

# LOCALIZED MODEL REDUCTION FOR NONLINEAR ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS: LOCALIZED TRAINING, PARTITION OF UNITY, AND ADAPTIVE ENRICHMENT\*

KATHRIN SMETANA<sup>†</sup> AND TOMMASO TADDEI<sup>‡</sup>

**Abstract.** We propose a component-based (CB) parametric model order reduction (pMOR) formulation for parameterized nonlinear elliptic partial differential equations. CB-pMOR is designed to deal with large-scale problems for which full-order solves are not affordable in a reasonable time frame or parameters' variations induce topology changes that prevent the application of monolithic pMOR techniques. We rely on the partition-of-unity method to devise global approximation spaces from local reduced spaces, and on Galerkin projection to compute the global state estimate. We propose a randomized data compression algorithm based on oversampling for the construction of the components' reduced spaces: the approach exploits random boundary conditions of controlled smoothness on the oversampling boundary. We further propose an adaptive residual-based enrichment algorithm that exploits global reduced-order solves on representative systems to update the local reduced spaces. We prove exponential convergence of the enrichment procedure for linear coercive problems; we further present numerical results for a two-dimensional nonlinear diffusion problem to illustrate the many features of our methodology and demonstrate its effectiveness.

**Key words.** parameterized PDEs, model order reduction, domain decomposition

**MSC codes.** 65N30, 41A45, 35J15

**DOI.** 10.1137/22M148402X

## 1. Introduction.

**1.1. Component-based model reduction for parameterized PDEs.** Numerical modeling and simulation is of paramount importance to predict the response, improve the design, monitor the structural health of engineering systems, and generate digital twins [48, 58]. Several problems of interest involve repeated solutions of a partial differential equation (PDE) for many values of the model parameters or require real-time responses: these tasks are prohibitively expensive for standard (e.g., finite element) methods. Parametric model order reduction (pMOR, [27, 30, 57]) aims to reduce the marginal cost associated with the solution of parameterized systems over a range of parameters. The goal of this paper is to develop a pMOR procedure for large-scale nonlinear elliptic PDEs with parameter-induced topology changes. This facilitates, e.g., building a digital twin from components equipped with local reduced-order models [36, 37] and the adaptation of the digital twin by exchanging components [36].

pMOR techniques rely on an offline/online decomposition to reduce marginal costs. During the offline phase, pMOR methods rely on several high-fidelity (HF)

---

\* Submitted to the journal's Methods and Algorithms for Scientific Computing section March 14, 2022; accepted for publication (in revised form) January 3, 2023; published electronically June 15, 2023.

<https://doi.org/10.1137/22M148402X>

**Funding:** The first author acknowledges support from the National Science Foundation under Award 2145364.

<sup>†</sup>Department of Mathematical Sciences, Stevens Institute of Technology, Hoboken, NJ 07030 USA (ksmetana@stevens.edu).

<sup>‡</sup>IMB, UMR 5251, Univ Bordeaux, 33400 Talence, France INRIA Bordeaux Sud-Ouest, Team MEMPHIS, 33400 Talence, France (tommaso.taddei@inria.fr).

solves to generate a reduced-order model (ROM) for the solution field. During the online phase, given a new value of the parameter, the ROM is solved to estimate the solution field and relevant quantities of interest. Monolithic pMOR methods rely on HF solves at the training stage, which might be unaffordable for very large-scale problems. Furthermore, pMOR methods rely on the assumption that the solution field is defined over a parameter-independent domain or over a family of diffeomorphic domains: they thus cannot deal with problems for which parametric variations induce topology changes.

To address these issues, several authors have proposed component-based (CB) pMOR procedures (cf. [33] and the review [11]). During the offline stage, a library of *archetype components* is defined and local reduced-order bases (ROBs) and local ROMs are built for each component; then, during the online stage, select components from the library are instantiated to form the global system and the global solution is estimated by coupling local ROMs. CB-pMOR strategies consist of two distinct building blocks: (i) a rapid and reliable domain decomposition (DD) strategy for online global predictions, and (ii) a localized training strategy exclusively based on local solves for the construction of the local approximations.

CB-pMOR shares important features with multiscale methods [3, 39, 66, 50, 51, 52, 53, 44, 65, 61, 41, 70, 40, 19, 17, 13, 18]. Similarly to CB-pMOR, multiscale methods rely on local solves to build suitable approximation spaces that are tailored to the problem of interest. The emphasis in CB-pMOR is to devise and then exploit a library of interoperable archetype components and associated ROMs that can be used for a broad range of potentially parameter dependent problems in a specific domain of interest.

**1.2. Domain decomposition strategies within CB-pMOR.** Since the seminal work by Maday and Rønquist [42]—that proposed a nonoverlapping nonconforming reduced basis element method based on mortar DD—several authors have combined DD methods with model reduction methods to devise effective CB-ROMs. As discussed in detail in the review [11], we can distinguish between conforming nonoverlapping approaches [33, 22, 65], nonconforming nonoverlapping approaches based on Lagrange multipliers [31, 34, 42, 54], nonconforming nonoverlapping approaches based on discontinuous Galerkin (DG) coupling [1, 2, 49], and overlapping methods [7, 12]. The vast majority of existing contributions (with few recent exceptions [5, 7, 31, 54]) are restricted to linear PDEs.

In this work, we rely on the partition-of-unity method (PUM) to devise global approximation spaces from local reduced spaces, and on Galerkin projection to compute the global state estimate. PUM was proposed by Babuška and Melenk in [4, 46] and further developed and analyzed in the framework of generalized finite element methods for multiscale problems (cf. [3]); PUM was also considered in the pMOR literature for linear elliptic and parabolic problems [12, 60]. In the CB-pMOR framework, PUM offers a general (i.e., independent of the underlying PDE) framework with strong theoretical guarantees.

**1.3. Localized training based on oversampling and randomization.** Given the domain  $\hat{\Omega}$  associated with a given archetype component, oversampling methods consist in (i) defining a patch  $\hat{\Omega}_{\text{ovr}} \supset \hat{\Omega}$  and a suitable local PDE problem in  $\hat{\Omega}_{\text{ovr}}$ , (ii) solving the local PDE for several choices of the boundary conditions on  $\partial\hat{\Omega}_{\text{ovr}}$  and then restricting the solution to  $\hat{\Omega}$ , and finally (iii) exploiting the results to build a local approximation space for the solution in  $\hat{\Omega}$ . Randomized methods rely on independent and identically distributed (iid) samples of the boundary conditions

on (a subset of)  $\partial\widehat{\Omega}_{\text{ovr}}$ : they thus require the introduction of a probability density function for the functions defined on  $\partial\widehat{\Omega}_{\text{ovr}}$ .

Oversampling methods exploit low-pass filtering properties of the differential operator to identify low-dimensional structures: we refer to [67, Chapter 5] and [65, Remark 3.3] for two representative working examples. In detail, Caccioppoli-type inequalities (see, e.g., [25]) provide the theoretical foundations for the application of oversampling methods to a particular class of PDEs. Oversampling methods have been suggested and used extensively in the context of multiscale methods (see, e.g., [32, 3, 29, 44] and references therein) and then used as well in CB-pMOR [22, 65] for linear PDEs.

As suggested in [12], randomized oversampling methods for linear parameter-independent PDEs can be linked to randomized singular value decomposition (SVD) techniques developed and analyzed in randomized numerical linear algebra [28, 45, 21, 43]: this link allows one to extend methodological and theoretical contributions in randomized linear algebra to CB-pMOR. In particular, we can exploit concentration inequalities to analyze the error of randomized techniques, and inform the choice of the sampling distribution. The influence of the choice of the sampling distribution for nonlinear PDEs remains an open question in CB-pMOR.

