Universität Münster

Two-Stage Model Reduction Approaches for Parametrized Optimal Control Problems

SIAM CSE 2025 – Minisymposium on “Modern Approaches in Machine Learning for Model Order Reduction”

Hendrik Kleikamp, Lukas Renelt (University of Münster)

March 5, 2025

living.knowledge

MM Mathematics Münster Cluster of Excellence

Parametrized optimal control problems

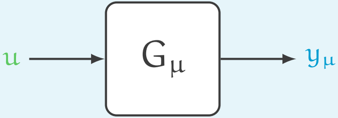
Control systems

$$u \longrightarrow \boxed{G_\mu} \longrightarrow y_\mu$$

- Parameter $\mu \in \mathcal{P}$
- Input/control $u\colon [0,T] \to U$
- Output $y_\mu\colon [0,T] \to Y$

Parametrized linear time-varying systems
General setting

Control systems



- Parameter $\mu \in \mathcal{P}$
- Input/control $u\colon [0,T] \to U$
- Output $y_\mu\colon [0,T] \to Y$

- State system relating controls to outputs:

$$G_\mu : \begin{cases} E\frac{d}{dt}x_\mu(t) &= A(\mu;t)x_\mu(t) + B(\mu;t)u(t) \\ y_\mu(t) &= Cx_\mu(t) \end{cases}$$

- (High-dimensional) state $x_\mu\colon [0,T] \to X$

Parametrized optimal control problems
General setting

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

General setting

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

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

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

hendrik.kleikamp@uni-muenster.de

Optimality system and final time adjoints

Theorem (K./Renelt'24)

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

$$
\begin{aligned}
E\frac{d}{dt}x_\mu(t) &= A(\mu;t)x_\mu(t) + B(\mu;t)u_\mu(t), && \textit{(state equation)}\\
u_\mu(t) &= -R(t)^{-1}B(\mu;t)^*\varphi_\mu(t), && \textit{(control equation)}\\
-E_{\mathrm{ad}}\frac{d}{dt}\varphi_\mu(t) &= A(\mu;t)^*\varphi_\mu(t), && \textit{(adjoint equation)}
\end{aligned}
$$

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,* $E_{\mathrm{ad}} \coloneqq \mathcal{R}_X E^* \mathcal{R}_X^{-1}$ *and* $M \coloneqq C^* \mathcal{R}_Y C$.

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!

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

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

where

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

$\Longrightarrow$ Costly to solve!

Reduced order models

Overview of the involved reduced order models

FOM

Optimal final time adjoint

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

Overview of the involved reduced order models

FOM

Optimal final time adjoint

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

Primal and adjoint systems

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

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

Overview of the involved reduced order models

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

FOM

Optimal final time adjoint

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

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

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

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

Overview of the involved reduced order models

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

FOM

Optimal final time adjoint

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

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

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

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

Overview of the involved reduced order models

Universität Münster

MM

FOM

Optimal final time adjoint

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

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

RB-ROM

Least-squares on reduced subspace

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

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

F-ROM

Combination of RB-ROM and system reduction

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

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

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

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

Overview of the involved reduced order models

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

FOM

Optimal final time adjoint

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

RB-ROM

Least-squares on reduced subspace

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

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

F-ROM

Combination of RB-ROM and system reduction

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

[Hesthaven/Ubbiali'18, K.'25]

ML-F-ROM

Approximate $\hat{\varphi}_\mu^{N,\mathrm{red}}$ 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

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

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

A posteriori error estimation for the reduced models

- Reliable a posteriori error estimators for RB-ROM, F-ROM and ML-F-ROM available

A posteriori error estimation for the reduced models

- Reliable a posteriori error estimators for RB-ROM, F-ROM and ML-F-ROM available

Theorem (Reliable a posteriori error estimator for the F-ROM)

For an approximate final time adjoint φ^N we have

$$\| \varphi_\mu^*(T) - \varphi^N \|_X \leqslant \eta_\mu^{\mathrm{red}}(\varphi^N).$$

- Reliable a posteriori error estimators for RB-ROM, F-ROM and ML-F-ROM available

