Localized Model Reduction in PDE Constrained Optimization

Mario Ohlberger, Michael Schaefer and Felix Schindler

Abstract. We present efficient localized model reduction approaches for PDE constraint optimization or optimal control. The first approach focuses on problems where the underlying PDE is given as a locally periodic elliptic multiscale problem. The second approach is more universal and focuses on general underlying multiscale or large scale problems. Both methods make use of reduced basis techniques and rely on efficient a posteriori error estimation for the approximation of the underlying parameterized PDE. The methods are presented and numerical experiments are discussed.

Mathematics Subject Classification (2010). Primary 65K10; Secondary 65K10.

Keywords. Localized model reduction, reduced basis methods, optimal control, PDE constrained optimization, LRBMS, heterogeneous multiscale method.

1. Introduction

In this contribution we are concerned with efficient approximation schemes for the following class of multiscale or large scale PDE constrained optimization problems.

$$\left.\begin{aligned} \text{Find} \quad & \mu^* = \arg\min J\big(u(\mu), \mu\big) \\ \text{subject to} \quad & C_j\big(u(\mu), \mu\big) \leq 0 \quad \forall j = 1, \dots, m, \\ & \mu \in \mathcal{P} \end{aligned}\right\} \tag{1.1}$$

with a compact *parameter set* $\mathcal{P} \subset \mathbb{R}^P$ for $P \in \mathbb{N}$. In (1.1), the *state variable* $u(\mu)$ is given as the solution of the following (parameterized) multiscale problem:

$$\left.\begin{aligned} -\nabla \cdot \big(A(\mu)\nabla u(\mu)\big) &= f(\mu) \quad (\text{in } \Omega) \\ u(\mu) &= 0 \qquad (\text{on } \partial\Omega) \end{aligned}\right\} \tag{1.2}$$

In (1.2), $\Omega \subset \mathbb{R}^d$ for $d = 1, 2, 3$ is a bounded domain and A denotes a diffusion tensor. We make use of the short notation $u(\mu) := u(\cdot; \mu)$ and will use analogue expressions for all functions that depend on both spatial variables and parameters.

We are particularly interested in multiscale or large scale applications in the sense that the diffusion tensor A has a rich structure that would lead to very high dimensional approximation spaces for the state space when approximated with classical finite element type methods. For parameter independent multiscale problems of this type there has been a tremendous development of suitable numerical multiscale methods in the last two decades including the multiscale finite element method (MsFEM) [24, 14, 15, 22], the heterogeneous multiscale method (HMM) [12, 13, 1, 3, 35, 20, 21], the variational multiscale method (VMM) [25, 26, 28, 29, 30, 31] or the more recent local orthogonal decomposition (LOD) [33, 19].

For parameterized partial differential equations, among others, reduced basis methods (RBM) have seen a great development in the last decade [8, 23, 41]. Meanwhile, also several applications of RBM in the context of multiscale methods have been proposed [9, 36, 37, 38, 39, 2, 10].

In this contribution we first derive a general framework for localized model reduction of multiscale or large scale PDE constrained optimization problems. We then focus on PDE constrained optimization, where the underlying multiscale problem has locally periodic structure, i.e. $A(x) = A^\varepsilon(x) = \hat{A}(x, \frac{x}{\varepsilon})$ and present an efficient combined RB-HMM approximation scheme. PDE constrained optimization for locally periodic structures has particular applications in shape optimization as, e.g., presented in [5, 17]. Finally, we also provide a localized approach for general multiscale or large scale problems based on the localized reduced basis multiscale method with online enrichment [38, 39]. This approach leads to a novel paradigm towards optimal complexity in PDE constrained optimization as recently introduced in [40].

2. Weak formulation for the parameterized multiscale problem and non-conforming reference approximation

Definition 2.1 (Weak solution of the multiscale problem). *We call $u(\mu) \in H^1_0(\Omega)$ weak solution of (1.2), if*

$$b\big(u(\mu), v; \mu\big) = L(v; \mu) \qquad \text{for all } v \in H^1_0(\Omega). \tag{2.1}$$

Here, the bilinear form b and the right hand side L are given as

$$b(v, w; \mu) := \int_\Omega A(\mu) \nabla v \cdot \nabla w, \quad L(v; \mu) := \int_\omega f v.$$

In order to derive a suitable formulation for (non-)conforming weak reference approximations for our model reduction approach, we first assume that a non-overlapping decomposition of the underlying domain Ω is given by a coarse grid $\mathcal{T}_H$ with cells $T_j \in \mathcal{T}_H, j = 1, \dots N_H$. Furthermore, each macro cell T_j is further decomposed by a local fine resolution triangulation $\tau_h(T_j)$,

that resolves all fine scale features of the multiscale problem. We then define the global fine scale partition τ_h as the union of all its local contributions, i.e., $\tau_h = \bigcup_{j=1}^{N_H} \tau_h(T_j)$. Let $H^p\big(\tau_h(\omega)\big) := \big\{v \in L^2(\omega) \;\big|\; v|_t \in H^p(t) \;\; \forall t \in \tau_h(\omega)\big\}$ for a triangulation $\tau_h(\omega)$ of some $\omega \subseteq \Omega$ denote the broken Sobolev space of order $p \in \mathbb{N}$ on τ_h, then $H^p(\tau_h)$ naturally inherits the decomposition $H^p(\tau_h) = \bigoplus_{j=1}^{N_H} H^p(\tau_h(T_j))$.

Definition 2.2 ((Non-)conforming approximations of the multiscale problem in broken spaces). *Let $V(\tau_h) \subset H^2(\tau_h)$ denote any approximate subset of the broken Sobolev space.*

We call $u_h(\mu) \in V(\tau_h)$ an approximate weak reference solution of (1.2), if

$$\mathcal{A}_{\mathrm{DG}}\big(u_h(\mu), v; \mu\big) = L_{\mathrm{DG}}(v;\mu) \qquad \text{for all } v \in V(\tau_h). \tag{2.2}$$

Here, the DG bilinear form $\mathcal{A}_{\mathrm{DG}}$ and the right hand side L_{DG} are given as

$$\begin{aligned}\mathcal{A}_{\mathrm{DG}}(v,w;\mu) &:= \sum_{t\in\tau_h}\int_t A(\mu)\nabla v\cdot\nabla w + \sum_{e\in\mathcal{F}(\tau_h)} \mathcal{A}^e_{\mathrm{DG}}(v,w;\mu)\\ L_{\mathrm{DG}}(v;\mu) &:= \sum_{t\in\tau_h}\int_t fv,\end{aligned}$$

where $\mathcal{F}(\cdot)$ denotes the set of all faces of a triangulation and the DG coupling bilinear form $\mathcal{A}^e_{\mathrm{DG}}$ for a face e is given by

$$\mathcal{A}^e_{\mathrm{DG}}(v,w;\mu) := \int_e \big\langle A(\mu)\nabla v\cdot\mathbf{n_e}\big\rangle[w] + \big\langle A(\mu)\nabla w\cdot\mathbf{n_e}\big\rangle[v] + \frac{\sigma_e(\boldsymbol{\mu})}{|e|^\beta}[v][w].$$

For any triangulation $\tau_h(\omega)$ of some $\omega \subseteq \Omega$, we assign to each face $e \in \mathcal{F}\big(\tau_h(\omega)\big)$ a unique normal $\mathbf{n_e}$ pointing away from the adjacent cell t^-, where an inner face is given by $e = t^- \cap t^+$ and a boundary face is given by $e = t^- \cap \partial\omega$, for appropriate cells $t^\pm \in \tau_h(\omega)$. In the above, the average and jump of a two-valued function $v \in H^2\big(\tau_h(\omega)\big)$ are given by $\big\langle v\big\rangle := \frac{1}{2}(v|_{t^-} + v|_{t^+})$ and $[v] := v|_{t^-} - v|_{t^+}$ for an inner face and by $\big\langle v\big\rangle := [v] := v$ for a boundary face, respectively. The parametric penalty function $\sigma_e(\mu)$ and the parameter β need to be chosen appropriately to ensure coercivity of $\mathcal{A}_{\mathrm{DG}}$ and may involve A. For simplicity, we restrict ourselves to the above symmetric interior penalty DG scheme; other DG variants can be easily accommodated and we refer to [16, 39] *and the references therein for further details.*

Note that Definition 2.2 contains both, continuous Galerkin finite element approximations, if $V(\tau_h) \subset H^2(\tau_h) \cap H^1_0(\Omega)$ and discontinuous Galerkin finite elements if $V(\tau_h) \subset H^2(\tau_h), V(\tau_h) \not\subset H^1_0(\Omega)$. In the continuous Galerkin case, we naturally have $\mathcal{A}_{\mathrm{DG}} \equiv b$ and $L_{\mathrm{DG}} \equiv L$.