**1.4. Contributions of the paper and outline.** In this work, we propose a CB-pMOR procedure based on the PUM for parametric nonlinear elliptic PDEs; we do not require the nonlinear operator to be monotone. The contributions of the paper are twofold. First, we propose a randomized data compression algorithm based on oversampling: the approach relies on random samples of local parameters and boundary conditions on the oversampling boundary. We propose a new sampler that controls the smoothness of the boundary condition, and we empirically demonstrate its effectiveness for a nonlinear diffusion problem. Second, we propose a basis enrichment algorithm that relies on global reduced solves to enrich the local reduced spaces. The algorithm relies on a local residual-based error indicator to identify boundary conditions for which the local ROM is inaccurate and a rigorous global a posteriori error bound as a termination criterion. We prove in-sample a priori exponential convergence of the enrichment algorithm for linear coercive problems; we further investigate performance for a nonlinear diffusion problem.

Our randomized algorithm reads as a randomized proper orthogonal decomposition [74] with respect to parameter and boundary conditions. On the other hand, the enrichment algorithm is closely related to the online enrichment strategy proposed in [49] for nonoverlapping DG DD, and to the residual-based online enrichment algorithm considered in [9] for linear problems. The major difference is that the enrichment is performed at a training stage and aims to update the local approximation spaces associated with the archetype components, rather than during the online stage on the “instantiated components.” We note that in [31] the authors employ the randomized boundary conditions proposed in [22] for nonlinear PDEs. In contrast, in this manuscript, we try to generalize the approach of [12] for which one can show a quasi-optimal convergence behavior for linear problems. The randomized approach in [18] also aims at approximating nonlinear PDEs. To that end, the authors create a dictionary of solutions by solving the PDE for random boundary conditions and construct tangential approximations to the nonlinear map, which are then used to obtain a solution with a Schwarz iteration. While the random boundary conditions in [18] are chosen uniformly on the unit sphere considering, for instance, the  $H^{1/2}$ -norm for semilinear elliptic equations, we propose in this article to employ random boundary

conditions of various prescribed smoothness to exploit the expected faster decay of high frequencies from the boundary towards the interior of the target subdomains.

The outline of the paper is as follows. In section 2, we introduce the model problem considered throughout the paper to illustrate the main definitions and to numerically validate our proposal: the model problem involves a high-dimensional ( $\mathcal{O}(10^2)$ ) parameterization and topology changes. In section 3, we present the main ideas and contributions of the paper. In section 4, we discuss the DD strategy based on the PUM and we introduce local and global discrete approximation spaces; in section 5, we discuss the randomized localized data compression; in section 6, we present the enrichment strategy; and in section 7 we present thorough numerical investigations for the model problem. Section 8 concludes the paper. Appendix A summarizes the notation.

**2. Model problem: Nonlinear diffusion.** Given  $n_{\text{dd}} \in \mathbb{N}$  and  $H = 0.1$ , we define the domains

$$(2.1) \quad \Omega_{i+(j-1)n_{\text{dd}}} = \{[x_1 + H(i-1), x_2 + H(j-1)] : x_1, x_2 \in (0, H)\}$$

for  $i, j = 1, \dots, n_{\text{dd}}$ , and introduce the global domain  $\Omega = \bigcup_{k=1}^{n_{\text{dd}}} \Omega_k$  with  $N_{\text{dd}} = n_{\text{dd}}^2$ .

We consider the problem of parametric nonlinear diffusion: for every parameter  $\mu$  in the compact set of admissible parameters  $\mathcal{P}_{\text{glo}}$ , find  $u_\mu$  such that

$$(2.2) \quad \begin{cases} -\nabla \cdot (\kappa_\mu(x, u_\mu) \nabla u_\mu) = f_\mu & \text{in } \Omega, \\ u_\mu = 0 & \text{on } \partial\Omega, \end{cases}$$

where for every  $\mu \in \mathcal{P}_{\text{glo}}$  the function  $\kappa_\mu : \mathbb{R} \rightarrow \mathbb{R}_+$  is  $C^2(\mathbb{R})$  and uniformly elliptic and bounded, i.e., there exist constants  $\kappa_k > 0$ ,  $k = 0, 1, 2, 3$ , such that for all  $\mu \in \mathcal{P}_{\text{glo}}$

$$(2.3) \quad \begin{aligned} \kappa_0 \|w\|_2^2 &\leq \kappa_\mu(u_\mu) w \cdot w, \quad \|\kappa_\mu(u_\mu) w\|_2 \leq \kappa_1 \|w\|_2, \\ \|\kappa'_\mu(u_\mu) w\|_2 &\leq \kappa_2 \|w\|_2, \quad \|\kappa''_\mu(u_\mu) w\|_2 \leq \kappa_3 \|w\|_2 \quad \forall w \in \mathbb{R}^2 \quad \forall u_\mu \in \mathbb{R}, \end{aligned}$$

where  $\|\cdot\|_2$  denotes the Euclidean norm in  $\mathbb{R}^2$ . We further assume that the restriction of  $\kappa_\mu(u_\mu)$  to a subdomain  $\Omega_k$  is the same for every subdomain and that the map  $\mu \mapsto \kappa_\mu$  is continuous.<sup>1</sup> Moreover, we require that for every  $\mu \in \mathcal{P}$  the function  $f_\mu$  is in  $L^2(\Omega)$  and that the function  $\mu \mapsto f_\mu$  is continuous. This yields the existence of a unique solution (see, e.g., [20]) of the variational form for any  $\mu \in \mathcal{P}$ , find  $u_\mu \in H_0^1(\Omega)$  such that

$$(2.4) \quad \int_{\Omega} \kappa_\mu(u_\mu) \nabla u_\mu \nabla v = \int_{\Omega} f_\mu v \quad \forall v \in H_0^1(\Omega).$$

While we focus on (2.2) in this paper to ease the exposition of ideas, we emphasize that the proposed methods can be readily applied to other nonlinear elliptic PDEs.

*Example 2.1.* As one specific example of (2.2) we consider the following model problem that has been previously considered in [64], and is inspired by the model

<sup>1</sup>We note that this continuity assumption is necessary to draw samples from the parameters later within the POD algorithm. We note that Lipschitz continuity in the parameter was assumed for linear parameter dependent elliptic problems to show results on the approximation error caused by the POD in the infinite dimensional setting [35]. So a theoretical analysis especially in infinite dimensions or one that is robust with respect to the mesh size will likely require higher regularity assumptions.

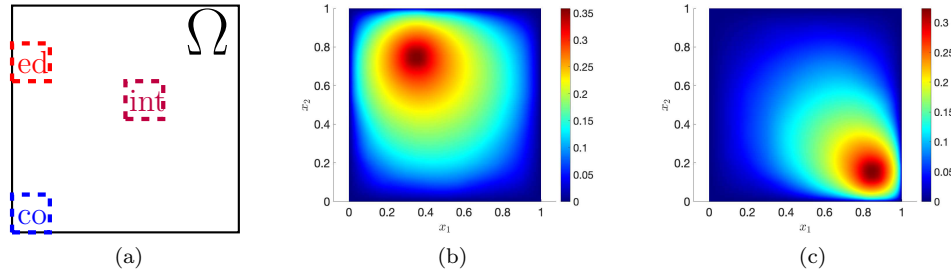


FIG. 1. *Nonlinear diffusion.* (a) *Instantiated archetype components.* (b)–(c) *solution fields for  $N_{\text{dd}} = 100$  and two choices of the parameters.*

for immiscible two-phase flows in porous media [47]. We introduce  $\widehat{\mathcal{P}} = [0.1, 0.2] \times [30, 40]$ , the permeability coefficient  $\kappa : \Omega \times \mathbb{R} \times \bigotimes_{i=1}^{N_{\text{dd}}} \widehat{\mathcal{P}} \rightarrow \mathbb{R}_+$  such that  $\kappa|_{\Omega_i} = \kappa(x; u, \mu^{(1)}, \dots, \mu^{(N_{\text{dd}})})|_{\Omega_i}$  satisfies

$$(2.5a) \quad \kappa|_{\Omega_i} = \frac{36}{\mu_2^{(i)}} \left( \frac{u(1-u)}{u^3 + \frac{12}{\mu_2^{(i)}}(1-u)^3} \right)^2 + \mu_1^{(i)}, \quad i = 1, \dots, N_{\text{dd}},$$

and the source term

$$(2.5b) \quad f(x; i^*) = 100 e^{-50 \|x - x_{c, i^*}\|_2^2} \mathbf{1}_{\Omega_{i^*}}(x).$$

The global parameter and the set of global admissible parameters for this problem are thus  $\mu = [\mu^{(1)}, \dots, \mu^{(N_{\text{dd}})}, i^*]$  and  $\mathcal{P}_{\text{glo}}(n_{\text{dd}}) := \bigotimes_{i=1}^{N_{\text{dd}}} \widehat{\mathcal{P}} \times \{1, \dots, N_{\text{dd}}\}$ . In Figures 1(b)–(c), we show the domain  $\Omega$  and selected solutions for different parameters  $\mu$  for  $N_{\text{dd}} = 100$ .

