Universität Münster

Adaptive Model Hierarchies for Multi-Query Scenarios

MATHMOD 2025 – Minisymposium on “Recent Advances in Model Order Reduction and Data-driven Modelling”

Hendrik Kleikamp, Mario Ohlberger (University of Münster)

February 20, 2025

living.knowledge

MM Mathematics Münster Cluster of Excellence

Motivation and general construction

Motivation for model hierarchies

Full order model, reduced model, machine learning surrogate, ...

- Availability of different models with different advantages and disadvantages, such as
  - Full order model
  - Projection-based reduced model
  - Machine learning surrogate
  - ...

hendrik.kleikamp@uni-muenster.de

Motivation for model hierarchies

Full order model, reduced model, machine learning surrogate, ...

- Availability of different models with different advantages and disadvantages, such as
  - Full order model
  - Projection-based reduced model
  - Machine learning surrogate
  - ...
- Typical strategy: Train a suitable surrogate offline and use it online.

hendrik.kleikamp@uni-muenster.de

Motivation for model hierarchies
Full order model, reduced model, machine learning surrogate, ...

- Availability of different models with different advantages and disadvantages, such as
  - Full order model
  - Projection-based reduced model
  - Machine learning surrogate
  - ...

- Typical strategy: Train a suitable surrogate offline and use it online.

- Idea: Train surrogates adaptively and try to leverage their advantages.

Motivation for model hierarchies

Full order model, reduced model, machine learning surrogate, ...

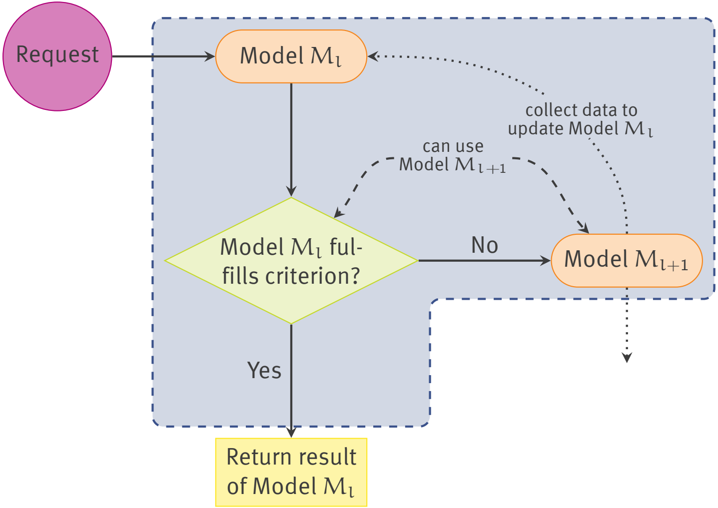
- Availability of different models with different advantages and disadvantages, such as
  - Full order model
  - Projection-based reduced model
  - Machine learning surrogate
  - ...
- Typical strategy: Train a suitable surrogate offline and use it online.
- Idea: Train surrogates adaptively and try to leverage their advantages.

$\Longrightarrow$ Combine all available models in an adaptive and certified hierarchy!

hendrik.kleikamp@uni-muenster.de

Description of the main building block

Multi-fidelity assumptions



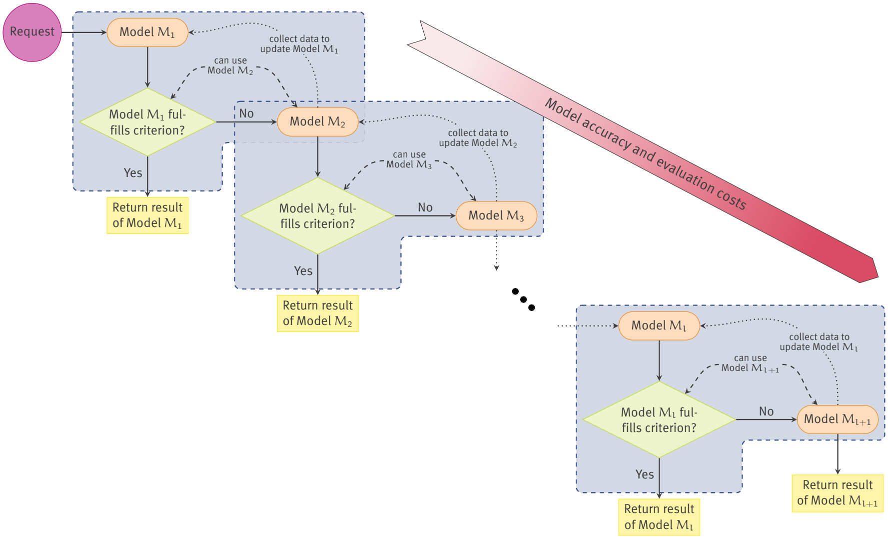
Assumptions:

- Model M_l can be solved faster than Model M_{l+1}
- Model M_{l+1} is more accurate than Model M_l
- Model M_l can be improved by means of information from Model M_{l+1}

hendrik.kleikamp@uni-muenster.de

General definition of an adaptive model hierarchy

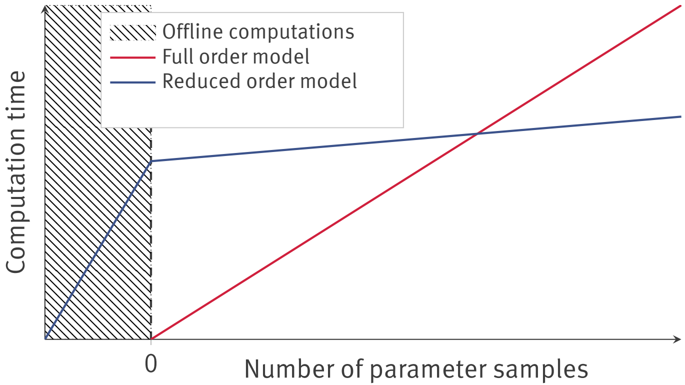
Combination of several components in multiple stages



hendrik.kleikamp@uni-muenster.de

Adaptive model hierarchy for parametrized problems

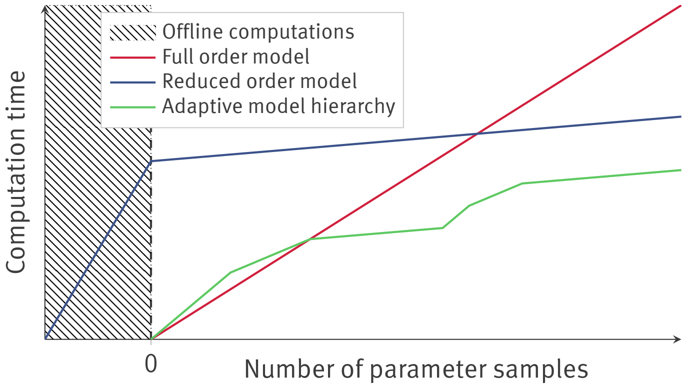
Adaptivity vs. offline-online decomposition



hendrik.kleikamp@uni-muenster.de

Adaptive model hierarchy for parametrized problems

Adaptivity vs. offline-online decomposition



Examples and numerical results

Example 1 – Parametrized parabolic PDEs

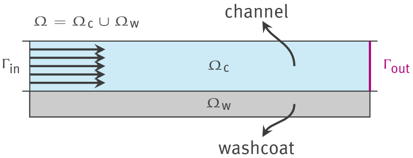
Problem setting – PDE-constrained optimization

Advection-diffusion-reaction equation with parameter $\mu = (\mathrm{Da}, \mathrm{Pe}) \in \mathcal{P} = [0.01, 10] \times [9, 11]$:

$$\partial_t u_\mu - \nabla \cdot (\kappa \nabla u_\mu) + \mathrm{Pe}\, \nabla \cdot (\vec{b}\, u_\mu) + \mathrm{Da}\, \chi_{\Omega_w} u_\mu = 0 \qquad \text{in } (0,T) \times \Omega,$$

$$(\kappa \nabla u_\mu) \cdot n_{\partial\Omega} = 0 \qquad \text{on } \Gamma_{\text{out}},$$

$$\Phi(\mu)(t) = \fint_{\Gamma_{\text{out}}} u_\mu(t)\, ds.$$



hendrik.kleikamp@uni-muenster.de

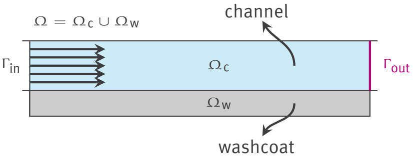
Problem setting – PDE-constrained optimization