Theorem (Reliable a posteriori error estimator for the F-ROM)

For an approximate final time adjoint φ^N we have

$$\left\| \varphi_\mu^*(T) - \varphi^N \right\|_X \leqslant \eta_\mu^{\text{red}}(\varphi^N).$$

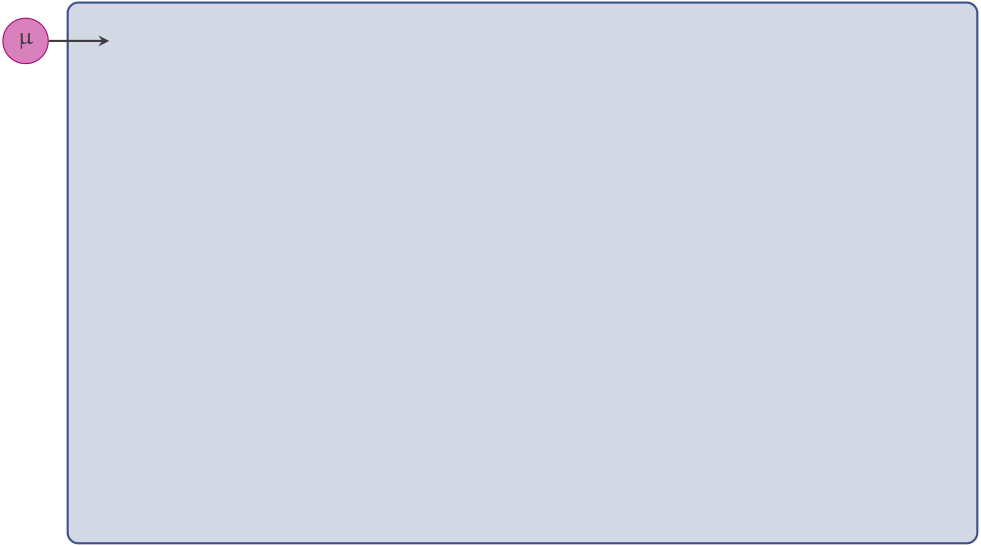
Main idea of the proof: Derive the bound

$$\begin{aligned}
\left\| \varphi_\mu^*(T) - \varphi^N \right\|_X \leqslant \boxed{\eta_\mu(\varphi^N)} &= C_{\text{op}} \left\| g(\mu) - S(\mu)\varphi^N \right\|_X \\
&= C_{\text{op}} \left\| \Big(g(\mu) - \hat{g}(\mu)\Big) + \Big(\hat{g}(\mu) - \hat{S}(\mu)\varphi^N\Big) + \Big(\hat{S}(\mu)\varphi^N - S(\mu)\varphi^N\Big) \right\|_{X'} \\
&\leqslant C \left(\Delta_\mu^{\text{pr}}[0](T) + \hat{\eta}_\mu(\varphi^N) + \Delta_\mu^\Lambda[\varphi^N]\right) \\
&=: \boxed{\eta_\mu^{\text{red}}(\varphi^N)}.
\end{aligned}$$

