac{\|R_{\mu,N}\|_* \|R^*_{\mu,N}\|_*}{\langle c^*_{\mu,N}, A_{\mu,N} c_{\mu,N}\rangle} = C^k(\mu) \eta^k_N(\mu) \tag{5}$$

with $\eta^k_N(\mu) := \dfrac{\|R_{\mu,N}\|_* \|R^*_{\mu,N}\|_*}{\langle c^*_{\mu,N}, A_{\mu,N} c_{\mu,N}\rangle}$ and $\|\cdot\|_*$ a norm on $\mathbb{R}^{N_h}$.

Error on the eigenvectors

- **Residuals**:

$$R_{\mu,N} = (B_{\mu,h} - k_{\mu,N} A_{\mu,h})\, u_{\mu,N}$$
$$R^*_{\mu,N} = \left(B^T_{\mu,h} - k_{\mu,N} A^T_{\mu,h}\right) u^*_{\mu,N} \tag{6}$$

Proposition. *A posteriori* error estimator

There exist positive constants $C^u(\mu), C^{u_*}(\mu) > 0$ (called the **prefactors**) such that for all $\mu \in \mathcal{P}$,

$$e^u_N(\mu) = \|u_{\mu,h} - u_{\mu,N}\| \leqslant C^u(\mu)\|R_{\mu,N}\|_* = C^u(\mu)\eta^u_N(\mu) \tag{7}$$

and

$$e^{u^*}_N(\mu) = \left\|u^*_{\mu,h} - u^*_{\mu,N}\right\| \leqslant C^{u^*}(\mu)\|R^*_{\mu,N}\|_* = C^{u^*}(\mu)\eta^{u^*}_N(\mu) \tag{8}$$

where $\|\cdot\|$ and $\|\cdot\|_*$ are dual norms of one another.

Example: If $\|\cdot\|$ is a discrete $H^1_0(\Omega)$ norm, $\|\cdot\|_*$ is a then discrete $H^{-1}(\Omega)$ norm.

How to build V_N ?

- **Practical *a posteriori* error estimator**:

$$\Delta_N^k(\mu) = \overline{C}_N^k \frac{\|R_{\mu,N}\| \|R^*_{\mu,N}\|}{\langle c^*_{\mu,N}, A_{\mu,N} c_{\mu,N} \rangle} = \overline{C}_N^k \eta_N^k(\mu)$$

  where $\overline{C}_N^k$ is a heuristic estimation of the prefactor $C^k(\mu)$
- $\Delta_N^k(\mu)$ can be efficiently computed with complexity $\mathcal{O}(n^2)$ if the data of the problem is separated.

How to build V_N ?

Actual Greedy algorithm

- Choose randomly $\mu_1 \in \mathcal{P}_{\text{train}}$.

$$V_1 = \text{Vect}\left\{u_{\mu_1,h}, u^*_{\mu_1,h}\right\}$$

- **Iteration** N**:** Choose $\mu_N \in \mathcal{P}_{\text{train}}$ such that

$$\mu_N \in \underset{\mu \in \mathcal{P}_{\text{train}}}{\text{argmax}}\ \Delta^k_{N-1}(\mu).$$

$$V_N = \text{Vect}\left\{u_{\mu_1,h}, u^*_{\mu_1,h}, \cdots, u_{\mu_N,h}, u^*_{\mu_N,h}\right\}$$