Advection-diffusion-reaction equation with parameter $\mu = (\mathrm{Da}, \mathrm{Pe}) \in \mathcal{P} = [0.01, 10] \times [9, 11]$:

$$\partial_t u_\mu - \nabla \cdot (\kappa \nabla u_\mu) + \mathrm{Pe}\, \nabla \cdot (\vec{b}\, u_\mu) + \mathrm{Da}\, \chi_{\Omega_w} u_\mu = 0 \qquad \text{in } (0,T) \times \Omega,$$

$$(\kappa \nabla u_\mu) \cdot n_{\partial\Omega} = 0 \qquad \text{on } \Gamma_{\mathrm{out}},$$

$$\Phi(\mu)(t) = \int_{\Gamma_{\mathrm{out}}} u_\mu(t)\, ds.$$



PDE-constrained optimization problem:

$$\min_{\mu \in \mathcal{P}} \mathcal{J}(\mu) := \|\hat{\Phi} - \Phi_{\mathrm{adapt}}(\mu)\|_{L^\infty([0,T])}.$$

Here, $\hat{\Phi} = \Phi(\hat{\mu})$ is a desired break-through curve with $\hat{\mu} = (5.005, 10)$ and $\Phi_{\mathrm{adapt}}(\mu)$ is the output of the adaptive model hierarchy.

M_3: Full order model (FOM)

+ arbitrarily accurate solutions (serves as the reference for error estimators)
– very slow for fine discretizations or many time steps

hendrik.kleikamp@uni-muenster.de

Example 1 – Parametrized parabolic PDEs
Three-stage hierarchy

M_3: Full order model (FOM)

- + arbitrarily accurate solutions (serves as the reference for error estimators)
- – very slow for fine discretizations or many time steps

M_2: Reduced basis ROM (RB-ROM)

- ▶ reduced basis model based on Galerkin projection
- ▶ construction of reduced basis using HaPOD
- + faster than FOM
- + residual-based a posteriori error estimator
- – slow iterative time stepping
- – requires solving dense linear system in each time step

hendrik.kleikamp@uni-muenster.de

Example 1 – Parametrized parabolic PDEs

Three-stage hierarchy

$\mathcal{M}_3$: Full order model (FOM)

+ arbitrarily accurate solutions (serves as the reference for error estimators)
– very slow for fine discretizations or many time steps

$\mathcal{M}_2$: Reduced basis ROM (RB-ROM)

- reduced basis model based on Galerkin projection
- construction of reduced basis using HaPOD

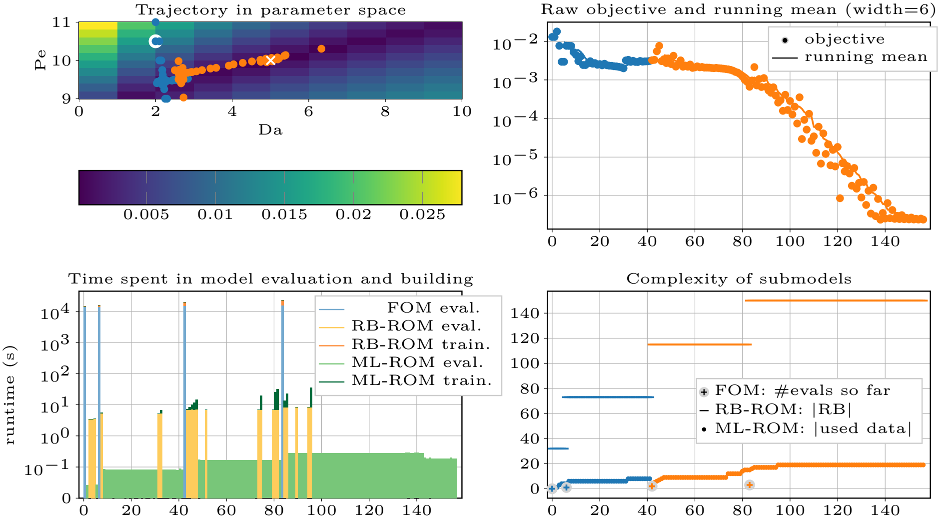
+ faster than FOM
+ residual-based a posteriori error estimator
– slow iterative time stepping
– requires solving dense linear system in each time step