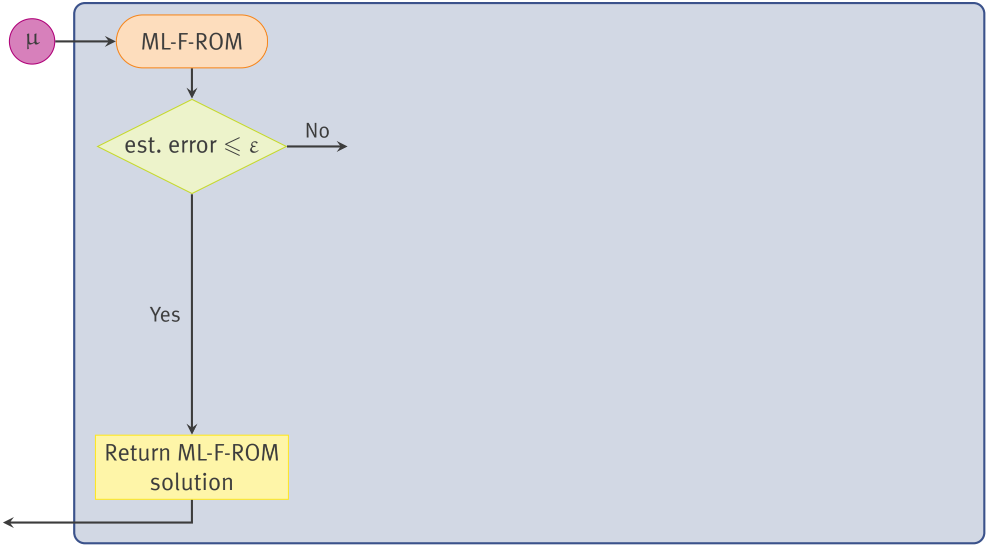
hendrik.kleikamp@uni-muenster.de

Adaptive model hierarchies

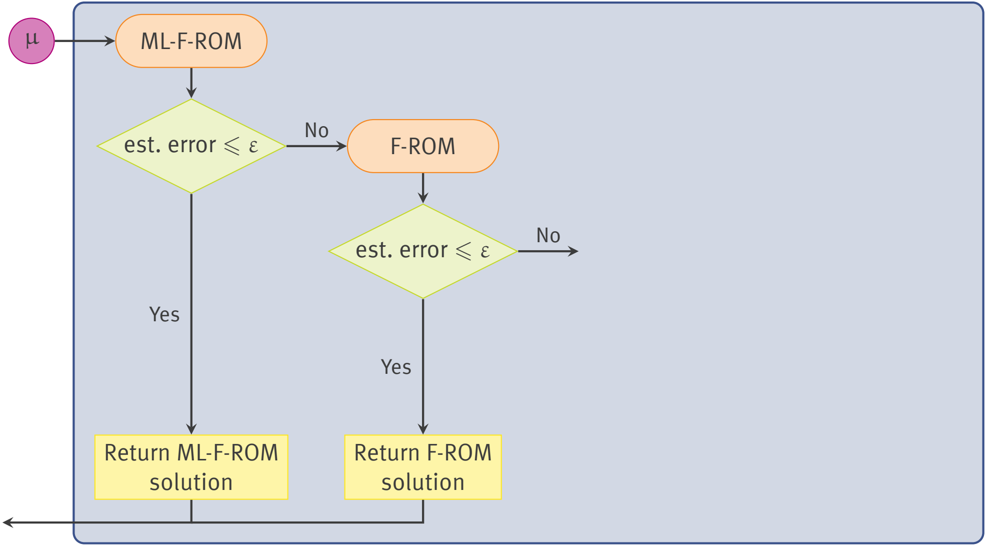
Four-stage hierarchy for optimal control problems



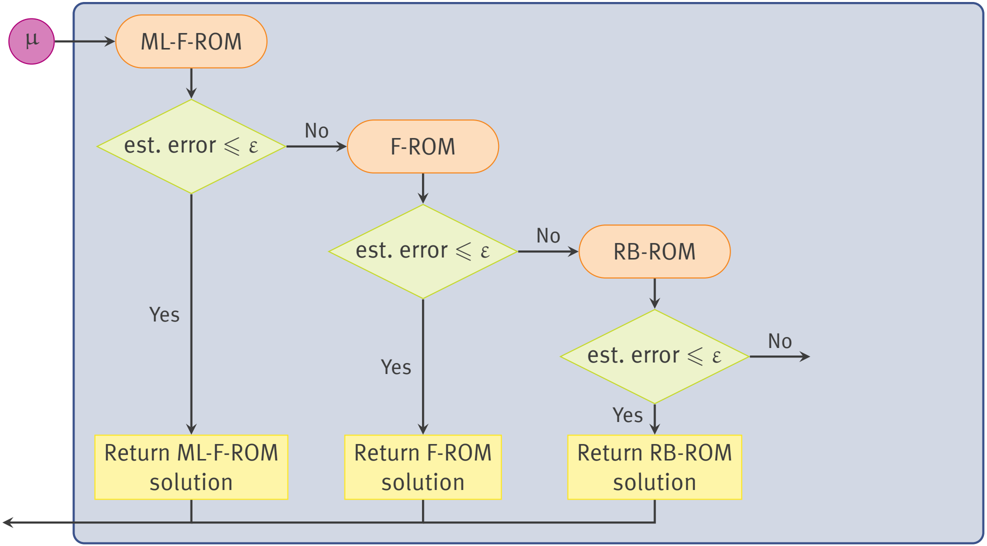
Four-stage hierarchy for optimal control problems