3. A general non-conforming weak formulation for numerical multiscale methods

We assume that for each quadrature point x_T there exists open environments M_T, O_T with $x_T \in M_T \subset O_T$. We call M_T the local reconstruction region and O_T the local oversampling region. The subsets M_T and O_T are further decomposed by local fine resolution grids $\tau_h(M_T), \tau_h(O_T)$, which resolve all fine scale features of the multiscale problem. For sake of simplicity we assume that $\tau_h(M_T) \subset \tau_h(O_T) \subset \tau_h$.

Let $V^c := V(\mathcal{T}_H) \subset H^1(\mathcal{T}_H)$ denote a global finite dimensional coarse scale space, i.e., a finite dimensional subset of the broken Sobolev space $H^2(\mathcal{T}_H)$. Furthermore, let $V^f(O_T) := V(\tau_h(O_T)) \subset H^1(\tau_h(O_T))$ denote a suitable finite dimensional local fine scale space, that resolves all features of the underlying multiscale problem. We define $V^f(M_T) := V(\tau_h(M_T)) := V(\tau_h(O_T))|_{M_T} \subset H^1(\tau_h(M_T))$.

Definition 3.1 (General non-conforming multiscale approximation). *The macroscopic part of the multiscale approximation $u^c \subset V^c$ is defined as the solution of*

$$\text{find } u^c(\mu) \in V^c : \qquad \mathcal{A}^c_{\mathrm{DG}}(u^c(\mu), \Phi; \mu) = L_{\mathrm{DG}}(\Phi; \mu) \quad \text{for all } \Phi \in V^c.$$

Here, the coarse DG bilinear form $\mathcal{A}^c_{\mathrm{DG}} : V^c \times V^c \times \mathcal{P} \to \mathbb{R}$ is given as

$$\mathcal{A}^c_{\mathrm{DG}}(v, w; \mu) := \sum_{T \in \mathcal{T}_H} \frac{|T|}{|M_T|} \Bigg\{ \sum_{t \in \tau_h(M_T)} \int_t A(\mu) \nabla \mathcal{R}^\mu(v) \cdot \nabla w + \sum_{e \in \mathcal{F}(\tau_h(M_T))} \mathcal{A}^e_{\mathrm{DG}}\big(\mathcal{R}^\mu(v), w; \mu\big) \Bigg\}.$$

Furthermore, for any $\Phi \in V^c, T \in \mathcal{T}_H$, the local reconstruction operators $\mathcal{R}^\mu|_T : V^c \to V^c|_{O_T} + V^f(O_T)$ are defined through the solution of the following local problem on the oversampling domain O_T:

$$\text{find } \mathcal{R}^\mu|_T(\Phi) \in \Phi|_{O_T} + V^f(O_T) \text{ such that}$$

$$\mathcal{A}_{O_T}(\mathcal{R}^\mu|_T(\Phi), \phi; \mu) = F_{O_T}(\phi) \quad \forall \phi \in V^f(O_T),$$

where the fine scale bilinear form $\mathcal{A}_{O_T}$ is defined as

$$\mathcal{A}_{O_T}(v, w; \mu) := \sum_{t \in \tau_h(O_T)} \int_t A(\mu) \nabla v \cdot \nabla w + \sum_{e \in \mathcal{F}(\tau_h(O_T))} \mathcal{A}^e_{\mathrm{DG}}(v, w; \mu)$$

and $F_{O_T} \subset V^f(O_T)'$ is a suitable fine scale right hand side that depends on the particular realization of the method.

Note that for $M_T := T$ for all $T \in \mathcal{T}_H$, we get from the definition above $\mathcal{A}^c_{\mathrm{DG}}(v, w : \mu) = \mathcal{A}_{\mathrm{DG}}(\mathcal{R}^\mu(v), w; \mu)$.

Particular instances of multiscale methods that fit in the general framework of Definition 3.1 are the Heterogeneous Multiscale Method (HMM) and the Discontinuous Galerkin Multiscale Finite Element Method (DG-MsFEM)

with oversampling. These methods are obtained with the following specifications:

Definition 3.2. *The HMM with oversampling is given through Definition* 3.1 *with the following choices of domains and spaces:*

$V^c := V^1(\mathcal{T}_H) \cap H_0^1(\Omega) \subset H_0^1(\Omega)$,
$M_T := Y_{T,\varepsilon} := \{y \in \mathbb{R}^n \mid |||y - x_T||_\infty \leq \varepsilon\}$,
$O_T := Y_{T,\delta} := \{y \in \mathbb{R}^n \mid ||y - x_T||_\infty \leq \delta\}$, $\varepsilon \leq \delta$,
$V^f(M_T) := V_\#^1(\tau_h(M_T)) \subset \tilde{H}_\#^1(M_T)$,
$V^f(O_T) := V_\#^1(\tau_h(O_T)) \cap \tilde{H}_\#^1(O_T) \subset \tilde{H}_\#^1(O_T)$,
$F_{O_T}(\phi) \equiv 0$.

Here $\tilde{H}_\#^1(O_T)$ denotes the space of H^1-functions on O_T with periodic boundary values and zero mean and $V^1(\tau_h(\omega))$ ($V_\#^1(\tau_h(\omega))$) denotes the standard conforming piecewise linear Lagrange finite element space on ω (with mean zero and periodic boundary condition).

Definition 3.3. *The DG-MsFEM with oversampling is given through Definition* 3.1 *with the following choices of domains and spaces:*

$V^c := Q^1(\mathcal{T}_H) \subset H^2(\mathcal{T}_H)$,
$M_T := T$,
$O_T := U_l(T)$,
$V^f(M_T) := Q^1(\tau_h(T)) \subset H^2(\tau_h(T))$,
$V^f(O_T) := Q^1(\tau_h(U_l(T))) \subset H^2(\tau_h(U_l(T)))$,
$F_{O_T}(\phi) \equiv \int_\Omega f\phi$.

Here $Q^1(\mathcal{T}_H), Q^1(\tau_h(\omega))$ denote the standard non-conforming piecewise linear discontinuous Galerkin finite element space on $\mathcal{T}_H$, $\tau_h(\omega)$ respectively. Furthermore, $U_l(T)$ denotes an environment of T, consisting of l additional layers of fine scale elements $t \in \tau_h$.

4. A general framework for localized model reduction based on numerical multiscale methods with oversampling

Based on the general framework for numerical multiscale methods from Definition 3.1, we define general localized reduced basis multiscale methods (LRBMS) as follows.