**3. Motivation and explanation of key ideas and contributions.** In this manuscript, we aim at approximating the global solution of the parametric nonlinear PDE by a CB ROM, where the local reduced basis functions are built from local solutions of the PDE. Therefore, the local reduced basis functions somehow have to capture all possible local behavior of the global solution. We illustrate why one can hope to accurately approximate all possible local behavior of the solution with relatively few local basis functions with an example [65, 60]: for the Laplacian on the domain  $\Omega = (-2, 2) \times (0, 1)$  with homogeneous Neumann boundary conditions at  $x_2 = 0$  and  $x_2 = 1$  and arbitrary Dirichlet conditions on  $x_1 = -2$  and  $x_1 = 2$  all solutions are of the form

$$(3.1) \quad u(x_1, x_2) = a_0 + b_0 x_1 + \sum_{n=1}^{\infty} \cos(n\pi x_2) [a_n \cosh(n\pi x_1) + b_n \sinh(n\pi x_1)],$$

where  $a_n, b_n \in \mathbb{R}$ ,  $n = 0, \dots, \infty$  are determined by the Dirichlet boundary data prescribed on  $x_1 = -2$  and  $x_1 = 2$ . We observe in Fig. 2(a) an exponential decay of the solutions in the interior of  $\Omega$  and that as a consequence most terms in the sum in (3.1) will be numerically zero around  $x_1 = 0$ . In addition, the more oscillations we have on the Dirichlet boundary (e.g.,  $-\cos(8\pi y)$  versus  $-\cos(2\pi y)$ ) the faster the decay is in the interior of  $\Omega$ . All this implies that already few local basis functions—here the functions  $\cos(\pi y), \cos(2\pi y), \cos(3\pi y), \dots$ —can approximate all possible local behavior of the solution well. One can show that these functions span an optimal

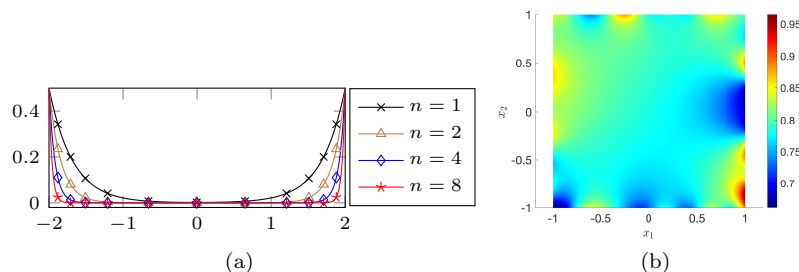


FIG. 2. (a) Solution  $u(x, 2/3)$  for Dirichlet boundary conditions  $-\cos(n\pi y)$  for  $n = 1, 2, 4, 8$  for the linear Laplace problem. (b) Solution  $u_\mu$  of Example 2.1 with  $\mu_1 = 0.1$ ,  $\mu_2 = 36$ , and random boundary conditions.

approximation space in the sense of Kolmogorov [38] in the sense that they minimize the approximation error among all spaces of the same dimension.

It has been shown that for general linear elliptic problems these optimal local approximation spaces are spanned by the left singular vectors of a linear compact operator that acts on the space of all local solutions of the PDE [3, 65, 41, 55]. In addition, it has been demonstrated in [12] that by solving the PDE with random boundary conditions on the boundary of a so-called oversampling domain that is strictly larger than the target subdomain for which we wish to construct our reduced space, and restricting the respective solutions to the target subdomain, one can obtain an approximation of the optimal local approximation spaces. Remarkably, this approximation provably converges with a convergence rate that is only slightly worse than the optimal rate obtained by the optimal local approximation spaces [12, 28].

As the linear compact operator used for the construction of the optimal local approximation spaces for linear PDEs becomes a nonlinear operator for nonlinear PDEs, a direct transfer of the above methods to nonlinear PDEs is not possible. However, we observe in Figure 2(b) for nonlinear problems also that we still have a rapid decay of high frequencies in the interior of the domain. We may thus hope that by prescribing random boundary conditions, the higher frequencies of these boundary conditions decay, and that we might get a good approximation of the functions relevant for approximation in the interior for nonlinear PDEs also; we confirm this in the numerical experiments in section 7. However, one also observes that the precise choice of the random boundary conditions for nonlinear problems is much more challenging than for linear problems as, e.g., for our considered problem here rapidly changing values of the solution lead to a rapidly changing diffusion. Motivated by the observation that higher frequencies seem to decay faster (cf. Figure 2(a)), we propose in subsection 5.2 to consider random boundary conditions such that the prescribed boundary datum interpolates with high probability a function of given smoothness on the boundary of the oversampling domain. We observe in the numerical experiments in Figure 9 that prescribing a higher smoothness might make enlarging the oversampling domain, and thus spending more computational resources, superfluous.

For linear elliptic PDEs solving the PDE for each basis function of the underlying HF discretization on the boundary of the oversampling domain and restricting these solutions to the target subdomain allows one to construct an approximation of all local solutions of the PDE of any desired accuracy. However, for nonlinear PDEs the main challenge is that the set of all local solutions of the PDE on the target subdomain is no longer a linear vector space. We thus have to ensure that our chosen

boundary conditions that we use to construct our local reduced spaces are a good representation of all boundary conditions that yield functions that are relevant for approximation or, in short, the training set has to be rich enough. This is challenging as we cannot assess at this point whether we have, e.g., chosen a good probability distribution. Therefore, after generating the local reduced spaces via randomized training we propose an adaptive algorithm in section 6 that iteratively (i) computes the global reduced approximation obtained using a generalized finite element method (gfem) with the current local approximation spaces; (ii) computes the local residuals on each subdomain and marks a certain percentage of subdomains with the largest residuals; (iii) solves the PDE locally on the marked subdomains prescribing the trace of the global reduced solution as boundary conditions; and (iv) enriches the local approximation spaces with a proper orthogonal decomposition (POD) basis for these solutions. The algorithm terminates if a rigorous global a posteriori error estimator based on local residuals that is derived in subsection 6.1 by means of the Brezzi–Rappaz–Raviart theory [8, 14, 15, 72] lies below the prescribed tolerance. We prove the exponential convergence of this adaptive algorithm for coercive linear problems in subsection 6.3, noting that to the best of our knowledge even the case of noncoercive linear problems is still an open problem. We demonstrate in numerical experiments a nearly exponential convergence of the adaptive algorithm for the nonlinear diffusion problem.

**4. CB-ROM.** In the following, we devise a CB-ROM for (2.2); we refer to Appendix A for a complete overview of the notation. To that end, we introduce the overlapping partition, which we will use in subsection 4.1 to define the partition of unity

$$(4.1) \quad \{\omega_i\}_{i=1}^{N_{\text{dd}}}, \quad \omega_i = \left\{ x \in \Omega : \min_{y \in \Omega_i} \|x - y\|_\infty < \delta_{\text{over}} \right\}, \quad i = 1, \dots, N_{\text{dd}},$$

where,  $\delta_{\text{over}}$  is the size of the overlap. Note that  $\bigcup_i \omega_i = \Omega$ .