$\mathcal{M}_1$: Machine learning (ML-ROM)

- approximates reduced coefficients of the RB-ROM
- based on the same reduced basis as the RB-ROM

+ faster than RB-ROM
+ parallel evaluation for many time instances possible
+ allows to reuse a posteriori error estimator of RB-ROM
– requires training of machine learning algorithm

Numerical results



Examples 2 and 3

Optimal control of (parametric) systems

Examples 2 and 3

Optimal control of (parametric) systems

Please come to my poster for examples 2 and 3!

hendrik.kleikamp@uni-muenster.de

Universität Münster

Thank you for your attention!

If you would like to discuss in more detail, come to my poster!

hendrik.kleikamp@uni-muenster.de

Example 2 – Parametrized optimal control problems

Problem setting – Monte Carlo simulations

$$\min_{u \in L^2([0,T];U)} \quad \frac{1}{2}\left[\underbrace{\| y_\mu(T) - y_\mu^T \|_Y^2}_{\substack{\text{deviation from target}\\ \text{at final time}}} + \underbrace{\int_0^T \langle u(t), R(t)u(t)\rangle_{U\times U'}\,dt}_{\text{control energy}}\right],$$

hendrik.kleikamp@uni-muenster.de

$$\min_{u\in L^2([0,T];U)} \quad \frac{1}{2}\Bigg[\underbrace{\left\| y_\mu(T) - y_\mu^T \right\|_Y^2}_{\text{deviation from target at final time}} + \underbrace{\int_0^T \langle u(t), R(t)u(t)\rangle_{U\times U'}\, dt}_{\text{control energy}}\Bigg],$$

$$\text{such that} \quad \begin{cases} E\frac{d}{dt}x_\mu(t) = A(\mu;t)x_\mu(t) + B(\mu;t)u(t) & \text{for } t\in[0,T], \\ y_\mu(t) = Cx_\mu(t) & \text{for } t\in[0,T], \\ x_\mu(0) = x_\mu^0, & \text{(initial data)} \end{cases}$$

Target output: $y_\mu^T = Cx_\mu^T$

Universität Münster MM

Theorem (K./Renelt'24)

The optimal state x_μ^**, control* u_μ^* *and adjoint* φ_μ^* *can be characterized as solution to the system:*

$$E\frac{d}{dt}x_\mu(t) = A(\mu;t)x_\mu(t) + B(\mu;t)u_\mu(t), \qquad \textit{(state equation)}$$

$$u_\mu(t) = -R(t)^{-1}B(\mu;t)^*\varphi_\mu(t), \qquad \textit{(control equation)}$$

$$-E_{ad}\frac{d}{dt}\varphi_\mu(t) = A(\mu;t)^*\varphi_\mu(t), \qquad \textit{(adjoint equation)}$$

for almost all $t \in [0,T]$ *with initial condition* $x_\mu(0) = x_\mu^0$ *and terminal condition*

$$\mathcal{R}_X E^*\varphi_\mu(T) = M\left(x_\mu(T) - x_\mu^T\right),$$

where $\mathcal{R}_X \colon X \to X'$ *is the Riesz map,* $E_{ad} \coloneqq \mathcal{R}_X E^* \mathcal{R}_X^{-1}$ *and* $M \coloneqq C^*\mathcal{R}_Y C$.