Definition 4.1 (Localized Model Reduction Multiscale Methods). *Let $V_N^c \subset V^c$, $V_N^f(O_T) \subset V^f(O_T), T \in \mathcal{T}_H$ denote suitably defined local reduced coarse scale and fine scale subspaces that are obtained e.g. by a greedy algorithm from snapshots $u^c(\mu), (\mathcal{R}^\mu|_T(u^c) - u^c)(\mu)$ for a number of suitably chosen parameters μ. Furthermore, define $V_N^f(M_T) := \{v|_{M(T)} | v \in V^f(O_T)\}$, such that $V_N^f(M_T) \subset V^f(M_T)$. Then the corresponding localized reduced basis multiscale method is defined as the DG-Galerkin projection onto these subspaces as follows.*

The macroscopic part of the localized multiscale model reduction approximation $u_N^c \subset V_N^c$ is defined as the solution of

$$\text{find } u_N^c(\mu) \in V_N^c: \qquad \mathcal{A}_{\mathrm{DG}}^c(u_N^c(\mu), \Phi; \mu) = L_{\mathrm{DG}}(\Phi; \mu) \quad \text{for all } \Phi \in V_N^c.$$

Furthermore, for any $\Phi \in V_N^c, T \in \mathcal{T}_H$ the local reconstruction operators $\mathcal{R}^\mu|_T : V_N^c \to V_N^c|_T + V_N^f(O_T)$ are defined through the solution of the following local problem on the oversampling domain O_T:

find $\mathcal{R}^\mu|_T(\Phi) \in \Phi|_T + V_N^f(O_T)$ such that

$$\mathcal{A}_{O_T}(\mathcal{R}^\mu|_T(\Phi), \phi; \mu) = F_{O_T}(\phi) \quad \forall \phi \in V_N^f(O_T).$$

In the next two sections we will study in some more detail two localized model reduction methods that can be seen as reduced approximations of the RB-HMM on the one hand and of the DG-MsFEM with oversampling on the other hand.

5. The Reduced Basis Heterogeneous Multiscale Method (RB-HMM)

In this section we first review the RB-HMM method from [37] in the given optimization context and finally present some new numerical experiments.

The RB-HMM is defined through Definition 4.1 with the HMM choice given in Definition 3.2 and a greedy construction of the corresponding reduced subspaces. In the case of locally periodic homogenization problems, i.e. $A^\varepsilon(\mu) = A(x, \frac{x}{\varepsilon}; \mu)$, where $A(x, y; \mu)$ is 1-periodic with respect to y, it has been shown in [37], that the resulting method for $\delta = \varepsilon$ can be equivalently formulated as a reduced Galerkin approximation of the two-scale homogenized limit equations with quadrature. To this end, let us first introduce the continuous two scale solution space

$$\mathcal{H} := H_0^1(\Omega) \times L^2(\Omega; \tilde{H}_\#^1(Y))$$

which is a Hilbert space equipped with the scalar product

$$(u, v)_{\mathcal{H}} := ((u^0, u^1), (v^0, v^1)) := \int_\Omega \nabla u^0 \cdot \nabla v^0 \,\mathrm{d}x + \int_\Omega \int_Y \nabla_y u^1 \cdot \nabla_y v^1 \,\mathrm{d}y \,\mathrm{d}x.$$

Definition 5.1 (Discrete two scale homogenized bilinear form and HMM approximation). *We first define a piecewise constant approximation of the coefficient function $A(x, y; \mu)$ through*

$$\mathcal{B}_h(x, y; \mu)\big|_{T \times t} := A(x_T, y_t; \mu) \quad \forall \ T \times t \in \mathcal{T}_H \times \tau_h(Y).$$

where y_t is the center of gravity of t. We then define the discrete two scale homogenized bilinear form $\mathcal{B}_h : \mathcal{H} \times \mathcal{H} \times \mathcal{P} \to \mathbb{R}$ as

$$\mathcal{B}_h(u, v; \mu) := \int_\Omega \int_Y \mathcal{B}_h(\mu) \left(\nabla u^0 + \nabla_y u^1\right) \cdot \left(\nabla v^0 + \nabla_y v^1\right) \, dy \, dx.$$

With $\mathcal{H}_H := V^1(\mathcal{T}_H) \times V^0(\Omega; V^1(\tau_h(Y))) \subset \mathcal{H}$ it has been shown in [35] that the HMM approximation $u^c \in V^1(\mathcal{T}_H)$ from Definitions 3.1, 3.2 can be equivalently defined as the solution of the following discrete two scale limit equation.

Find $u_H(\mu) = (u^c(\mu), u^f(\mu)) \in \mathcal{H}_H$ such that for all $v_H := (v^c, v^f) \in \mathcal{H}_H$ we have

$$\mathcal{B}_h(u_H(\mu), v_H; \mu) = L(v^c; \mu). \tag{5.1}$$

Finally we obtain the following equivalent formulation of the RB-HMM method defined above.

Definition 5.2 (RB-HMM in the two scale limit formulation). *Consider a* sample set $\mathcal{S} = \{\mu_1, \ldots, \mu_N\} \subset \mathcal{P}$ *chosen, e.g., from a greedy algorithm. We define the* RB-HMM space $\mathcal{H}_N \subset \mathcal{H}_H$ *as*

$$\mathcal{H}_N = span\{u_H(\mu_i), i = 1, \ldots, N\} \tag{5.2}$$

with $u_H(\mu_i) \in \mathcal{H}_H$ *solving* (5.1) *for* $\mu_i \in \mathcal{S}$. *The RB-HMM approximation* $u_N(\mu) \in \mathcal{H}_N$ *of* $u_H(\mu)$ *for any* $\mu \in \mathcal{P}$ *is then defined as the Galerkin projection of* (5.1) *onto* $\mathcal{H}_N$, *i.e.*

$$\mathcal{B}_h(u_N(\mu), v_N; \mu) = L(v_N) \qquad (v_N \in \mathcal{H}_N) \tag{5.3}$$

5.1. A posteriori error estimation

Based on the Galerkin framework it is hence straight forward to derive a robust and efficient a posteriori error estimation that allows for an efficient offline/online decomposition as has been detailed in [37]. To this end, we introduce the *residual* $\mathrm{Res} : \mathcal{P} \to \mathcal{H}^{-1}$ via

$$-\langle \mathrm{Res}(\mu), v\rangle_{\mathcal{H}^{-1}\times\mathcal{H}} := L(v; \mu) - \mathcal{B}_h(u_N(\mu), v; \mu) \tag{5.4}$$

With the Hilbert space structure of $\mathcal{H}_H$ we further define its Riesz representative $v_H^{\mathrm{Res}}(\mu) \in \mathcal{H}_H$ through

$$\left(v_H^{\mathrm{Res}}(\mu), v_H\right)_{\mathcal{H}} = \langle \mathrm{Res}(\mu), v_H\rangle_{\mathcal{H}^{-1}\times\mathcal{H}}$$

for all $v_H \in \mathcal{H}_H$. If we assume $A(x, y; \mu)\xi \cdot \xi \geq \alpha(\mu)\,|\xi|^2$ for all $(x, y, \mu) \in \Omega \times Y \times \mathcal{P}$ and $\xi \in \mathbb{R}^d$ with $\inf_{\mu\in\mathcal{P}} \alpha(\mu) > 0$, the bilinear form $\mathcal{B}_h$ is coercive and we get the following theorem.

Theorem 5.3 (A posteriori error estimate).

$$||u_H(\mu) - u_N(\mu)||_{\mathcal{H}} \leq \frac{\left|\left|v_H^{Res}(\mu)\right|\right|_{\mathcal{H}}}{\alpha(\mu)} =: \Delta_N(\mu) \tag{5.5}$$

It is this a posteriori error estimator that we use in the greedy construction of the RB-HMM space.