Four-stage hierarchy for optimal control problems



Four-stage hierarchy for optimal control problems

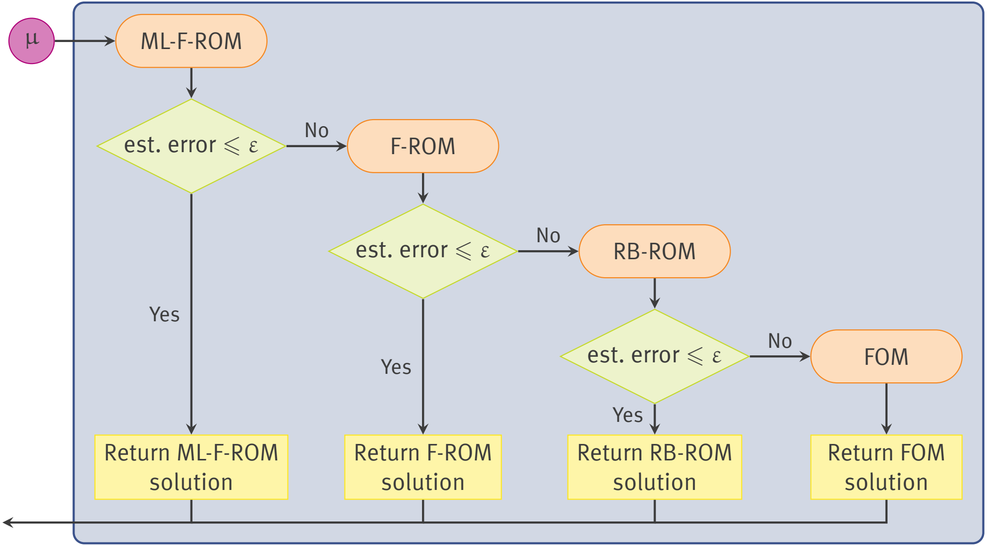


hendrik.kleikamp@uni-muenster.de

Four-stage hierarchy for optimal control problems