As motivated in section 3, we aim at generating local reduced models on each subdomain  $\omega_i$ ,  $i = 1, \dots, N_{\text{dd}}$ . Thanks to the assumptions in section 2, we may choose to construct one reduced model for components in the interior, components in the corner, or at the boundary, respectively; see Figure 1(a). Therefore, we introduce the following archetype components: the “corner” (co) component is associated with the corner elements of the partition  $\{\omega_i\}_i$ ; the “edge” (ed) component is associated with the edge elements of  $\{\omega_i\}_i$ ; the “internal” (int) component is associated with the internal elements of  $\{\omega_i\}_i$  (see Figure 1(a)).

We denote by  $\widehat{\Omega}^{\text{co}}, \widehat{\Omega}^{\text{int}}, \widehat{\Omega}^{\text{ed}}$  the spatial domains associated with the three archetypal components. For edge and corner components, we denote by  $\widehat{\Gamma}_{\text{dir}}^{\text{ed}}, \widehat{\Gamma}_{\text{dir}}^{\text{co}}$  the Dirichlet boundaries; furthermore, we introduce the local HF spaces associated with the underlying HF discretization

$$(4.2) \quad \mathcal{Y}^{\text{int}} \subset H^1(\widehat{\Omega}^{\text{int}}), \quad \mathcal{Y}^{\text{ed}} \subset H_{0, \widehat{\Gamma}_{\text{dir}}^{\text{ed}}}^1(\widehat{\Omega}^{\text{ed}}), \quad \mathcal{Y}^{\text{co}} \subset H_{0, \widehat{\Gamma}_{\text{dir}}^{\text{co}}}^1(\widehat{\Omega}^{\text{co}})$$

and the corresponding (semi)norms  $\|\cdot\|_\bullet$  with  $\bullet \in \{\text{co}, \text{ed}, \text{int}\}$  that will be introduced in (4.9). We denote by  $\mathcal{L} = \{\text{co}, \text{ed}, \text{int}\}$  the library of archetype components, and by  $L : \{1, \dots, N_{\text{dd}}\} \rightarrow \mathcal{L}$  the function that associates each element of the partition  $\{\omega_i\}_i$  to the corresponding label; we further denote by  $\Phi_i : \widehat{\Omega}^{L_i} \rightarrow \omega_i$  the mapping from the (appropriate) component to the  $i$ th element of the partition. We remark that the mappings  $\Phi_i$  are simple translations for all internal components, while they are

the composition of a rigid translation and a rotation for boundary (edge and corner) components to ensure that  $\Phi_i(\widehat{\Gamma}_{\text{dir}}^{L_i}) \subset \partial\Omega$  and thus to facilitate the imposition of Dirichlet conditions.

**4.1. Partition of unity method for localized model reduction.** In [4, 46], Babuška and Melenk proposed the PUM to construct ansatz spaces with local properties. As discussed in [4], the PUM is designed to include a priori knowledge about the PDE in the ansatz spaces, and it provides a systematic way to construct ansatz spaces of any desired regularity. In the framework of CB-pMOR, the PUM provides a systematic framework to construct global ansatz spaces and offers strong theoretical guarantees concerning approximation and robustness.

Given the overlapping cover of  $\Omega$ ,  $\{\omega_i\}_{i=1}^{N_{\text{dd}}}$ , we denote by  $M$  the minimum constant such that

$$(4.3a) \quad \forall x \in \Omega, \quad \text{card} \{i \in \{1, \dots, N_{\text{dd}}\} : x \in \omega_i\} \leq M,$$

where  $\text{card}(A)$  denotes the cardinality of the discrete set  $A$ . Then, we define the partition of unity (PoU)  $\{\phi_i\}_{i=1}^{N_{\text{dd}}}$  such that

$$(4.3b) \quad \begin{cases} \text{supp}(\phi_i) \subset \bar{\omega}_i, \quad 0 \leq \phi_i(x) \leq 1, \quad \|\nabla \phi_i\|_{L^\infty(\Omega)} \leq C_i, \\ \sum_{j=1}^{N_{\text{dd}}} \phi_j(x) = 1, \quad x \in \Omega, \quad i = 1, \dots, N_{\text{dd}}. \end{cases}$$

We say that  $\{\phi_i\}_{i=1}^{N_{\text{dd}}}$  is of degree  $m$  if  $\{\phi_i\}_{i=1}^{N_{\text{dd}}} \subset C^m(\Omega; \mathbb{R})$ . Then, we define the PUM spaces

$$(4.4) \quad \mathcal{X}_{\text{pum}} := \left\{ \sum_{i=1}^{N_{\text{dd}}} \phi_i \psi_i : \psi_i \in \mathcal{X}_i \right\} \subset H_0^1(\Omega),$$

where  $\mathcal{X}_i = \{\zeta \circ \Phi_i^{-1} : \zeta \in \mathcal{Y}^i\}$ . Note that by construction  $\phi_i \zeta \circ \Phi_i^{-1} \in H_0^1(\omega_i)$  and can thus be trivially extended to  $\mathbb{R}^d$ . Next, given the reduced spaces  $\{\mathcal{Z}^\bullet\}_{\bullet \in \mathfrak{L}}$  such that  $\mathcal{Z}^\bullet \subset \mathcal{Y}^\bullet$ , we define the global reduced space

$$(4.5) \quad \mathcal{Z}_{\text{gfem}} := \left\{ \sum_{i=1}^{N_{\text{dd}}} \phi_i \zeta_i \circ \Phi_i^{-1} : \zeta_i \in \mathcal{Z}^{L_i} \right\} \subset \mathcal{X}_{\text{pum}}.$$

Theorem 4.1 provides a rigorous upper bound for the approximation properties of the PUM space in  $\Omega$ —the local approximation condition (4.6) provides the foundations for the localized data compression strategy proposed in section 5.

**THEOREM 4.1** (see [4, Theorem 1]). *Let  $u \in H_0^1(\Omega)$ . Assume that there exist  $\zeta_1, \dots, \zeta_{N_{\text{dd}}}$  such that  $\zeta_i \circ \Phi_i \in \mathcal{Z}^{L_i}$  and*

$$(4.6) \quad \|u - \zeta_i\|_{L^2(\Omega \cap \omega_i)} \leq \epsilon_i, \quad \|\nabla u - \nabla \zeta_i\|_{L^2(\Omega \cap \omega_i)} \leq \epsilon_{\nabla, i}, \quad i = 1, \dots, N_{\text{dd}},$$

*for some positive constants  $\{\epsilon_i\}_i$  and  $\{\epsilon_{\nabla, i}\}_i$ . Then, the function  $u_{\text{gfem}} = \sum_{i=1}^{N_{\text{dd}}} \phi_i \zeta_i \in \mathcal{Z}_{\text{gfem}}$  satisfies*

$$(4.7) \quad \begin{cases} \|u - u_{\text{gfem}}\|_{L^2(\Omega)} \leq \sqrt{M} \sqrt{\sum_{i=1}^{N_{\text{dd}}} \epsilon_i^2}, \\ \|\nabla u - \nabla u_{\text{gfem}}\|_{L^2(\Omega)} \leq \sqrt{2M} \sqrt{\sum_{i=1}^{N_{\text{dd}}} C_i^2 \epsilon_i^2 + \epsilon_{\nabla, i}^2}. \end{cases}$$

**4.2. Discrete variational formulation and functional norms.** We introduce the functions  $\{\widehat{\phi}^\bullet\}_\bullet$  such that  $0 \leq \widehat{\phi}^\bullet(x) \leq 1$  in  $\mathbb{R}^d$ ,  $\widehat{\phi}^\bullet(x) = 0$  if  $x \notin \widehat{\Omega}^\bullet$ ,  $\|\nabla \widehat{\phi}^\bullet\|_{L^2(\mathbb{R}^2)} \leq C_\bullet$ . We observe that, if we define the functions  $\phi_i : \mathbb{R}^d \rightarrow \mathbb{R}_+$  satisfying  $\widetilde{\phi}_i|_{\omega_i} = \widehat{\phi}^{L_i} \circ \Phi_j^{-1}$  and  $\widetilde{\phi}_i|_{\mathbb{R}^d \setminus \omega_i} \equiv 0$ , we can show that the set  $\{\phi_i\}_i$  such that

$$(4.8) \quad \phi_i = \frac{1}{\sum_{j \in \text{Neigh}(i)} \widetilde{\phi}_j} \widetilde{\phi}_i \quad \forall i = 1, \dots, N_{\text{dd}},$$

is a PoU subordinate to the cover  $\{\omega_i\}_i$ . We introduce the local seminorms

$$(4.9) \quad \|w\|_\bullet = \|\widehat{\phi}^\bullet w\|_{H^1(\widehat{\Omega}^\bullet)}, \quad \bullet \in \mathfrak{L}.$$