How to build $\overline{C}_N^k$? Heuristic procedure

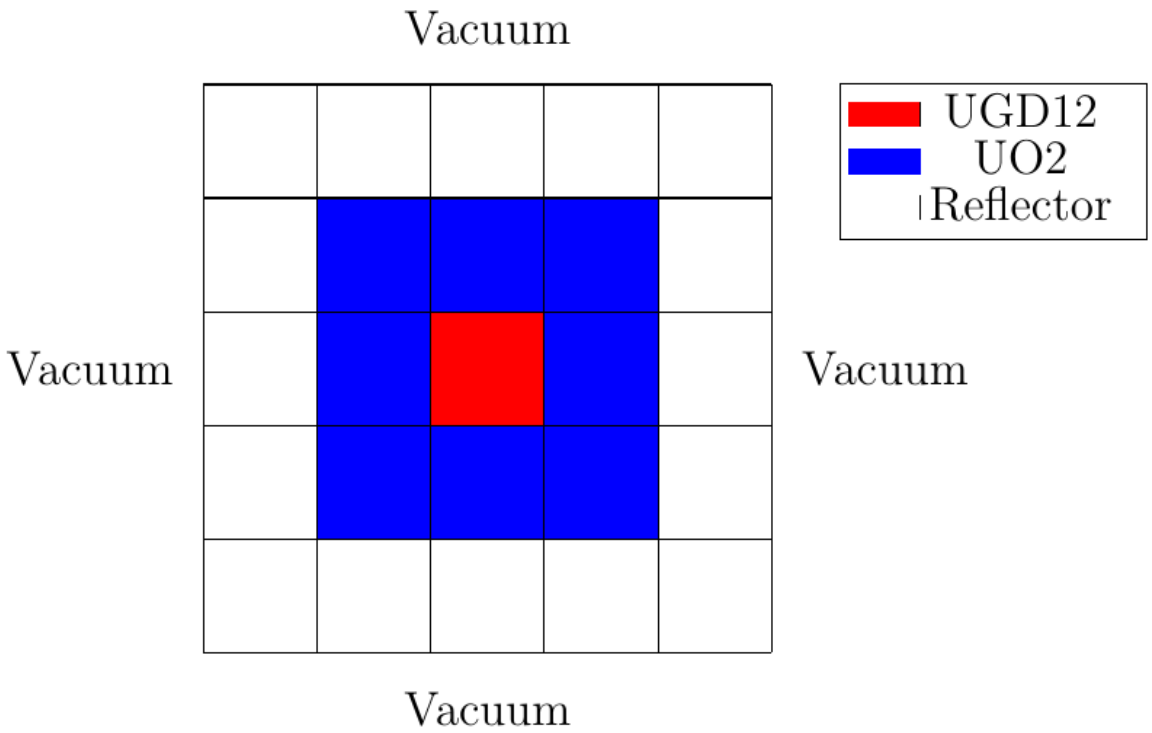
Need to choose a finite subset $\mathcal{P}_{\text{pref}} \subset \mathcal{P}$, called **prefactor set** such that $\mathcal{P}_{\text{pref}} \cap \mathcal{P}_{\text{train}} = \emptyset$.

- Solve the high-fidelity problem for all $\mu \in \mathcal{P}_{\text{pref}}$.
- Compute for all $\mu \in \mathcal{P}_{\text{pref}}$, $\mathcal{E}_N^k(\mu) := \frac{|k_{\mu,h} - k_{\mu,N}|}{\eta_N^k(\mu)}$.
- Define

$$\overline{C}_N^k := \max_{\mu \in \mathcal{P}_{\text{pref}}} \mathcal{E}_N^k(\mu).$$

Outline

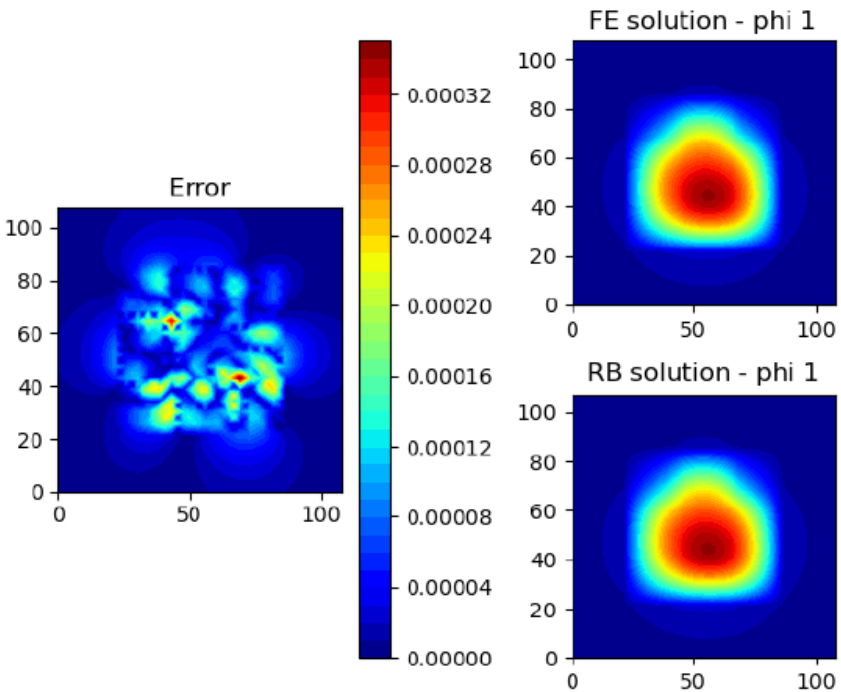
First toy test case: the MiniCore problem