Example 2 – Parametrized optimal control problems
Linear system for optimal final time adjoints

Lemma (K./Renelt'24)

The optimal final time adjoint $\varphi_\mu^*(T) \in X$ *solves the linear system*

$$\boxed{S(\mu)\varphi_\mu^*(T) = g(\mu)}$$

where

- $S(\mu)$ *involves solving high-dimensional parametric systems,*
- $g(\mu)$ *is a parameter-dependent right hand side.*

Linear system for optimal final time adjoints

Lemma (K./Renelt'24)

The optimal final time adjoint $\varphi_\mu^*(T) \in X$ solves the linear system

$$\boxed{S(\mu)\varphi_\mu^*(T) = g(\mu)}$$

where

- $S(\mu)$ *involves solving high-dimensional parametric systems,*
- $g(\mu)$ *is a parameter-dependent right hand side.*

$\Longrightarrow$ Costly to solve!

Example 2 – Parametrized optimal control problems

Linear system for optimal final time adjoints

Lemma (K./Renelt'24)

The optimal final time adjoint $\varphi_\mu^*(T) \in X$ *solves the linear system*

$$S(\mu)\varphi_\mu^*(T) = g(\mu) \qquad \Longleftrightarrow \qquad \varphi_\mu^*(T) = \arg\min_{p \in X} \| g(\mu) - S(\mu)p \|_X^2$$

where

- $S(\mu)$ *involves solving high-dimensional parametric systems,*
- $g(\mu)$ *is a parameter-dependent right hand side.*

$\Longrightarrow$ Costly to solve!

hendrik.kleikamp@uni-muenster.de

Example 2 – Parametrized optimal control problems

Four-stage hierarchy

M_4: FOM

Optimal final time adjoint

$$\varphi_\mu^*(T) = \operatorname*{arg\,min}_{p \in X} \| g(\mu) - S(\mu)p \|_X^2$$

hendrik.kleikamp@uni-muenster.de

Example 2 – Parametrized optimal control problems

Four-stage hierarchy

M_4: FOM

Optimal final time adjoint

$$\varphi_\mu^*(T) = \operatorname*{arg\,min}_{p \in X} \| g(\mu) - S(\mu)p \|_X^2$$

Primal and adjoint systems

$$E \frac{d}{dt} x_\mu(t) = A(\mu; t) x_\mu(t) + B(\mu; t) u_\mu(t)$$

$$-E_{ad} \frac{d}{dt} \varphi_\mu(t) = A(\mu; t)^* \varphi_\mu(t)$$

Four-stage hierarchy

[Lazar/Zuazua'16, K./Lazar/Molinari'24]

M_4: FOM

Optimal final time adjoint

$$\varphi_\mu^*(T) = \arg\min_{p \in X} \| g(\mu) - S(\mu)p \|_X^2$$

→

M_3: RB-ROM

Least-squares on reduced subspace

$$\tilde{\varphi}_\mu^N = \arg\min_{p \in X^N} \| g(\mu) - S(\mu)p \|_X^2$$

Primal and adjoint systems

$$E \frac{d}{dt} x_\mu(t) = A(\mu; t) x_\mu(t) + B(\mu; t) u_\mu(t)$$

$$-E_{ad} \frac{d}{dt} \varphi_\mu(t) = A(\mu; t)^* \varphi_\mu(t)$$

Example 2 – Parametrized optimal control problems
Four-stage hierarchy

[Lazar/Zuazua'16, K./Lazar/Molinari'24]

M_4: FOM

Optimal final time adjoint

$$\varphi_\mu^*(T) = \arg\min_{p \in X} \| g(\mu) - S(\mu)p \|_X^2$$

→

M_3: RB-ROM

Least-squares on reduced subspace

$$\tilde{\varphi}_\mu^N = \arg\min_{p \in X^N} \| g(\mu) - S(\mu)p \|_X^2$$

[Haasdonk/Ohlberger'11]