Note that for this choice of the local norms, since the mappings  $\{\Phi_i\}_i$  are rototranslations, if  $\{\zeta_j^\bullet\}_{j=1}^n$  are orthonormal bases with respect to  $\|\cdot\|_\bullet$ , then  $\{\phi_i \zeta_j^{L_i}\}_{j=1}^n$  is orthonormal in  $H^1(\omega_i)$  for  $i = 1, \dots, N_{\text{dd}}$ . Given the spaces  $\mathcal{X}_{i,0} := \{\phi_i \zeta \circ \Phi_i^{-1} : \zeta \in \mathcal{U}^i\}$  for  $i = 1, \dots, N_{\text{dd}}$ , we further introduce the inner products and induced norms

$$(4.10) \quad (w, v)_{1, \omega_i} = \int_{\omega_i} \nabla w \cdot \nabla v + wv \, dx, \quad \|w\|_{1, \omega_i} = \sqrt{(w, w)_{1, \omega_i}}, \quad w, v \in \mathcal{X}_{i,0};$$

the global norm  $\|w\|_{1, \Omega} = \sqrt{\int_{\Omega} \|\nabla w\|_2^2 + w^2 \, dx}$ ; and the dual norms

$$(4.11) \quad \|f\|_{-1, \omega_i} = \sup_{v \in \mathcal{X}_{i,0}} \frac{f(v)}{\|v\|_{1, \omega_i}}, \quad \|F\|_{-1, \Omega} = \sup_{v \in \mathcal{X}_{\text{pum}}} \frac{f(v)}{\|v\|_{1, \Omega}}, \quad i = 1, \dots, N_{\text{dd}},$$

for  $f \in \mathcal{X}'_{i,0}$  and  $F \in \mathcal{X}'_{\text{pum}}$ .

Then, we introduce the HF problem: given  $\mu \in \mathcal{P}_{\text{glo}}$ , find  $u_\mu \in \mathcal{X}_{\text{pum}}$  such that

$$(4.12a) \quad \mathfrak{R}_\mu(u_\mu, v) = 0 \quad \forall v \in \mathcal{X}_{\text{pum}},$$

where

$$(4.12b) \quad \mathfrak{R}_\mu(w, v) := \int_{\Omega} \eta_\mu(x; w, v) \, dx \quad \text{with} \quad \eta_\mu(x; w, v) = \kappa_\mu(x; w) \nabla w \cdot \nabla v - f_\mu v$$

and  $\mathfrak{R}_\mu : \mathcal{X}_{\text{pum}} \rightarrow \mathcal{X}'_{\text{pum}}$ .

**4.3. Residual assembly and algebraic formulation of the reduced order model.** We omit dependence of  $\Omega$  and  $\mathcal{P}_{\text{glo}}$  on  $n_{\text{dd}}$  to shorten notation. We consider the Galerkin ROM

$$(4.13) \quad \text{find } \widehat{u}_\mu \in \mathcal{Z}_{\text{gfem}} : \mathfrak{R}_\mu(\widehat{u}_\mu, v) = 0 \quad \forall v \in \mathcal{Z}_{\text{gfem}}.$$

Given the local approximation spaces  $\{\mathcal{Z}^\bullet\}_{\bullet \in \mathfrak{L}}$  with bases<sup>2</sup>  $\{\zeta_i^\bullet\}_{i=1}^n$ , we define the basis of  $\mathcal{Z}_{\text{gfem}}$   $\{\zeta_{i,j}\}_{i,j}$  such that

$$(4.14a) \quad \zeta_{i,j} = \zeta_i^{L_j} \circ \Phi_j^{-1} \phi_j, \quad i = 1, \dots, n, \quad j = 1, \dots, N_{\text{dd}}.$$

Given  $u \in \mathcal{Z}_{\text{gfem}}$ , we set  $N := nN_{\text{dd}}$  and we denote by  $\mathbf{u} \in \mathbb{R}^N$  the vector of coefficients such that

$$(4.14b) \quad u = \sum_{j=1}^{N_{\text{dd}}} \sum_{i=1}^n (\mathbf{u})_{i+(j-1)n} \zeta_{i,j}.$$

<sup>2</sup>Here, we choose  $n = \dim(\mathcal{Z}^{\text{int}}) = \dim(\mathcal{Z}^{\text{ed}}) = \dim(\mathcal{Z}^{\text{co}})$ . This choice simplifies notation and is also convenient for code vectorization. The extension to reduced spaces of arbitrary size is straightforward.

Then, we introduce the discrete residual  $\mathbf{R} : \mathbb{R}^N \times \mathcal{P}_{\text{glo}} \rightarrow \mathbb{R}^N$  such that

$$(4.15a) \quad (\mathbf{R}_\mu(\mathbf{u}))_{i+(j-1)n} = \mathfrak{R}_\mu(u, \zeta_{i,j})$$

and the algebraic nonlinear problem associated with (4.13),

$$(4.15b) \quad \text{find } \hat{\mathbf{u}}_\mu \in \mathbb{R}^N \text{ such that } \mathbf{R}_\mu(\hat{\mathbf{u}}_\mu) = \mathbf{0}.$$

In order to discuss the practical evaluation of the discrete residual  $\mathbf{R}_\mu$  in (4.15a), we define  $\text{Neigh}_i = \{j : \omega_i \cap \omega_j \neq \emptyset\}$ . Then, we observe that

$$(4.16a) \quad \begin{aligned} \mathfrak{R}_\mu(\hat{u}_\mu, \zeta_{i,j}) &= \int_{\omega_i} \eta_\mu(x; \hat{u}_\mu|_{\omega_i}, \zeta_{i,j}) dx \\ &= \int_{\hat{\Omega}^{L_i}} \hat{\eta}_\mu^{(i)}(x; (\hat{u}_\mu|_{\omega_i}) \circ \Phi_i, \zeta_i^{L_j} \hat{\phi}^{L_j}) dx, \end{aligned}$$

where

$$(4.16b) \quad \hat{\eta}_\mu^{(i)}(x; w, v) = \left( \kappa_\mu(\Phi_i(x); w) \nabla \Phi_i^{-1} \nabla \Phi_i^{-T} \nabla w \cdot \nabla v - \tilde{f}_\mu v \right) \det(\nabla \Phi_i)$$

with  $\tilde{f}_\mu = f_\mu \circ \Phi_i$ . Since  $\{\Phi_i\}_i$  are rototranslations, (4.16b) reduces to

$$(4.16c) \quad \hat{\eta}_\mu^{(i)}(x; w, v) = \kappa_\mu(\Phi_i(x); w) \nabla w \cdot \nabla v - \tilde{f}_\mu v.$$

We observe that the Jacobian  $\mathbf{J}_\mu(\cdot)$  of the algebraic residual  $\mathbf{R}_\mu(\cdot)$  is sparse for large values of  $N_{\text{dd}}$ . More precisely, exploiting (4.16) and  $\hat{u}_\mu|_{\omega_i} = \sum_{j \in \text{Neigh}_i} \sum_{i=1}^n (\hat{\mathbf{u}}_\mu)_{i+(j-1)n} \zeta_{i,j}$ , it is easy to verify that the number of nonzero elements of  $\mathbf{J}_\mu(\cdot)$  is bounded by

$$(4.17) \quad \text{nnz}(\mathbf{J}_\mu(\mathbf{u})) \leq \sum_{i=1}^{N_{\text{dd}}} n^2 \text{card}(\text{Neigh}_i) = \mathcal{O}(n^2 N_{\text{dd}}) \quad \forall \mathbf{u} \in \mathbb{R}^N.$$

For the model problem considered in this work we have  $\text{card}(\text{Neigh}_i) \leq 9$  for  $i = 1, \dots, N_{\text{dd}}$ .