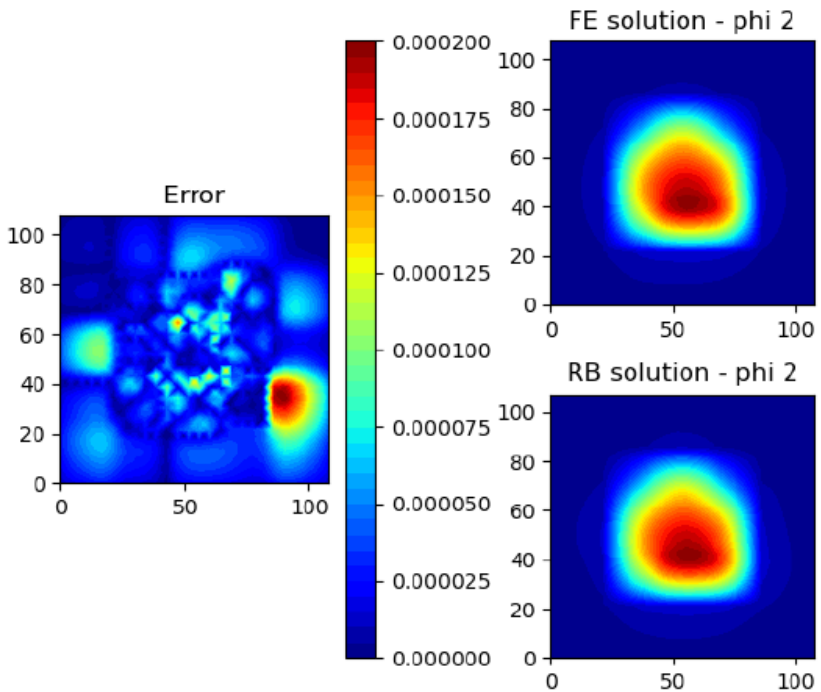


- 25 spatial regions
- $L = 107.52\ cm$
- UGD12: mix of uranium dioxyde and Galinium oxyde
- UO2: uranium dioxyde
- BC: $u_\mu(x) = 0, \quad x \in \partial\Omega$
- $N_h = 2602$ DoFs per group
- Training set of parameters $\mathcal{P}_{\text{train}}$ of cardinality 1000 generated **randomly**

High-fidelity and reduced solver

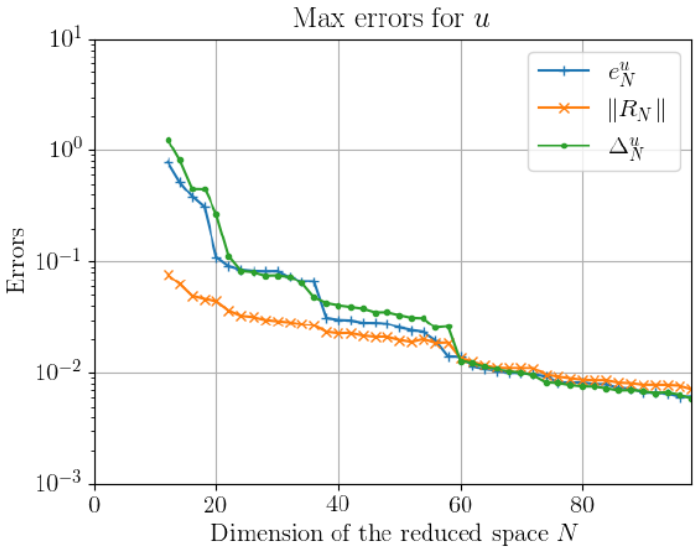
Reduced-order model obtained with $N = 100$