5.2. Optimization and computation of derivative information

Following [37], the optimization problem (1.1) will be approximated in our RB-HMM setting as follows. We define the RB-HMM approximations of the functionals J and C_j through

$$\begin{aligned} J_N(\mu) &:= J(u_N^c(\mu), \mu), \\ C_{N,j}(\mu) &:= C_j(u_N^c(\mu), \mu). \end{aligned}$$

The *reduced optimization problem* then reads

$$\begin{aligned} \text{Find} \quad & \mu_N^* = \arg\min J_N(\mu) \\ \text{subject to} \quad & C_{N,j}(\mu) \leq 0 \quad \forall j = 1, \dots, m, \\ & \mu \in \mathcal{P} \end{aligned} \tag{5.6}$$

Note that the reduced functionals involve only the macroscopic part of $u_N(\mu)$. For an efficient optimization process, the involved functionals must allow an offline/online splitting. We thus assume that J_N and $C_{N,j}$ allow a representation as follows

$$\begin{aligned} J_N(\mu) &= \sum_{q=1}^{Q_J} \sigma_J^q(\mu) J_N^q\left(\varphi_1^N, \dots, \varphi_N^N\right), \\ C_{N,j}(\mu) &= \sum_{q=1}^{Q_{C_j}} \sigma_{C_j}^q(\mu) C_{N,j}^q\left(\varphi_1^N, \dots, \varphi_N^N\right), \end{aligned}$$

where the mappings J_N^q and $C_{N,j}^q$ can (in principle) be arbitrary. These expansions ensure that we can separate the parts of the functionals that depend on the basis functions from those depending on parameters which is necessary to achieve the offline/online splitting.

The only thing that is still missing are the parameter derivatives of the state $u_N(\mu)$. Equations for these quantities are established by differentiating the defining equation (5.3) of $u_N(\mu)$ with respect to μ_i. This results in a weak formulation for $\partial_{\mu_i} u_N(\mu) \in \mathcal{H}_N$:

$$\mathcal{B}_h(\partial_{\mu_i} u_N(\mu), v_N; \mu) = -\partial_{\mu_i} \mathcal{B}_h(u_N(\mu), v_N; \mu) \tag{5.7}$$

for all $v_N \in \mathcal{H}_N$, if the functional L does not depend on μ. Since $\mathcal{B}_h$ is affinely decomposable, we have

$$\partial_{\mu_i} \mathcal{B}_h(u, v; \mu) = \sum_{q=1}^{Q_A} \partial_{\mu_i} \sigma_q^A(\mu) \mathcal{B}_h^q(u, v).$$

Hence, we can reuse both the same reduced spaces and precomputed reduced system matrices as for the approximation of $u_N(\mu)$. Thus the computational costs for the parameter derivatives are negligible. Higher order derivatives can be specified by further differentiation of (5.7).

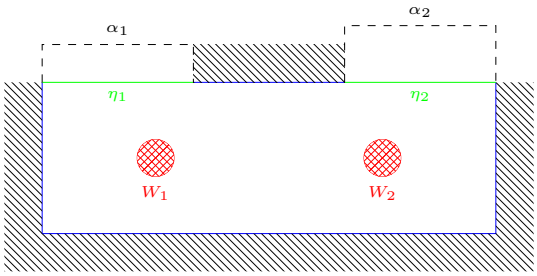

FIGURE 1. Geometry of the model problem: Two heat sources (red), no flow (blue) and Robin (green) boundary conditions. The head transfer coefficients and surrounding temperatures are $\eta_1 = 0.3$, $\eta_2 = 0.4$ resp. $\alpha_1 = \alpha_2 = 300$.

	detailed [s]	**reduced [s]**	**speed-up**	
single solve	0.5	0.1	~ 5	$(h = 4 \cdot 10^{-3})$
	1.5	0.1	~ 15	$(h = 2 \cdot 10^{-3})$
	6.4	0.1	~ 64	$(h = 1 \cdot 10^{-3})$
optimization	341	24	~ 14	$(h = 4 \cdot 10^{-3})$
	999	24	~ 42	$(h = 2 \cdot 10^{-3})$

TABLE 1. Runtime comparison between detailed and reduced simulations for a single forward solve and the complete optimization problem.

5.3. Numerical experiments

As a benchmark for the RB-HMM in a homogenization setting we look at the following model problem. For some scale parameter $\varepsilon > 0$ let $A^\varepsilon : \Omega \times \mathcal{P} \to \mathbb{R}^{d\times d}$ be a parameterized rapidly oscillating diffusion tensor, $\eta, g_N : \Gamma_N \to \mathbb{R}$ and f some right hand side. For $\mu \in \mathcal{P}$ we seek $u^\varepsilon(\mu)$ as solution of

$$\begin{aligned} -\nabla \cdot (A^\varepsilon(\mu)\nabla u(\mu)) &= f && \text{in } \Omega, \\ u^\varepsilon(\mu) &= 0 && \text{on } \Gamma_D, \\ -A^\varepsilon(\mu)\nabla u(\mu) \cdot n &= \eta\ (u^\varepsilon(\mu) - g_N) && \text{on } \Gamma_N, \end{aligned} \tag{5.8}$$

where $n(x)$ is the outer normal on Γ_D.

In particular we choose $\Omega \equiv (0, 0.6) \times (0, 0.2) \subset \mathbb{R}^2$ and define $\Gamma_{N_1} \equiv [0, 0.2] \times \{0.2\}$, $\Gamma_{N_2} \equiv [0.4, 0.6] \times \{0.2\}$ and $\Gamma_N \equiv \partial\Omega \setminus (\Gamma_{N_1} \cup \Gamma_{N_2})$ (see Figure 1). With $r_1 = r_2 = 0.025$ and $x_1 = (0.15, 0.1), x_2 = (0.45, 0.1) \in \mathbb{R}^2$ we introduce the source term $f : \Omega \to \mathbb{R}$, the Neumann value function $g_N : \partial\Omega \to \mathbb{R}$ and the heat transfer coefficient function $\eta : \partial\Omega \to \mathbb{R}$ as

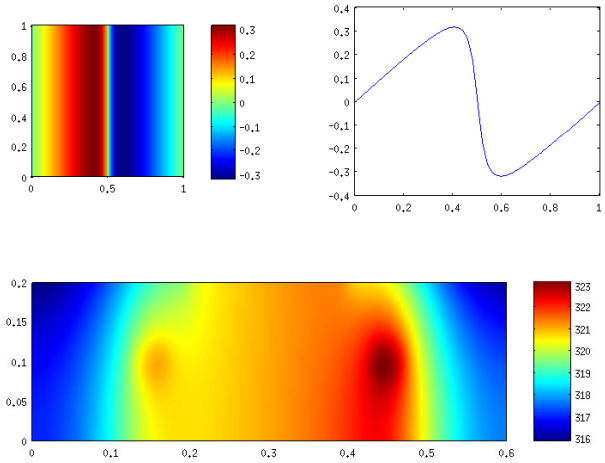

FIGURE 2. Example of a solution of the test problem for $\mu = (0.9112, 0.0062)$. Top left: $\chi^1_H(\mu)$, top right: cross-section plot of $\chi^1_H(\mu)$ in y_1-direction; bottom: $u^0_H(\mu)$. The effective diffusion tensor is $\mathcal{B}^0_h(\mu) = \mathrm{diag}(0.069, 0.429)$.