Assembly of the residual in (4.16) is extremely expensive due to the need to integrate over all instantiated components  $\{\hat{\Omega}^{L_i}\}$ . To speed up computations, we should thus resort to hyperreduction techniques [6, 16, 24, 59, 73]. The choice of the hyperreduction procedure strongly depends on the PDE model of interest, on the underlying HF numerical scheme, and on the geometrical parameterization; we refer to [69] for a discussion on the treatment of geometry parameterizations. We further observe that evaluation of (4.16a) involves evaluation of  $\hat{u}_\mu$  in the mapped quadrature points of the mesh  $\hat{\Omega}^{L_i}$ ; this evaluation is extremely expensive for unstructured meshes and thus requires a specialized treatment. The development of specialized hyperreduction techniques for CB-pMOR is part of ongoing research and is not addressed in the present work.

**5. Data compression: Randomized localized training.** The aim of this section is to devise an actionable procedure to build the local approximation spaces  $\mathcal{Z}^\bullet \subset \mathcal{Y}^\bullet$  for  $\bullet \in \mathcal{L}$  such that

$$(5.1) \quad \min_{\zeta \in \mathcal{Z}^{L_i}} \|u_\mu|_{\omega_i} - \zeta \circ \Phi_i^{-1}\|_{1, \omega_i} \leq \varepsilon_{\text{tol}} \quad \text{for } i = 1, \dots, N_{\text{dd}}, \quad \mu \in \mathcal{P}_{\text{glo}}(n_{\text{dd}}),$$

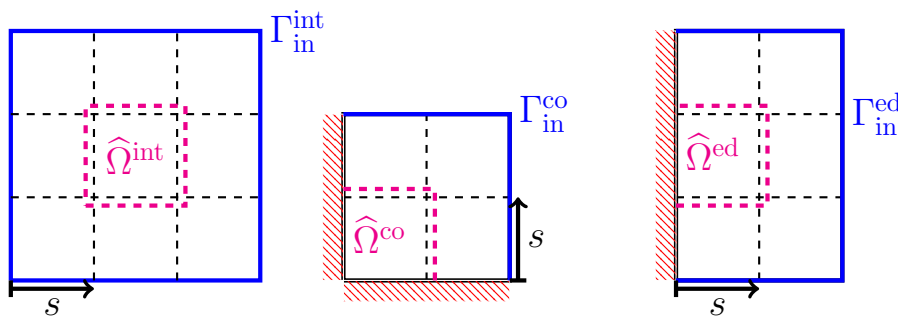


FIG. 3. Nonlinear diffusion. Archetype components with corresponding oversampling domain.

where  $\varepsilon_{\text{tol}} > 0$  is a prescribed tolerance. Condition (5.1) implies that the local spaces  $\mathcal{Z}^{\text{int}}, \mathcal{Z}^{\text{co}}, \mathcal{Z}^{\text{ed}}$  should approximate the manifolds

$$(5.2) \quad \mathcal{M}^\bullet = \left\{ u_\mu|_{\omega_i} \circ \Phi_i : \mathbf{L}_i = \bullet, \mu \in \mathcal{P}_{\text{glo}}(n_{\text{dd}}), n_{\text{dd}} \in \mathbb{N} \right\} \subset \mathcal{Y}^\bullet$$

for  $\bullet \in \mathcal{L}$ . The computation of snapshots that belong to the manifolds  $\{\mathcal{M}^\bullet\}_\bullet$  requires solving global problems and is thus unfeasible in our framework. Instead, in subsection 5.1, we propose to rely on oversampling to identify an actionable localized manifold  $\widetilde{\mathcal{M}}$  for which we can compute snapshots; then, in subsection 5.2, we propose a randomized training algorithm to construct local approximation spaces.

**5.1. Oversampling.** We fix  $i \in \{1, \dots, N_{\text{dd}}\}$  such that  $\mathbf{L}_i = \bullet$ , and we define the patch  $\widehat{\Omega}_{\text{ovr}}^\bullet \subset \mathbb{R}^2$  with input boundary  $\widehat{\Gamma}_{\text{in}}^\bullet \subset \partial\widehat{\Omega}_{\text{ovr}}^\bullet$ . We extend the mapping  $\Phi_i$  to  $\widehat{\Omega}_{\text{ovr}}^\bullet$  and we define  $\Omega_{\text{ovr},i} := \Phi_i(\widehat{\Omega}_{\text{ovr}}^\bullet)$ —for the considered model problem, the mappings  $\{\Phi_i\}_i$  are linear maps that can be trivially extended to  $\mathbb{R}^2$ . As depicted in Figure 3, we consider  $\Omega_{\text{ovr},i} = \bigcup_{j \in \text{Neigh}_i} \Omega_j$ , where  $\text{Neigh}_i = \{j : \omega_i \cap \omega_j \neq \emptyset\}$ .

We denote by  $u_{i,\mu}$  the restriction of the solution  $u_\mu$  to  $\Omega_{\text{ovr},i}$  and we define  $\widetilde{u}_{i,\mu} := u_{i,\mu} \circ \Phi_i$ . We observe that  $\widetilde{u}_{i,\mu}$  solves the problem (cf. (4.16c)):

$$(5.3a) \quad \int_{\widehat{\Omega}_{\text{ovr}}^\bullet} \kappa_\mu(\widetilde{u}_{i,\mu}) \nabla \widetilde{u}_{i,\mu} \cdot \nabla v \, dx = \int_{\widehat{\Omega}_{\text{ovr}}^\bullet} f_\mu v \, dx \quad \forall v \in \mathcal{Y}_{i,0}^{\text{ovr}}$$

with  $\widetilde{u}_{i,\mu}|_{\widehat{\Gamma}_{\text{in}}^\bullet} = u_{i,\mu} \circ \Phi_i$  and  $\mathcal{Y}_{i,0}^{\text{ovr}} = \{v \circ \Phi_i : v|_{\Omega_{\text{ovr},i}} \in \mathcal{X}_{\text{pvm}}, v|_{\partial\Omega_{\text{ovr},i}} = 0\} \subset H_0^1(\widehat{\Omega}_{\text{ovr}}^\bullet)$ . Therefore,  $u_{i,\mu}$  is a function of the subset of parameters that are active in  $\Omega_{\text{ovr},i}$  and of the boundary datum. We denote by  $\mathcal{P}^\bullet$  the active parameters and by  $\mathcal{G}^\bullet \subset H^{1/2}(\widehat{\Gamma}_{\text{in}}^\bullet)$  the domain for which (5.3a) is well-posed. Then, we introduce the transfer operator  $T^\bullet : \mathcal{G}^\bullet \times \mathcal{P}^\bullet \rightarrow \mathcal{Y}^\bullet$  such that  $T_\mu^\bullet(g) = u|_{\widehat{\Omega}^\bullet}$  where  $u$  satisfies (5.3a) with  $u|_{\partial\widehat{\Omega}_{\text{ovr}}^\bullet \setminus \widehat{\Gamma}_{\text{in}}^\bullet} = 0$  and  $u|_{\widehat{\Gamma}_{\text{in}}^\bullet} = g$ .

To provide a concrete reference, for the model problem considered in the numerical investigations, we find  $\mathcal{P}^\bullet = \bigotimes_{i=1}^{N_{\text{dd}}} \widehat{\mathcal{P}} \times \{1, \dots, N_{\text{dd}}^\bullet, 0\}$ , where  $N_{\text{dd}}^\bullet = \text{card}(\text{Neigh}_i)$  and  $i^* = 0$  means that the source term is outside the patch. Note that the parameterization  $\mathcal{P}^\bullet$  is associated with the archetype component of interest and is independent of the size of the system (i.e., the number of subdomains  $N_{\text{dd}}$ ).