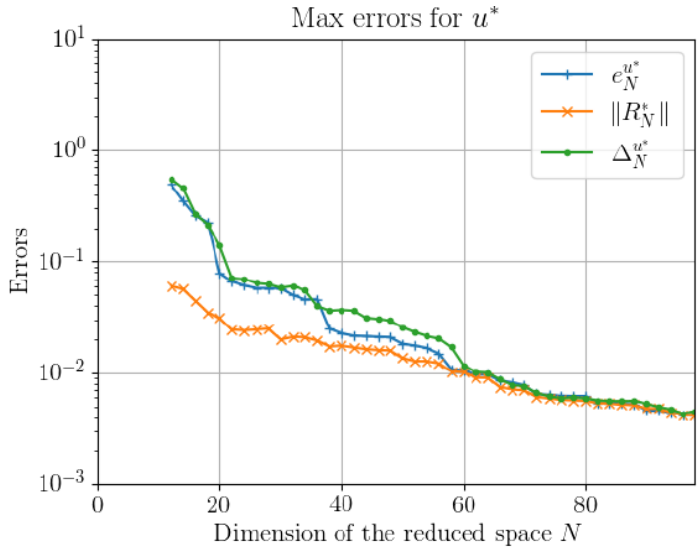


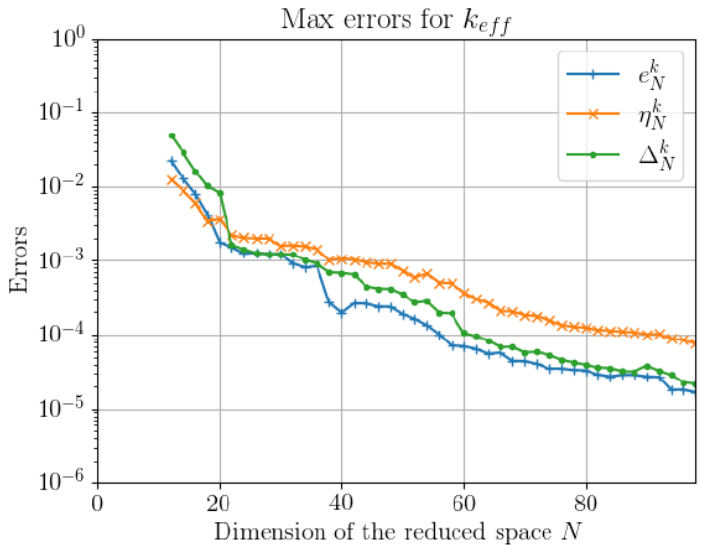


Convergence of the reduced basis : mean relative errors over $\mathcal{P}_{\text{test}}$

- $\mathcal{P}_{\text{test}} \subset \mathcal{P}$ with cardinality 50 (test set)
- $\mathcal{P}_{\text{pref}} \subset \mathcal{P}$ with cardinality 10 (prefactor set)