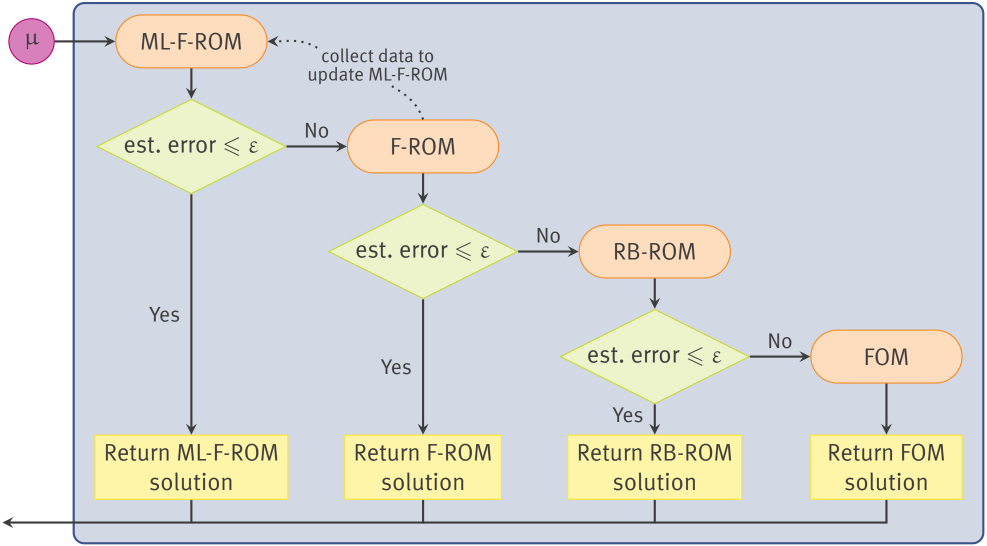
Universität Münster

MM

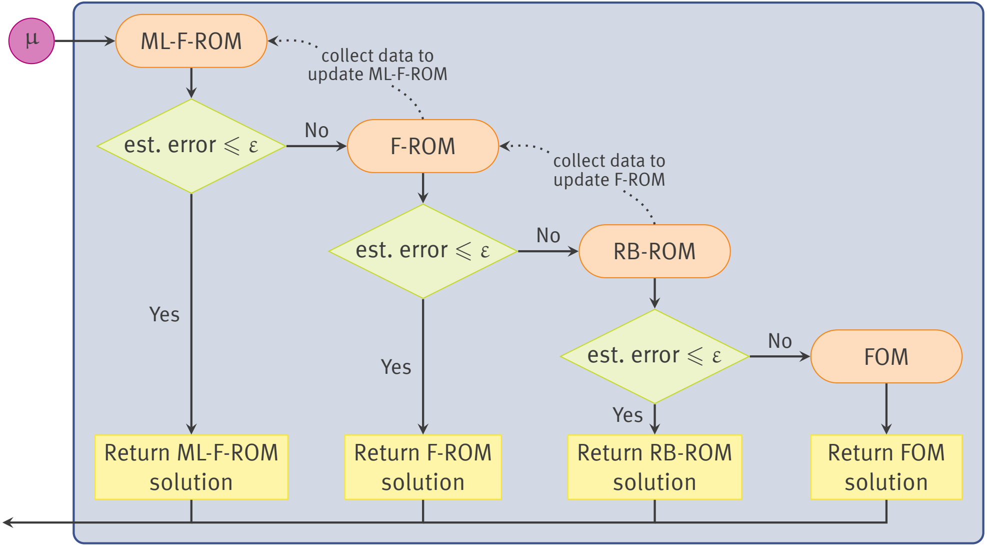


hendrik.kleikamp@uni-muenster.de

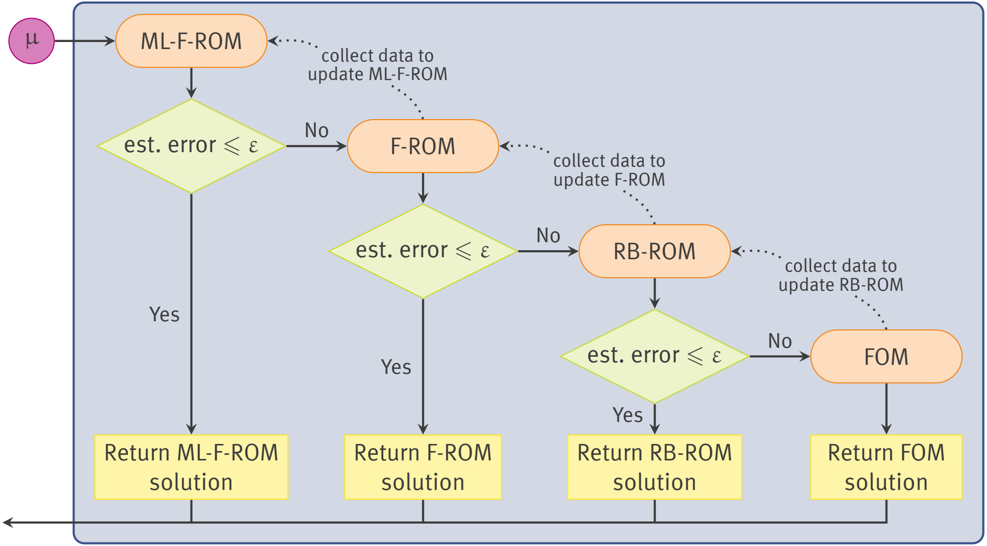
Four-stage hierarchy for optimal control problems