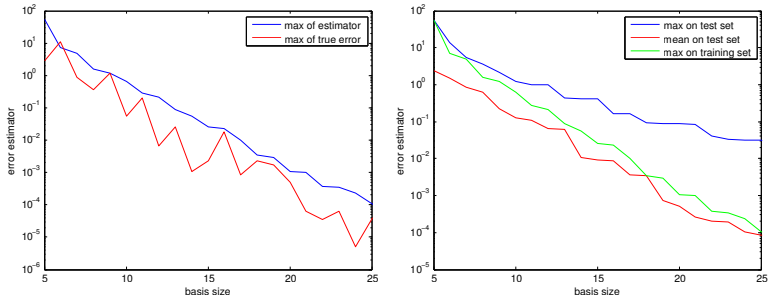

FIGURE 3. Convergence of the error estimator. Left: Estimator and true error on training set. Right: Estimator on test set.

$$f(x) \equiv 500 \cdot \mathbf{1}_{\mathrm{B}_{r_1}(x_1)}(x) + 800 \cdot \mathbf{1}_{\mathrm{B}_{r_2}(x_2)}(x), \tag{5.9}$$

$$g_N(x) \equiv 300 \cdot \mathbf{1}_{\Gamma_{N_1} \cup \Gamma_{N_2}}(x), \tag{5.10}$$

$$\eta(x) \equiv 0.3 \cdot \mathbf{1}_{\Gamma_{N_1}}(x) + 0.4 \cdot \mathbf{1}_{\Gamma_{N_2}}(x). \tag{5.11}$$

Additionally, we define the parameter set $\mathcal{P} \equiv [0.001, 1]^2$ and the parameterized diffusion tensor as

$$A^\varepsilon(\mu) = A(x, \frac{x}{\varepsilon}; \mu) \text{ with } A(x, y; \mu) = 16 y_1^2 (1 - y_1)^2 (\mu_2 - \mu_1) + \mu_1.$$

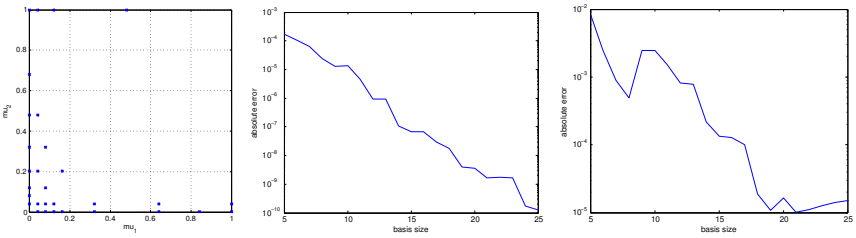

FIGURE 4. Parameters from the training set chosen by the greedy algorithm (left). Convergence of the reduced objective functional (middle) and optimal parameter value (right) with respect to the high dimensional reference optimization problem.

As an objective functional for the optimization we choose

$$J(u,\mu) = \int_\Omega (u - u_{\text{ref}})^2 \,\mathrm{d}x + \frac{\alpha}{2} \,||\mu||^2$$

with a reference solution $u_{\text{ref}} = u^0(\mu), \mu = (0.9112, 0.0062)$.

Note that since the diffusion tensor does not depend on y_2, the cell solutions $\chi^2(\mu)$ vanish. Samples of the macroscopic and the first cell solution are displayed in Figure 2 (the macroscopic solution is also used as target function for the optimization problems). Additionally, we impose the restriction $\mu \in \mathcal{P}$.

The convergence behavior of the a posteriori error estimator is shown in Figure 3. Figure 4 shows the parameters chosen by the greedy algorithm in the offline phase for the reduced basis construction. Table 1 shows a runtime comparisons. For that we have run the detailed and reduced optimization procedure for different underlying grid widths. The detailed optimization scales more or less quadratically with the grid width while in the reduced setting it stays constant. In the detailed optimization, gradient and Hessian information are approximated with finite differences, while in the reduced setting they can be easily generated directly form the corresponding reduced model problem. This makes the detailed approach more sensitive to the discretization, resulting in the necessity of more iterations and function evaluations which slows down the whole process significantly. Finally, in Figure 4 we display the convergence of the reduced objective functional and the computed optimal parameter value with respect to the high dimensional reference computation in dependence of the basis size of the RB-HMM.

6. The Localized Reduced Basis Multiscale Method (LRBMS) with Online Enrichment

In this section we first review the LRBMS method with online enrichment from [38, 39] in the given optimization context. The LRBMS is defined

through Definition 4.1 with the DG-MsFEM choice given in Definition 3.3. The construction of the corresponding local reduced subspaces is done using a localized a posteriori error estimator and a greedy online enrichment strategy during the optimization loop. This procedure follows a new paradigm that we recently introduced in [40].

The online enrichment procedure that we describe in Subsection 6.2 below constructs appropriate low dimensional local approximation spaces $U_N^j \subset V^c|_{T_j} \oplus V^f(T_j) \subset H^1(\tau_h^j)$ of local dimensions N_j that form the global reduced solution space via

$$U_H^N = \bigoplus_{j=1}^{N_H} U_N^j, \qquad N := \dim(U_H^N) = \sum_{j=1}^{N_H} N_j. \tag{6.1}$$

Once such a reduced approximation space is constructed, the LRBMS approximation is defined as follows.

Definition 6.1 (The localized reduced basis multiscale method). *We call $u_N(\mu) \in U_H^N$ a localized reduced basis multiscale approximation of (2.2) if it satisfies*

$$\mathcal{A}_{\mathrm{DG}}(u_N(\mu), v_N; \mu) = L_{\mathrm{DG}}(\mu; v_N) \qquad \text{for all } v_N \in U_H^N. \tag{6.2}$$

Note that $u_N(\mu)$ defined through (6.2) corresponds to $\mathcal{R}^\mu(u_N^c(\mu))$, where $u_N^c(\mu)$ is defined in 4.1. It solves a globally coupled reduced problem, where all arising quantities can nevertheless be locally computed w.r.t. the local reduced spaces U_N^j.

6.1. A posteriori error estimation

In recent contributions [27, 4, 39, 10] we discussed several possibilities to construct local reduced spaces U_N^j from global or localized snapshot computations. Thereby, in the concept presented above, it is possible to use finite volume, DG or conforming finite element approximations on the underlying fine partition τ_h or restrictions thereof to a local neighborhood of the macro elements $T_j \in \mathcal{T}_H$. In what follows, we consider the iterative construction of reduced approximation spaces and related surrogate models based on localized a posteriori error control and local enrichment as recently introduced in [39]. In these circumstances we obtain the following estimate on the error w.r.t. the unknown weak solution of (1.2).

Theorem 6.2 (Localizable a posteriori error estimate). *With the assumptions and the notation of* [39, Cor. 4.5]*, the following estimate on the full approximation error in the energy semi-norm $\|\!|v|\!\|_{\overline{\mu}} := \sum_{t\in\tau_h} \int_t (A^\varepsilon(\overline{\mu})\nabla v)\cdot\nabla v$ holds for arbitrary $\mu, \overline{\mu}, \hat{\mu} \in \mathcal{P}$,*

$$\|\!|u(\mu) - u_N(\mu)|\!\|_{\overline{\mu}} \leq \eta\big(u_N(\mu)\big) := C(\mu,\overline{\mu},\hat{\mu}) \Bigg[\sum_{j=1}^{N_H} \Big(\eta_j^{nc}(u_N(\mu))^2\Big)^{1/2} + \sum_{j=1}^{N_H} \Big(\eta_j^{r}(u_N(\mu))^2\Big)^{1/2} + \sum_{j=1}^{N_H} \Big(\eta_j^{df}(u_N(\mu))^2\Big)^{1/2} \Bigg]$$

with a computable constant $C(\mu, \overline{\mu}, \hat{\mu}) > 0$ and fully computable local indicators η_j^{nc}, η_j^r and η_j^{df} corresponding to the local non-conformity errors, residual errors, and diffusive flux reconstruction errors, respectively.