Parametric variability of the prefactor

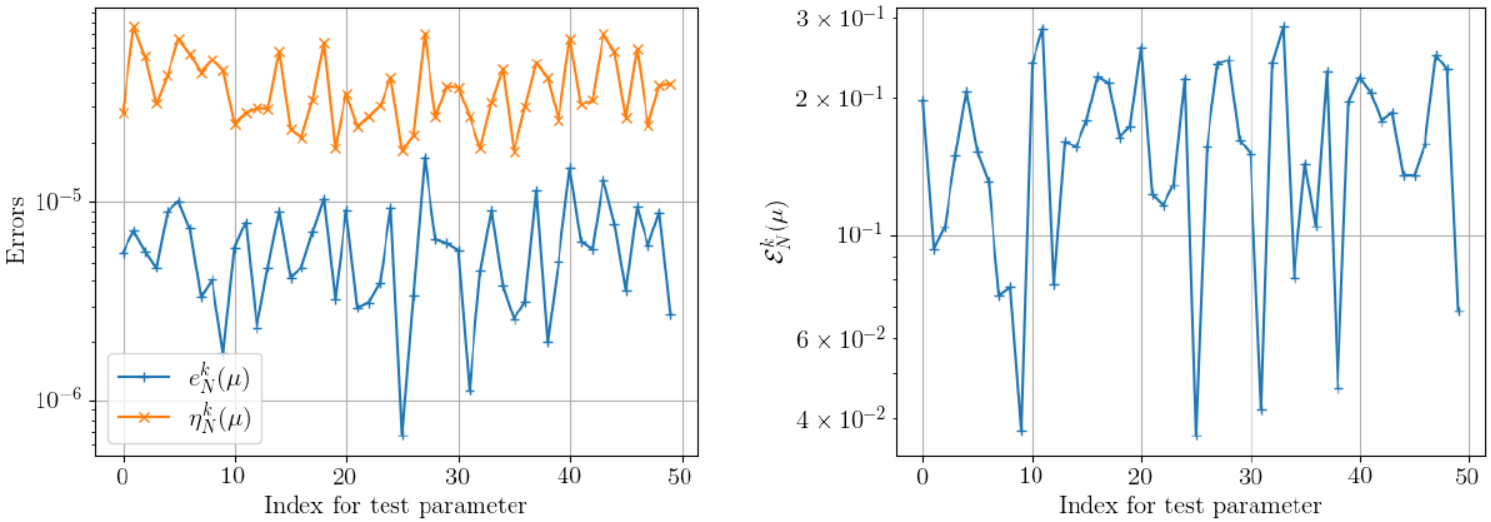


Figure: Parametric variability of the prefactor