Primal and adjoint systems

$$E \frac{d}{dt} x_\mu(t) = A(\mu; t) x_\mu(t) + B(\mu; t) u_\mu(t)$$

$$-E_{ad} \frac{d}{dt} \varphi_\mu(t) = A(\mu; t)^* \varphi_\mu(t)$$

→

Reduced primal and adjoint systems

$$\hat{E}_{pr} \frac{d}{dt} \hat{x}_\mu(t) = \hat{A}_{pr}(\mu; t) \hat{x}_\mu(t) + \hat{B}_{pr}(\mu; t) u_\mu(t)$$

$$-\hat{E}_{ad} \frac{d}{dt} \hat{\varphi}_\mu(t) = \hat{A}_{ad}(\mu; t)^* \hat{\varphi}_\mu(t)$$

Example 2 – Parametrized optimal control problems

Four-stage hierarchy

M_4: FOM

Optimal final time adjoint

$$\varphi_\mu^*(T) = \underset{p \in X}{\arg\min} \| g(\mu) - S(\mu)p \|_X^2$$

[Lazar/Zuazua'16, K./Lazar/Molinari'24]

M_3: RB-ROM

Least-squares on reduced subspace

$$\tilde{\varphi}_\mu^N = \underset{p \in X^N}{\arg\min} \| g(\mu) - S(\mu)p \|_X^2$$

[Fabrini/Iapichino/Volkwein'18, K./Renelt'24]

M_2: F-ROM

Combination of RB-ROM and system reduction

$$\hat{\varphi}_\mu^{N,\text{red}} = \underset{p \in X^N}{\arg\min} \left\| \hat{g}(\mu) - \hat{S}(\mu)p \right\|_X^2$$

Primal and adjoint systems

$$E \frac{d}{dt} x_\mu(t) = A(\mu; t) x_\mu(t) + B(\mu; t) u_\mu(t)$$

$$-E_{ad} \frac{d}{dt} \varphi_\mu(t) = A(\mu; t)^* \varphi_\mu(t)$$

[Haasdonk/Ohlberger'11]

Reduced primal and adjoint systems

$$\hat{E}_{pr} \frac{d}{dt} \hat{x}_\mu(t) = \hat{A}_{pr}(\mu; t) \hat{x}_\mu(t) + \hat{B}_{pr}(\mu; t) u_\mu(t)$$

$$-\hat{E}_{ad} \frac{d}{dt} \hat{\varphi}_\mu(t) = \hat{A}_{ad}(\mu; t)^* \hat{\varphi}_\mu(t)$$

hendrik.kleikamp@uni-muenster.de

Example 2 – Parametrized optimal control problems

Four-stage hierarchy

M_4: FOM

Optimal final time adjoint

$$\varphi_\mu^*(T) = \underset{p \in X}{\arg\min} \| g(\mu) - S(\mu)p \|_X^2$$

[Lazar/Zuazua'16, K./Lazar/Molinari'24]

M_3: RB-ROM

Least-squares on reduced subspace

$$\tilde{\varphi}_\mu^N = \underset{p \in X^N}{\arg\min} \| g(\mu) - S(\mu)p \|_X^2$$

[Fabrini/Iapichino/Volkwein'18, K./Renelt'24]

M_2: F-ROM

Combination of RB-ROM and system reduction

$$\hat{\varphi}_\mu^{N,\text{red}} = \underset{p \in X^N}{\arg\min} \left\| \hat{g}(\mu) - \hat{S}(\mu)p \right\|_X^2$$

M_1: ML-F-ROM

Approximate $\hat{\varphi}_\mu^{N,\text{red}}$ as a function of the parameter using machine learning

Primal and adjoint systems

$$E \frac{d}{dt} x_\mu(t) = A(\mu; t) x_\mu(t) + B(\mu; t) u_\mu(t)$$

$$-E_{ad} \frac{d}{dt} \varphi_\mu(t) = A(\mu; t)^* \varphi_\mu(t)$$

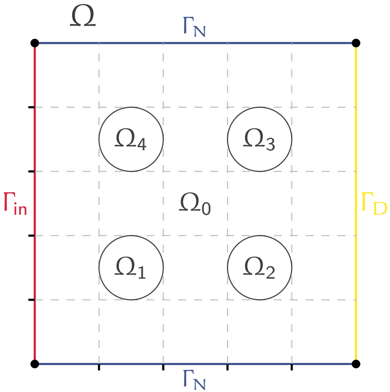
[Haasdonk/Ohlberger'11]