We refer to [39] for a more detailed presentation and derivation of this result and the definition of the corresponding local indicators. Note that a similar result holds for a full norm comprising the above semi-norm and a DG-jump semi-norm.

6.2. Optimization and Online Enrichment

Within an optimization loop, typically only parameters along a path towards the optimal parameter are depicted. As introduced in [40], we thus suggest a new iterative procedure to successively build up or enhance the surrogate model (6.2) by using the concept of local enrichment from [39]. Hence, only localized snapshot computations for the parameters that are selected during the optimization loop are computed. The resulting approach is thus tailored towards the specific optimization problem in an a posteriori manner.

Algorithm 6.1 Parameter optimization with adaptive enrichment.

Require: $\boldsymbol{\mu}^{(0)} \in \mathcal{P}$, initial local bases $\Phi_j{}^{(0)}$, $\Delta_{\text{model}}, \Delta_{\text{opt}} > 0$, a marking strategy `MARK` and an orthonormalization procedure `ONB` (see [39, Sec. 5]), an optimization routine `OPT` (returning a new parameter and status of convergence).

$n \leftarrow 0$, $U_H^{N\,(0)} \leftarrow \bigoplus_{j=1}^{N_H} \operatorname{span}\big(\Phi_j{}^{(0)}\big)$

repeat

  $m \leftarrow n$

  Solve (6.2) *for* $u_N^{(m)}(\boldsymbol{\mu}^{(n)}) \in U_H^{N\,(n)}$.

  while $\eta(\boldsymbol{\mu}^{(n)}) > \Delta_{\text{model}}$ **do**

    for all $j = 1, \dots, N_H$ **do**

      Compute local error indicators $\eta_j(\boldsymbol{\mu}^{(n)})$ *according to* [39, Cor. 4.5].

    end for

    $\tilde{\mathcal{T}}_H \leftarrow \texttt{MARK}\big(\mathcal{T}_H, \{\eta_j(\boldsymbol{\mu}^{(n)})\}_{j=1}^{N_H}\big)$

    for all $T_j \in \tilde{\mathcal{T}}_H$ **do**

      Solve locally on O_{T_j} *for enhanced local snapshot* $u_h^j(\boldsymbol{\mu}^{(n)}) \in H^1(\tau_h^j)$.

      $\Phi_j{}^{(m+1)} \leftarrow \texttt{ONB}\big(\{\Phi_j{}^{(m)}, u_h^j(\boldsymbol{\mu}^{(n)})\}\big)$

    end for

    $U_H^{N\,(m+1)} \leftarrow \bigoplus_{T_j \in \tilde{\mathcal{T}}_H} \operatorname{span}\big(\Phi_j{}^{(m+1)}\big) \oplus \bigoplus_{T_j \in \mathcal{T}_H \setminus \tilde{\mathcal{T}}_H} \operatorname{span}\big(\Phi_j{}^{(m)}\big)$

    Solve (6.2) *for* $u_N^{(m+1)}(\boldsymbol{\mu}^{(n)}) \in U_H^{N\,(m+1)}$.

    $m \leftarrow m + 1$

  end while

  $(\mu^{(n+1)}, \texttt{success}) \leftarrow \texttt{OPT}\big(\boldsymbol{\mu}^{(n)}, \Delta_{\text{opt}}, u_N^{(m)}(\boldsymbol{\mu}_n)\big)$

  $n \leftarrow n + 1$

until `success`

return optimal parameter $\boldsymbol{\mu}^{(n)}$ and state $u_N^{(m)}(\boldsymbol{\mu}^{(n)})$

In more detail, in a first step we initialize the local reduced spaces U_N^j with a classical polynomial coarse scale DG basis of prescribed order, thus ensuring that any reduced solution of the state equation is at least as good as

a DG solution on the coarse grid $\mathcal{T}_H$. During the following optimization loop, given any $\boldsymbol{\mu} \in \mathcal{P}$ from the optimization algorithm, we compute a reduced solution $u_N(\boldsymbol{\mu}) \in U_H^N$ and efficiently assess its quality using the localized a posteriori error estimator $\eta(\boldsymbol{\mu}) := (\sum_{j=1}^{N_H} \eta_j(\boldsymbol{\mu})^2)^{1/2}$ derived in [38, 39]. If the estimated error is above a prescribed tolerance, $\Delta > 0$, we start an intermediate local enrichment phase to enhance the surrogate model in the SEMR (solve → estimate → mark → refine) spirit of adaptive mesh refinement. We refer to Algorithm 6.1 and [39] for a detailed description and evaluation of this enrichment procedure that only involves local snapshot computations for the given parameter $\boldsymbol{\mu}$ on the local neighborhoods $O_{T_j}, T_j \in \mathcal{T}_H$ with Dirichlet boundary values obtained from the insufficient previous reduced surrogate. The algorithm calls a routine `OPT` that performs one optimization step with a descent method based on the old parameter value and the corresponding state with respect to the parameters. It returns the new parameter value and `success=true`, if the optimization criteria have been met.

6.3. Numerical experiment

To investigate the performance of the online enrichment, we consider (1.2) on $\Omega := [-1, 1]^2$ with right hand side $f(x, y) := \frac{1}{2}\pi^2 \cos(\frac{1}{2}\pi x)\cos(\frac{1}{2}\pi y)$, a parameter space $\mathcal{P} := [0, \pi]^2$ and a parametric scalar diffusion $A(\mu) := \sum_{\xi=1}^{2} \theta_\xi(\mu) A_\xi$ with coefficients $A_1 := \chi_{\Omega\setminus\omega}$, $A_2 := \chi_\omega$ and parameter functionals $\theta_1(\mu) := 1.1 + \sin(\mu_0)\mu_1$, $\theta_2(\mu) := 1.1 + \sin(\mu_1)$, where χ denotes an indicator function for the given domain and $\omega := [-\frac{2}{3}, -\frac{1}{3}]^2 \cup ([-\frac{2}{3}, -\frac{1}{3}] \times [\frac{1}{3}, \frac{2}{3}])$, compare Figure 5, top left. We are interested in minimizing the compliant quantity of interest (QoI) $J(u(\mu); \mu) := \int_\Omega f u(\mu)$ over $\mathcal{P}$. While this problem does not contain any multiscale features, it may serve as a model problem to compare model reduction using standard global RB methods with localized RB methods in the context of PDE-constrained optimization.

We discretize Ω by a triangulation τ_h with 2018 simplices and approximate the solution of (1.2) using a P^1-SWIPD scheme [16] (similar to Definition 2.2) with 6144 unknowns and use an L-BFGS-B algorithm [11] with a finite difference approximation of the objectives derivatives as optimization routine. We compare four different scenarios (compare Tab. 2), which we discuss further below: (i) using only the reference discretization ("standard FEM"); (ii) using a reduced order model (ROM) based on a single reduced basis with global support ("standard RB"); (iii) using a ROM based on a local reduced basis on each subdomain, containing only the constant 1 ("localized RB: Q^0-basis"); (iv) same as (iii), but with adaptive online enrichment of the local bases according to Algorithm 6.1 ("localized RB: adaptive").

Using the standard FEM approach (i) and starting from an initial guess of $\mu^{(0)} = (0.25, 0.5)$ we obtain the reference minimizer $\mu_h^* \approx (\frac{\pi}{2}, \pi)$ after 7 iterations of the optimization routine, with an additional 32 evaluations of J for the approximation of its derivatives (resulting in 39 evaluations of the reference discretization).