Gain in computational time

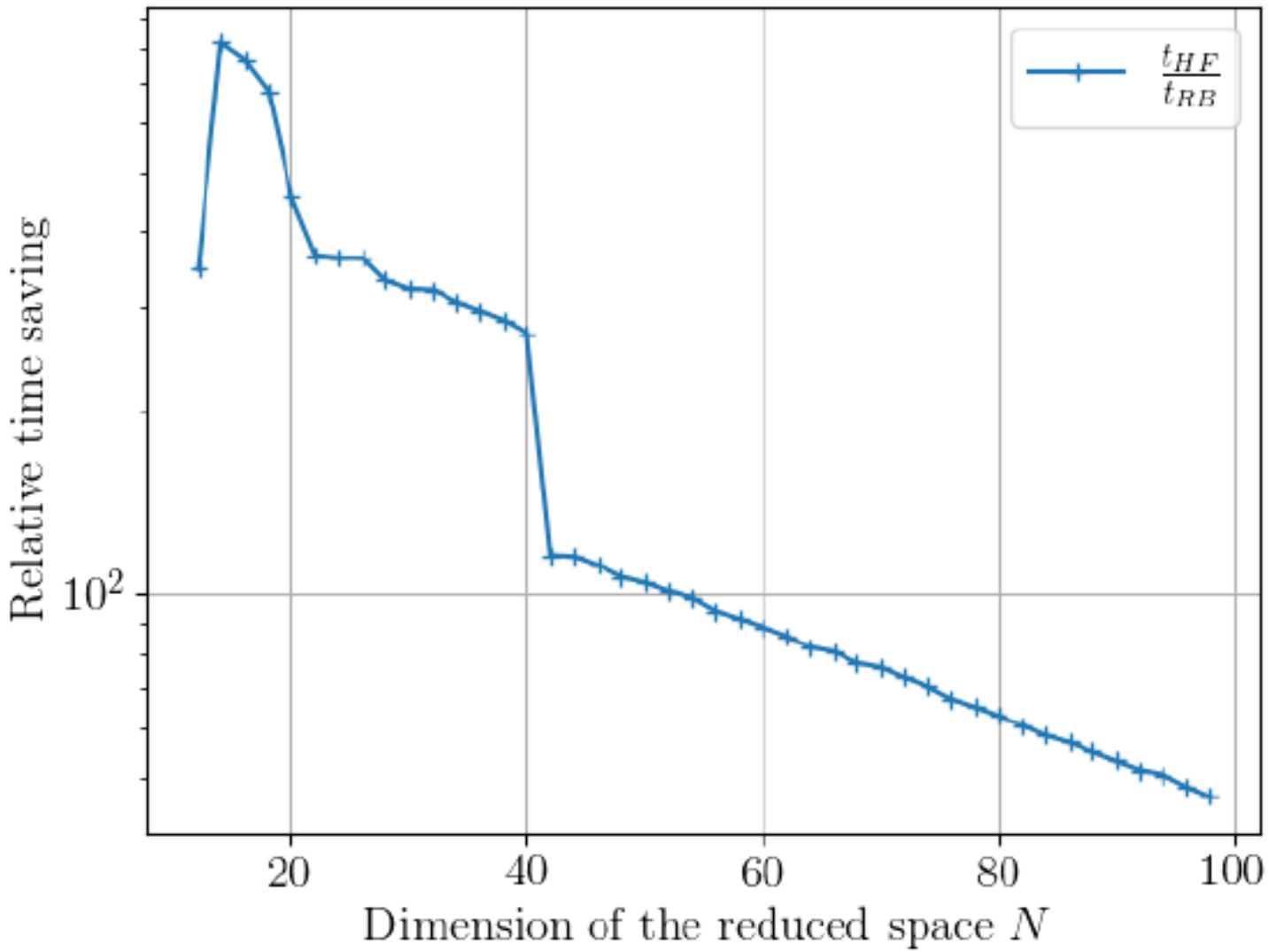


Figure: Relative time saving of the reduced solver