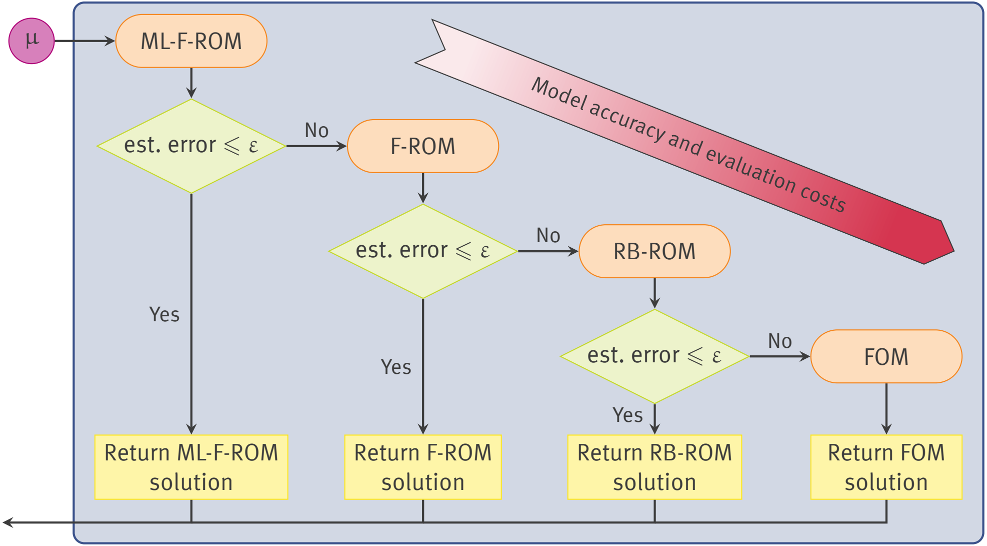
Four-stage hierarchy for optimal control problems



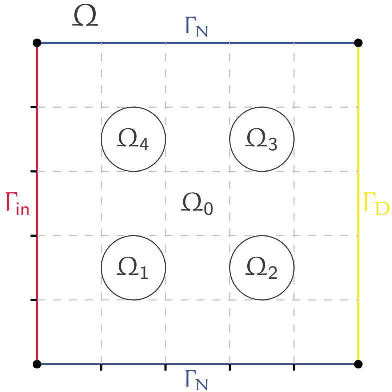
Four-stage hierarchy for optimal control problems



Four-stage hierarchy for optimal control problems



Numerical experiment



- Heat equation as state system:

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

  \+ 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

Numerical results for the four-stage hierarchy

Results for querying the hierarchy for 10,000 random parameters

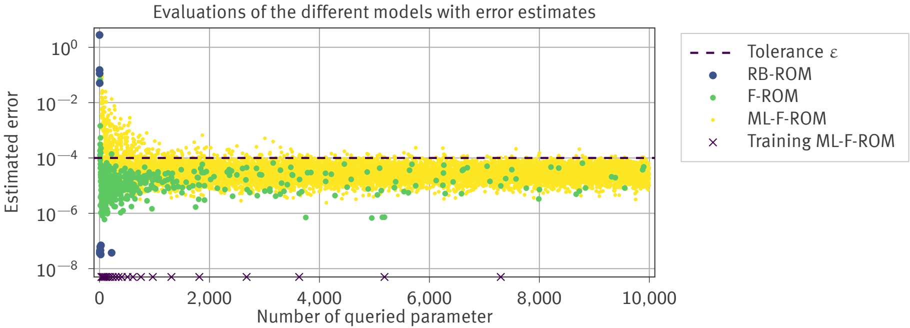
Model	Number of solves	Number of error estimates	Total time for error est. and solving [s]	Average time for error est. and solving per solve [s]
FOM	4	—	304.95	76.24
RB-ROM	12	16	234.56	19.55
F-ROM	437	453	450.39	1.03
ML-F-ROM	9,547	10,000	5,136.95	0.54

hendrik.kleikamp@uni-muenster.de

Numerical results for the four-stage hierarchy

Results for querying the hierarchy for 10,000 random parameters

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



Thank you for your attention!

Papers and code:

hendrik.kleikamp@uni-muenster.de