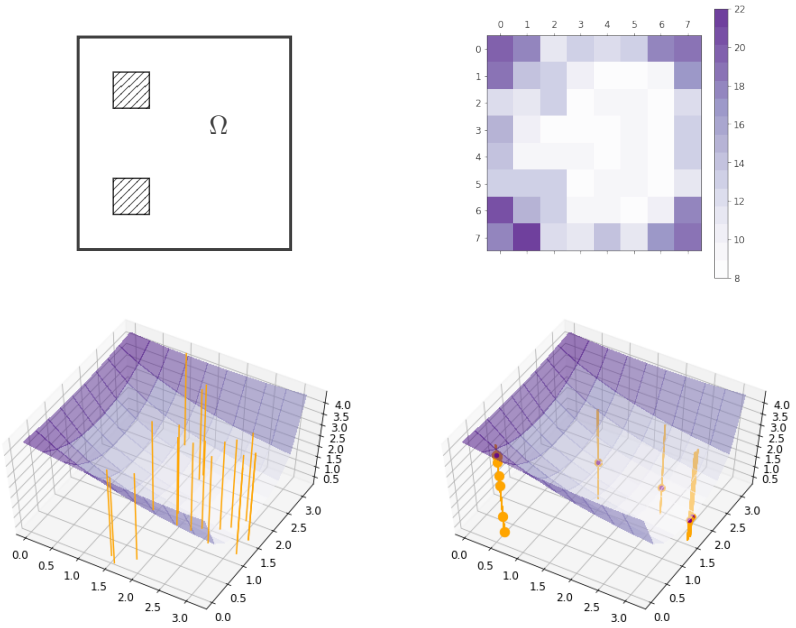

FIGURE 5. Physical domain and coarse grid (top row) and QoI J (purple surface, computed with a reference discretization) over parameter space (bottom row). Top left: Ω and ω (shaded regions). Top right: $\mathcal{T}_H$ with 8×8 subdomains and sizes of the local reduced bases (between 8 and 22) at the end of the computation. Bottom left: selected parameters (orange bars) during greedy basis generation of a standard RB approximation. Bottom right: selected parameters (orange bars) during optimization, intermediate evaluations of the QoI $J\big(u_N^{(m)}(\mu^{(n)}), \mu^{(n)}\big)$ during the adaptive enrichment (orange circles) and those given to `OPT` (purple dots).

In the standard RB approach (ii), we employ a weak greedy algorithm (using the standard residual based a posteriori error estimate on the model reduction error w.r.t. the energy product induced by the SWIPDG bilinear form for a fixed parameter $\overline{\mu} = (0, 0)$) to build a single reduced basis with global support, requiring 18 evaluations of the reference discretization to reach a model reduction error of $1.77 \cdot 10^{-11}$ over the training set of 625 equally distributed parameters. We did not employ an a posteriori error estimate on the quantity of interest, since the estimate on the state is an equivalent one in the present compliant setting (compare [18]). As we observe in Tab. 2, third column, using this ROM during the optimization yields very satisfactory results (thus justifying this choice). Compared to the standard FEM approach, we require only half of the evaluations of the reference discretization (since the

TABLE 2. Comparison of the computational effort (in terms of global/local PDE solutions) and accuracy (in terms of relative errors in the QoI and minimizer) of different approaches in the context of PDE-constrained optimization.

	standard FEM	standard RB	localized RB	
			Q^0-basis	adaptive
#evaluations of the reference discretization	$7 + 32$	18	0	0
#local corrector problems	-	-	-	709
relative error in the minimizer w.r.t. μ_h^*	-	$4.53 \cdot 10^{-5}$	$9.18 \cdot 10^{-5}$	$4.50 \cdot 10^{-5}$
relative error in the QoI w.r.t. $J(\mu_h^*)$	-	$9.01 \cdot 10^{-9}$	$9.47 \cdot 10^{-1}$	$1.36 \cdot 10^{-5}$

finite difference approximation of the objectives derivative is now performed using the ROM). However, the purpose of the greedy algorithm is to build a reduced basis that is equally valid for the whole parameter space and the reference discretization is thus evaluated over a large part of the parameter space (compare Figure 5, bottom left) that is not required for the optimization (although a certain symmetry in the parameterization is detected). One could thus argue that too many evaluations of the reference discretization are required, compared to the (unknown) trajectory of the optimizer through the parameter space.

Using the localized RB approach only with the Q^0-basis without adaptive enrichment (iii) can be interpreted as a Finite Volume scheme on an 8×8 grid with a high quadrature. If one is only interested in finding the minimizer, this approach would already be sufficient (compare Tab. 2, fourth column), without any evaluation of the reference discretization. It is thus subject of future work to establish an a posteriori error estimate on the QoI for the localized RB approach, to automatically detect this scenario.

Finally, in the online adaptive localized RB approach (iv), in each step of the optimization routine, we use the localized a posteriori error estimate from Theorem 6.2 to select a subset of the subdomains by a Dörfler and age-based marking strategy and enrich the corresponding local reduced bases by solutions to local corrector problems posed on the neighborhood of each of these subdomains (containing all adjacent subdomains), see [39] and the references therein. Using this approach, we obtain a satisfactory approximation of the minimum as well as the minimizer (compare Tab. 2, fifth column), without requiring any evaluation of the reference discretization. However, we did require the solution of 709 local corrector problems.

It is clear that for the simple example at hand, the adaptive localized RB approach does not pay off in terms of computational time when compared to the standard FEM or standard RB approach. However, if applied to a large real-world multiscale problem, the solution of which requires the use of large computing clusters, the localized nature of this approach should show significant benefits over the other approaches.

We used the generic discretization toolbox `dune-gdt`[1], based on the `dune-xt`ensions [32] and the `DUNE` software framework [7, 6], together with the model reduction package `pyMOR` [34]. To reproduce the experiments follow the instructions on:

`https://github.com/ftschindler-work/proceedings-mbour-2017-lrbms-control`

References

[1] A. Abdulle. The finite element heterogeneous multiscale method: a computational strategy for multiscale PDEs. In *Multiple scales problems in biomathematics, mechanics, physics and numerics*, volume 31 of *GAKUTO Internat. Ser. Math. Sci. Appl.*, pages 133–181. Gakkōtosho, Tokyo, 2009.

[2] A. Abdulle and P. Henning. A reduced basis localized orthogonal decomposition. *J. Comput. Phys.*, 295:379–401, 2015.

[3] A. Abdulle and A. Nonnenmacher. Adaptive finite element heterogeneous multiscale method for homogenization problems. *Comput. Methods Appl. Mech. Engrg.*, 200(37-40):2710–2726, 2011.

[4] F. Albrecht, B. Haasdonk, S. Kaulmann, and M. Ohlberger. The localized reduced basis multiscale method. In *Proceedings of Algoritmy 2012, Conference on Scientific Computing, Vysoke Tatry, Podbanske, September 9-14, 2012*, pages 393–403. Slovak University of Technology in Bratislava, Publishing House of STU, 2012.

[5] G. Allaire, E. Bonnetier, G. Francfort, and F. Jouve. Shape optimization by the homogenization method. *Numer. Math.*, 76(1):27–68, 1997.

[6] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöfkorn, R. Kornhuber, M. Ohlberger, and O. Sander. A generic grid interface for parallel and adaptive scientific computing. II. Implementation and tests in DUNE. *Computing*, 82(2-3):121–138, 2008.

[7] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöfkorn, M. Ohlberger, and O. Sander. A generic grid interface for parallel and adaptive scientific computing. I. Abstract framework. *Computing*, 82(2-3):103–119, 2008.

[8] P. Benner, A. Cohen, M. Ohlberger, and K. Willcox. *Model Reduction and Approximation: Theory and Algorithms*, volume 15 of *Computational Science and Engineering*. SIAM Publications, Philadelphia, PA, 2017.