3D test case in APOLLO3 code (MINARET solver)

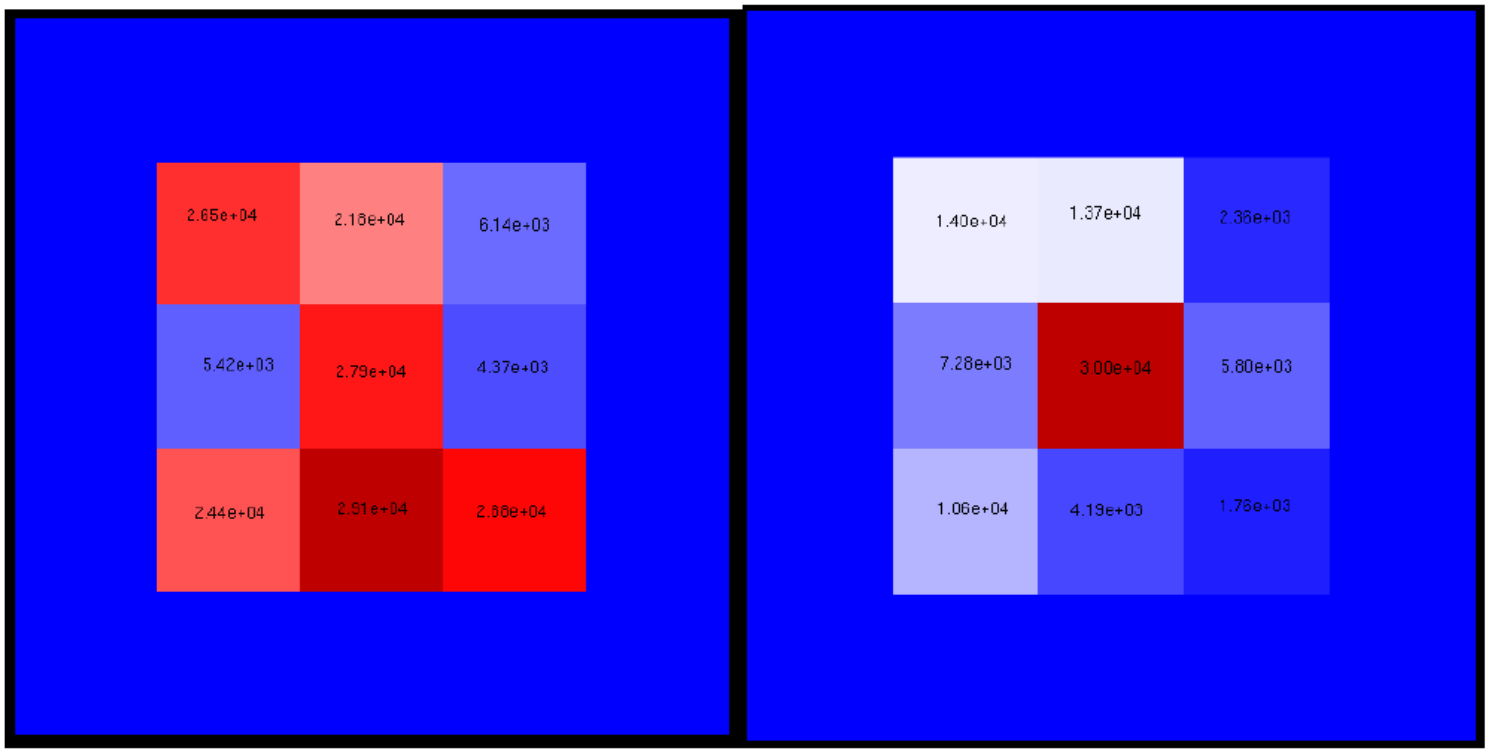
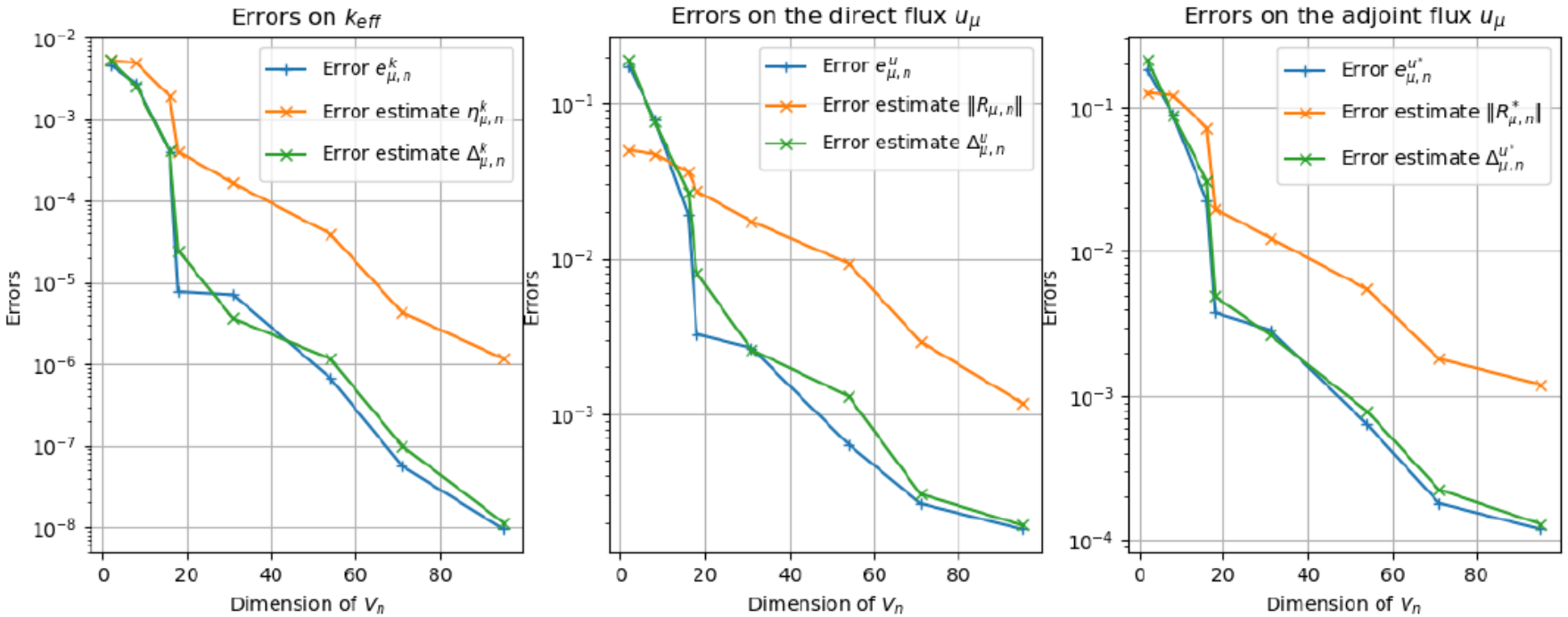