We define the (unknown) set  $\mathcal{G}^{\text{true},\bullet} \subset H^{1/2}(\widehat{\Gamma}_{\text{in}}^\bullet)$  that contains all possible restrictions of the solution field to the input boundary for all instantiated com-

ponents  $\omega_i$  of type  $\bullet$ , all parameters, and all choices of  $n_{\text{dd}}$ ; clearly, we have  $\mathcal{M}^\bullet = \{T_\mu^\bullet(g) : g \in \mathcal{G}^{\text{true}, \bullet}, \mu \in \mathcal{P}^\bullet\}$ . If we introduce the “approximation”  $\tilde{\mathcal{G}}^\bullet$  of  $\mathcal{G}^{\text{true}, \bullet}$ , we obtain the localized manifold

$$(5.3b) \quad \tilde{\mathcal{M}}^\bullet = \left\{ T_\mu^\bullet(g) : g \in \tilde{\mathcal{G}}^\bullet, \mu \in \mathcal{P}^\bullet \right\}.$$

We observe that snapshots of  $\tilde{\mathcal{M}}^\bullet$  can be computed by solving local problems in the patch  $\hat{\Omega}_{\text{ovr}}^\bullet$  for prescribed choices of the active parameters  $\mu \in \mathcal{P}^\bullet$  and the boundary conditions. The patch  $\hat{\Omega}_{\text{ovr}}^\bullet$  should be significantly smaller than  $\Omega$  to ensure rapid computations; at the same time,  $\hat{\Omega}_{\text{ovr}}^\bullet$  should be large enough to ensure decay of high-frequency modes on  $\hat{\Gamma}_{\text{in}}^\bullet$ . Recently it was shown in [41] that for linear diffusion problems the local approximation error decays exponentially in the distance between  $\hat{\Omega}_{\text{ovr}}^\bullet$  and the target subdomain. Here, we investigate the effect of the size of the oversampling domain numerically in section 7 to provide some guidance on the choice of the oversampling size for the nonlinear diffusion problem of section 2.

The choice of the set of boundary conditions  $\tilde{\mathcal{G}}^\bullet$  is of paramount importance; clearly,  $\tilde{\mathcal{G}}^\bullet$  should be rich enough to ensure that  $\sup_{w \in \mathcal{M}^\bullet} \text{dist}(w, \tilde{\mathcal{M}}^\bullet) \leq \varepsilon_{\text{tol}}$ . Since the problem is nonlinear, generating a discrete representative approximation of the high-dimensional set  $\tilde{\mathcal{G}}^\bullet$  is also particularly challenging. In the next section, we directly prescribe probability density functions (pdfs)  $p_{\text{bc}}^\bullet$  of the space of boundary conditions for all  $\bullet \in \mathcal{L}$ : the set  $\tilde{\mathcal{G}}^\bullet$  is thus defined as the support of the pdf  $p_{\text{bc}}^\bullet$ .

**5.2. Randomized training.** Algorithm 5.1 illustrates the randomized training procedure. The algorithm reads as a randomized POD [74] with respect to parameter and boundary conditions: the inputs of the algorithm are the number of training points  $n_{\text{train}}$ , the size of the sought reduced spaces  $n$ , the inner product  $(\cdot, \cdot)_\bullet$ , and the pdfs  $\{p_\mu^\bullet, p_{\text{bc}}^\bullet\}$  for the archetype component  $\bullet \in \mathcal{L}$ ; the output is the reduced space  $\mathcal{Z}^\bullet$ . Here, the notation

$$\mathcal{Z} = \text{POD} \left( \{u^{(i)}\}_{i=1}^{n_{\text{train}}}, (\cdot, \cdot)_\bullet, n \right)$$

means that  $\mathcal{Z}$  is the  $n$ -dimensional POD space associated with the snapshot set  $\{u^{(i)}\}_{i=1}^{n_{\text{train}}}$  and the inner product  $(\cdot, \cdot)_\bullet$ .

It is well known that the POD is optimal in  $L^2(p_\mu^\bullet \times p_{\text{bc}}^\bullet)$  in the limit  $n_{\text{train}} \rightarrow \infty$ ; however, since the pdfs  $p_\mu^\bullet, p_{\text{bc}}^\bullet$  are chosen a priori, they might not be representative of the true distributions for the global systems. Provided that additional information on the class of global systems of interest is available, these observations motivate the enrichment strategy proposed in section 6.

---

**Algorithm 5.1** Randomized localized training.

---

*Inputs:*  $n_{\text{train}}$  size of training set,  $n$  size of the ROB,  $(\cdot, \cdot)_\bullet$  inner product,  $p_\mu^\bullet, p_{\text{bc}}^\bullet$  pdfs for the component  $\bullet \in \mathcal{L}$ .

*Output:*  $\mathcal{Z}^\bullet$  local approximation space for the component  $\bullet \in \mathcal{L}$ .

- 1: Generate  $\mu^{(k)} \stackrel{\text{iid}}{\sim} p_\mu^\bullet, g^{(k)} \stackrel{\text{iid}}{\sim} p_{\text{bc}}^\bullet, k = 1, \dots, n_{\text{train}}$ .
  - 2: Compute  $u^k = T_{\mu^{(k)}}^\bullet(g^{(k)})$  for  $k = 1, \dots, n_{\text{train}}$ .
  - 3:  $\mathcal{Z}^\bullet = \text{POD}(\{u^k\}_{k=1}^{n_{\text{train}}}, (\cdot, \cdot)_\bullet, n)$ .
-

*Remark 5.1* (probabilistic a posteriori error estimation). Given  $\bullet \in \mathfrak{L}$ ,  $n_{\text{test}}$  additional simulations  $\{u^{(i)}\}_{i=1}^{n_{\text{test}}} \subset \widetilde{\mathcal{M}}^\bullet$ , and the space  $\mathcal{Z}^\bullet$ , we introduce the error indicator

$$(5.4) \quad \widehat{E}^\bullet := \frac{1}{n_{\text{test}}} \sum_{i=1}^{n_{\text{test}}} \frac{\|u^{(i)} - \Pi_{\mathcal{Z}^\bullet}^\bullet u^{(i)}\|_\bullet}{\|u^{(i)}\|_\bullet},$$

which measures the average relative projection error on the test set  $\{u^{(i)}\}_i$ . Here,  $\Pi_{\mathcal{Z}^\bullet}^\bullet : \mathcal{Y}^\bullet \rightarrow \mathcal{Z}^\bullet$  is the projection operator on  $\mathcal{Z}^\bullet$ . Provided that  $u^{(i)} = T_{\mu^{(i)}}^\bullet(g^{(i)})$  with  $\mu^{(i)} \stackrel{\text{iid}}{\sim} p_\mu^\bullet$  and  $g^{(i)} \stackrel{\text{iid}}{\sim} p_{\text{bc}}^\bullet$ , then (5.4) is an unbiased estimator of the expected relative projection error

$$(5.5) \quad E^\bullet := \mathbb{E}_{\mu \sim p_\mu^\bullet, g \sim p_{\text{bc}}^\bullet} \left[ \frac{\|T_\mu^\bullet(g) - \Pi_{\mathcal{Z}^\bullet}^\bullet T_\mu(g)\|_\bullet}{\|T_\mu^\bullet(g)\|_\bullet} \right].$$

Note that the error indicator provides a measure of the performance of  $\mathcal{Z}^\bullet$  for the particular choice of the sampling distribution.

**Choice of random parameters.** The oversampling domains  $\{\widehat{\Omega}_{\text{ovr}}^\bullet\}_{\bullet \in \mathfrak{L}}$  in Figure 3 contain  $N_{\text{dd}}^\bullet$  subdomains (cf. Figure 3). For the model problem considered in this work, we set

$$(5.6) \quad \begin{aligned} \mu &= \left[ \mu^{(1)}, \dots, \mu^{(N_{\text{dd}}^\bullet)}, i^\star \right], \quad \mu^{(i)} \stackrel{\text{iid}}{\sim} \text{Uniform}(\widehat{\mathcal{P}}), \\ \Pr(i^\star = t) &= \begin{cases} \frac{p_{\text{src}}}{N_{\text{dd}}^\bullet}, & t = 1, \dots, N_{\text{dd}}^\bullet, \\ 1 - p_{\text{src}}, & t = 0, \end{cases} \end{aligned}$$

where  $p_{\text{src}}$  is the probability that a source term is present in the patch. If  $N_{\text{dd}}$  is known a priori, we might set  $p_{\text{s}} = \frac{N_{\text{dd}}^\bullet}{N_{\text{dd}}}$ . In this work, however, we consider  $p_{\text{src}} = 0.5$ .