Reduced primal and adjoint systems

$$\hat{E}_{pr} \frac{d}{dt} \hat{x}_\mu(t) = \hat{A}_{pr}(\mu; t) \hat{x}_\mu(t) + \hat{B}_{pr}(\mu; t) u_\mu(t)$$

$$-\hat{E}_{ad} \frac{d}{dt} \hat{\varphi}_\mu(t) = \hat{A}_{ad}(\mu; t)^* \hat{\varphi}_\mu(t)$$

Problem setting



- Heat equation as state system:

$$\partial_t \theta(t;\mu) - \nabla \cdot (\sigma(t;\mu)\nabla\theta(t;\mu)) = 0 \qquad \text{in } \Omega,$$
$$\sigma(t;\mu)\nabla\theta(t;\mu)\cdot\vec{n} = u(t) \qquad \text{on } \Gamma_{in},$$

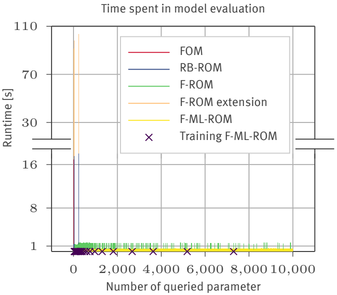
  \+ homogeneous initial, Dirichlet and Neumann conditions

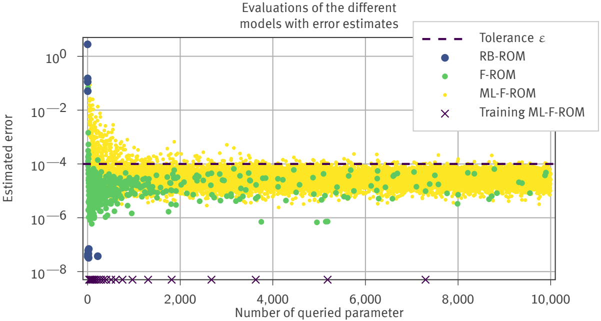
- Parametric and time-dependent diffusivity σ
- Output quantities: Average temperature in the four cookies
- Target state: Prescribed average temperature in all cookies

Example 2 – Parametrized optimal control problems

Numerical results

Model	Number of solves	Number of error estimates	Total time for error est. and solving [s]	Average time for error est. and solving per solve [s]
FOM	4	—	304.95	76.24
RB-ROM	12	16	234.56	19.55
F-ROM	437	453	450.39	1.03
ML-F-ROM	9,547	10,000	5,136.95	0.54





hendrik.kleikamp@uni-muenster.de

For more details, see:

- B. Haasdonk, H. Kleikamp, M. Ohlberger, F. Schindler, and T. Wenzel. *A new certified hierarchical and adaptive RB-ML-ROM surrogate model for parametrized PDEs, (2023)*. SIAM J. Sci. Comput., 45(3), A1039–A1065.
- T. Wenzel, B. Haasdonk, H. Kleikamp, M. Ohlberger, and F. Schindler. *Application of Deep Kernel Models for Certified and Adaptive RB-ML-ROM Surrogate Modeling, (2024)*. Proceeding of the LSSC 2023 Conference.
- H. Kleikamp. *Application of an adaptive model hierarchy to parametrized optimal control problems, (2024)*. Proceedings of the Conference Algoritmy, 66–75.
- H. Kleikamp and L. Renelt. *Two-stage model reduction approaches for the efficient and certified solution of parametrized optimal control problems, (2024)* arXiv preprint.

The source code for the papers is available open source:

- Parametrized parabolic PDEs
  - Three-stage hierarchy: `https://github.com/ftschindler/paper-2022-certified-adaptive-RB-ML-ROM-hierarchy`
- Parametrized optimal control problems
  - Three-stage hierarchy: `https://doi.org/10.5281/zenodo.10669855`
  - Four-stage hierarchy: `https://doi.org/10.5281/zenodo.13652744`