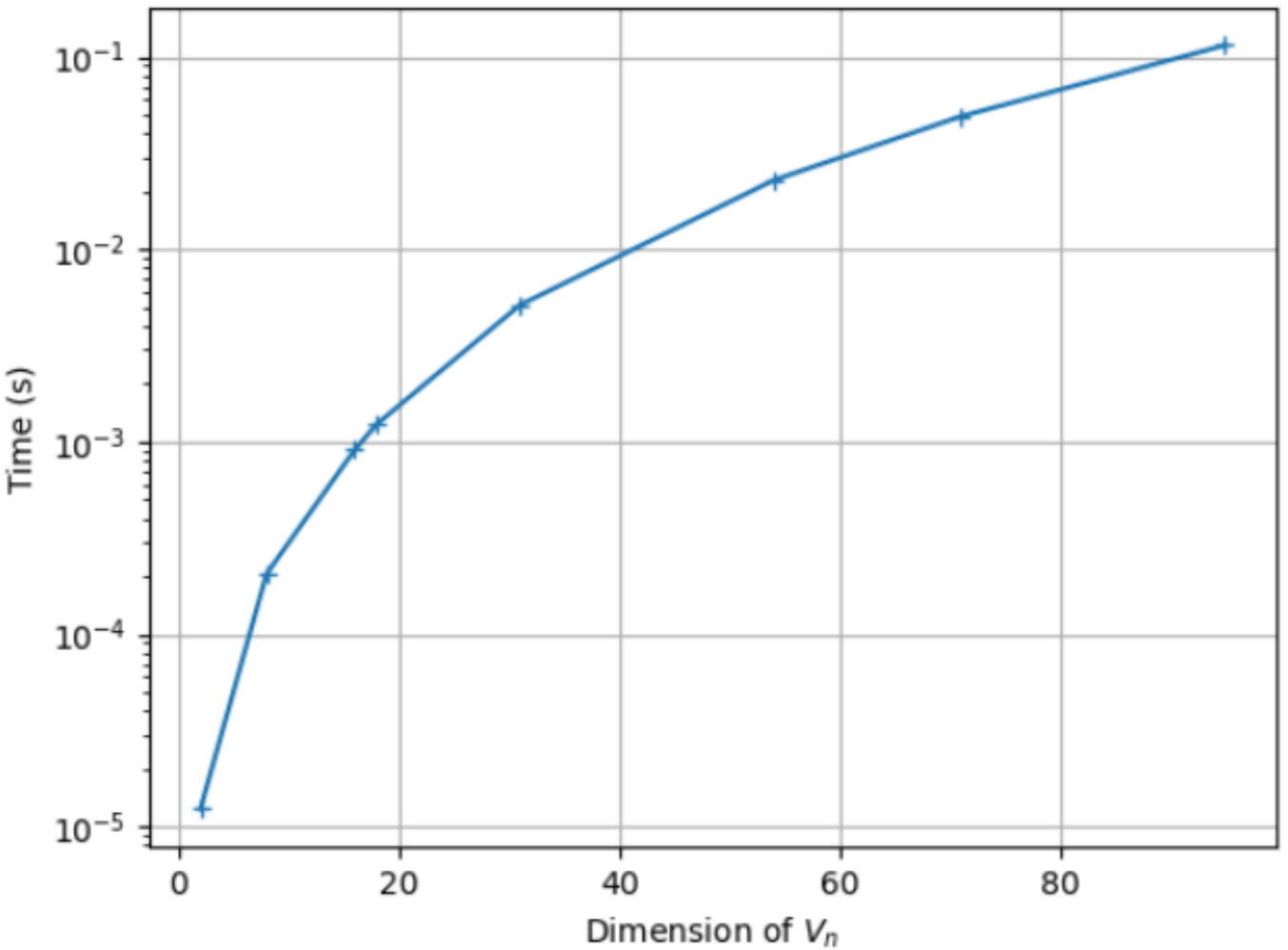
Figure: Cross-sectional views of the 3D core

- 9 spatial regions
- BC: $u_\mu(x) = 0, \quad x \in \partial\Omega$
- $N_h = 108800$ DoFs per group
- Training set of parameters $\mathcal{P}_{\text{train}}$ of cardinality 100 generated **randomly**
- Prefactor set $\mathcal{P}_{\text{pref}}$ of cardinality 5
- Test set $\mathcal{P}_{\text{test}}$ of cardinality 10

Convergence of the reduced basis approximation



Computational runtime of the reduced-order model



Conclusions and perspectives

Conclusions:

- Efficient reduced-order model for criticity calculations in neutronics using the reduced basis method
- Very encouraging results obtained on two-group diffusion models with the APOLLO3 code
- Heuristic procedure to estimate the value of the prefactor which yields accurate estimations of the true error

Perpsectives:

- More complex parametric variability in the 3D APOLLO test case
- Specific method to explore the parameter space: choice of $\mathcal{P}_{\text{train}}$, $\mathcal{P}_{\text{pref}}$?
- Rigorous justification of the heuristic procedure used to estimate the prefactor
- Efficient implementation of the reduced order model in the APOLLO code
- Application of the reduced basis method for the reduction of transport models (more complex than two-group diffusion models, like Boltzmann model)

Thank you for your attention!