[9] S. Boyaval. Reduced-basis approach for homogenization beyond the periodic setting. *Multiscale Model. Simul.*, 7(1):466–494, 2008.

[10] A. Buhr, C. Engwer, M. Ohlberger, and S. Rave. ArbiLoMod, a Simulation Technique Designed for Arbitrary Local Modifications. *SIAM J. Sci. Comput.*, 39(4):A1435–A1465, 2017.

[11] R. H. Byrd, P. Lu, J. Nocedal, and C. Y. Zhu. A limited memory algorithm for bound constrained optimization. *SIAM J. Sci. Comput.*, 16(5):1190–1208, 1995.

[12] W. E and B. Engquist. The heterogeneous multiscale methods. *Commun. Math. Sci.*, 1(1):87–132, 2003.

[1] `https://github.com/dune-community/dune-gdt`

[13] W. E and B. Engquist. The heterogeneous multi-scale method for homogenization problems. In *Multiscale methods in science and engineering*, volume 44 of *Lect. Notes Comput. Sci. Eng.*, pages 89–110. Springer, Berlin, 2005.

[14] Y. Efendiev, T. Hou, and V. Ginting. Multiscale finite element methods for nonlinear problems and their applications. *Commun. Math. Sci.*, 2(4):553–589, 2004.

[15] Y. Efendiev and T. Y. Hou. *Multiscale finite element methods*, volume 4 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer, New York, 2009. Theory and applications.

[16] A. Ern, A. F. Stephansen, and P. Zunino. A discontinuous galerkin method with weighted averages for advection–diffusion equations with locally small and anisotropic diffusivity. *IMA J. Numer. Anal.*, 29(2):235–256, 2009.

[17] B. Geihe and M. Rumpf. A posteriori error estimates for sequential laminates in shape optimization. *Discrete Contin. Dyn. Syst. Ser. S*, 9(5):1377–1392, 2016.

[18] P. Haasdonk. *Reduced Basis Methods for Parametrized PDEsA Tutorial Introduction for Stationary and Instationary Problems*, chapter 2, pages 65–136.

[19] P. Henning, A. Malqvist, and D. Peterseim. A localized orthogonal decomposition method for semi-linear elliptic problems. *ESAIM Math. Model. Numer. Anal.*, 48(5):1331–1349, 2014.

[20] P. Henning and M. Ohlberger. The heterogeneous multiscale finite element method for elliptic homogenization problems in perforated domains. *Numer. Math.*, 113(4):601–629, 2009.

[21] P. Henning and M. Ohlberger. The heterogeneous multiscale finite element method for advection-diffusion problems with rapidly oscillating coefficients and large expected drift. *Netw. Heterog. Media*, 5(4):711–744, 2010.

[22] P. Henning, M. Ohlberger, and B. Schweizer. An adaptive multiscale finite element method. *Multiscale Model. Simul.*, 12(3):1078–1107, 2014.

[23] J. S. Hesthaven, G. Rozza, and B. Stamm. *Certified reduced basis methods for parametrized partial differential equations*. SpringerBriefs in Mathematics. Springer, Cham; BCAM Basque Center for Applied Mathematics, Bilbao, 2016. BCAM SpringerBriefs.

[24] T. Y. Hou and X.-H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.*, 134(1):169–189, 1997.

[25] T. J. R. Hughes. Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. *Comput. Methods Appl. Mech. Engrg.*, 127(1-4):387–401, 1995.

[26] T. J. R. Hughes, G. R. Feijóo, L. Mazzei, and J.-B. Quincy. The variational multiscale method - a paradigm for computational mechanics. *Comput. Methods Appl. Mech. Engrg.*, 166(1-2):3–24, 1998.

[27] S. Kaulmann, M. Ohlberger, and B. Haasdonk. A new local reduced basis discontinuous Galerkin approach for heterogeneous multiscale problems. *C. R. Math. Acad. Sci. Paris*, 349(23-24):1233–1238, 2011.

[28] M. G. Larson and A. Malqvist. Adaptive variational multiscale methods based on a posteriori error estimation: duality techniques for elliptic problems. In *Multiscale methods in science and engineering*, volume 44 of *Lect. Notes Comput. Sci. Eng.*, pages 181–193. Springer, Berlin, 2005.

[29] M. G. Larson and A. Malqvist. Adaptive variational multiscale methods based on a posteriori error estimation: energy norm estimates for elliptic problems. *Comput. Methods Appl. Mech. Engrg.*, 196(21-24):2313–2324, 2007.

[30] M. G. Larson and A. Malqvist. An adaptive variational multiscale method for convection-diffusion problems. *Comm. Numer. Methods Engrg.*, 25(1):65–79, 2009.

[31] M. G. Larson and A. Malqvist. A mixed adaptive variational multiscale method with applications in oil reservoir simulation. *Math. Models Methods Appl. Sci.*, 19(7):1017–1042, 2009.

[32] T. Leibner, R. Milk, and F. Schindler. Extending dune: The dune-xt modules. *Archive of Numerical Software*, 5:193–216, 2017.

[33] A. Malqvist and D. Peterseim. Localization of elliptic multiscale problems. *Math. Comp.*, 83(290):2583–2603, 2014.

[34] R. Milk, S. Rave, and F. Schindler. pyMOR – generic algorithms and interfaces for model order reduction. *SIAM Journal on Scientific Computing*, 38(5):S194–S216, jan 2016.

[35] M. Ohlberger. A posteriori error estimates for the heterogeneous multiscale finite element method for elliptic homogenization problems. *Multiscale Model. Simul.*, 4(1):88–114, 2005.

[36] M. Ohlberger and M. Schaefer. A reduced basis method for parameter optimization of multiscale problems. In *Proceedings of Algoritmy 2012, Conference on Scientific Computing, Vysoke Tatry, Podbanske, September 9-14, 2012*, pages 272–281, september 2012.

[37] M. Ohlberger and M. Schaefer. Error control based model reduction for parameter optimization of elliptic homogenization problems. In Yann Le Gorrec, editor, *1st IFAC Workshop on Control of Systems Governed by Partial Differential Equations, CPDE 2013; Paris; France; 25 September 2013 through 27 September 2013; Code 103235*, volume 1, pages 251–256. International Federation of Automatic Control (IFAC), october 2013.

[38] M. Ohlberger and F. Schindler. A-posteriori error estimates for the localized reduced basis multi-scale method. In J. Fuhrmann, M. Ohlberger, and C. Rohde, editors, *Finite Volumes for Complex Applications VII-Methods and Theoretical Aspects*, volume 77 of *Springer Proceedings in Mathematics & Statistics*, pages 421–429. Springer International Publishing, may 2014.

[39] M. Ohlberger and F. Schindler. Error control for the localized reduced basis multi-scale method with adaptive on-line enrichment. *SIAM J. Sci. Comput.*, 37(6):A2865–A2895, december 2015.

[40] M. Ohlberger and F. Schindler. Non-conforming localized model reduction with online enrichment: Towards optimal complexity in pde constrained optimization. volume 200, pages 357–365, 2017.

[41] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced basis methods for partial differential equations*, volume 92 of *Unitext*. Springer, Cham, 2016. An introduction, La Matematica per il 3+2.

Mario Ohlberger
Applied Mathematics Muenster
Einsteinstr. 62
D-48149 Münster
Germany
e-mail: mario.ohlberger@uni-muenster.de

Michael Schaefer
Applied Mathematics Muenster
Einsteinstr. 62
D-48149 Münster
Germany
e-mail: michael.schaefer@uni-muenster.de

Felix Schindler
Applied Mathematics Muenster
Einsteinstr. 62
D-48149 Münster
Germany
e-mail: felix.schindler@uni-muenster.de