**Random boundary conditions.** As motivated in section 3 we aim at defining a random boundary datum that interpolates with high probability a function of given smoothness in order to benefit from the expected more rapid decay of higher frequencies; cf. Figure 2(b). To that end, we rely on the definition of fractional Sobolev spaces via the Fourier transform.

In detail, we introduce the curvilinear coordinate  $s \in [0, 1]$  (cf. Figure 3); then, given  $N_{\text{f}} \in \mathbb{N}$  and  $\alpha \in \mathbb{R}_+$ , we define the complex-valued random field  $\widetilde{g}$  such that

$$(5.7) \quad \widetilde{g}(s; \mathbf{c}^{\text{re}}, \mathbf{c}^{\text{im}}) = \sum_{k=0}^{N_{\text{f}}-1} \frac{c_{k+1}^{\text{re}} + \mathbf{i}c_{k+1}^{\text{im}}}{\sqrt{1 + (2\pi k)^{2\alpha}}} e^{2\pi k s \mathbf{i}}, \quad c_k^{\text{re}}, c_k^{\text{im}} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1).$$

Recalling that for any  $k, k' = 0, \dots, N_{\text{f}} - 1$  and  $\alpha \in \mathbb{N}$ , we have

$$\begin{aligned} \int_0^1 e^{2\pi k s \mathbf{i}} e^{-2\pi k' s \mathbf{i}} ds &= \delta_{k, k'}, \quad \frac{d^\alpha}{ds^\alpha} e^{2\pi k s \mathbf{i}} = (2\pi k \mathbf{i})^\alpha e^{2\pi k s \mathbf{i}}, \\ (2\pi k \mathbf{i})^\alpha (-2\pi k' \mathbf{i})^\alpha &= (4\pi^2 k k')^\alpha, \end{aligned}$$

we find that

$$\begin{aligned} \|\widetilde{g}(\cdot; \mathbf{c}^{\text{re}}, \mathbf{c}^{\text{im}})\|_{H^\alpha(0,1)}^2 &= \|\widetilde{g}(\cdot; \mathbf{c}^{\text{re}}, \mathbf{c}^{\text{im}})\|_{L^2(0,1)}^2 + \|\widetilde{g}^{(\alpha)}(\cdot; \mathbf{c}^{\text{re}}, \mathbf{c}^{\text{im}})\|_{L^2(0,1)}^2 \\ &= \sum_{k=0}^{N_{\text{f}}-1} (c_{k+1}^{\text{re}} + \mathbf{i}c_{k+1}^{\text{im}}) (c_{k+1}^{\text{re}} - \mathbf{i}c_{k+1}^{\text{im}}) = \sum_{k=1}^{N_{\text{f}}} (c_k^{\text{re}})^2 + (c_k^{\text{im}})^2. \end{aligned}$$

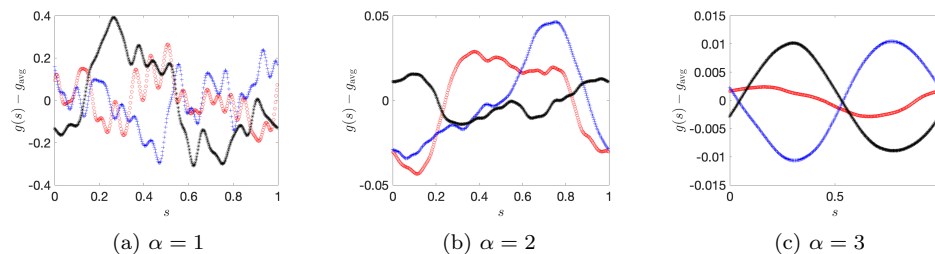


FIG. 4. Samples of random boundary conditions for 3 choices of  $\alpha$ .

---

**Algorithm 5.2** Random sample generator of boundary conditions.

---

*Inputs:*  $N_f, \alpha$  (cf. (5.7)),  $\bar{u}_{\max} \in (0, 1]$ ,  $\bullet \in \mathfrak{L}$ .

*Output:*  $g : [0, 1] \rightarrow [0, 1]$  boundary condition.

- 1: Draw  $\mathbf{c}^{\text{re}}, \mathbf{c}^{\text{im}} \in \mathbb{R}^{N_f}$  s.t.  $c_k^{\text{re}}, c_k^{\text{im}} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$ .
  - 2: Draw  $X_1, X_2, X_3 \stackrel{\text{iid}}{\sim} \text{Uniform}(0, \bar{u}_{\max})$ , set  $a = \min\{X_1, X_2\}$ ,  $b = \max\{X_1, X_2\}$ .
  - 3: Set  $g^{(1)} = \text{Real}[\tilde{g}(\cdot; \mathbf{c}^{\text{re}}, \mathbf{c}^{\text{im}})]$ .
  - 4: **if**  $\bullet = \text{int}$  **then**
  - 5:      $g = a + \frac{b-a}{\max g^{(1)} - \min g^{(1)}} (g^{(1)} - \min g^{(1)})$ .
  - 6: **else**
  - 7:      $g^{(2)}(s) = g^{(1)}(0.7s)$ .
  - 8:      $g^{(3)}(s) = \left(a + \frac{b-a}{\max g^{(2)} - \min g^{(2)}} (g^{(2)} - \min g^{(2)})\right) s(1-s)$ .
  - 9:      $g = \frac{X_3}{\max g^{(3)}} g^{(3)}$ .
  - 10: **end if**
- 

The latter implies that the random variable  $X := \|\tilde{g}(\cdot; \mathbf{c}^{\text{re}}, \mathbf{c}^{\text{im}})\|_{H^\alpha(0,1)}^2$  is distributed as a  $\chi^2$  distribution with  $2N_f$  degrees of freedom; therefore, the parameter  $\alpha$  in (5.7) controls (in a probabilistic sense) the Sobolev regularity of the datum  $\tilde{g}$ . The function in  $\tilde{\mathcal{G}}^\bullet$  that we prescribe as (random) boundary conditions then interpolates the real part of  $\tilde{g}$  times some scaling factor as discussed in the next paragraph. Thus, the prescribed boundary datum interpolates with high probability a function of given smoothness on the boundary of the oversampling domain. We illustrate this in Figure 4, which depicts samples of the random field  $g - g_{\text{avg}}$  with  $g = \text{Real}[\tilde{g}(\cdot; \mathbf{c}^{\text{re}}, \mathbf{c}^{\text{im}})]$  and  $g_{\text{avg}} = \int_0^1 g(s) ds$  for  $N_f = 20$  and increasing values of the smoothness parameter  $\alpha$ . We observe that, as  $\alpha$  increases, the samples become increasingly smooth.

In order to choose  $p_{\text{bc}}^\bullet$ , exploiting a physical argument—the solution  $u_\mu$  to (2.2) represents water saturation—we anticipate that  $u_\mu \in [0, \bar{u}_{\max}]$  for some  $\bar{u}_{\max} < 1$ . Furthermore, we wish to devise samplers that reflect the Sobolev regularity of the datum  $g$ . For these reasons, we propose to consider the procedure in Algorithm 5.2 to generate random samples of the boundary condition. We first generate a sample of the random field  $\tilde{g}$  in (5.7) and we extract its real part (cf. line 3). If  $\bullet = \text{int}$ , we simply rescale the datum to ensure that the image of  $g$ ,  $\text{Im}[g]$ , is contained in  $[0, \bar{u}_{\max}]$  (cf. line 5); if  $\bullet \in \{\text{co}, \text{ed}\}$ , since by construction  $\frac{d^k}{ds^k} g^{(1)}(0) = \frac{d^k}{ds^k} g^{(1)}(1)$  for  $k \in \mathbb{N}$ , we

define  $g^{(2)}(s) = g^{(1)}(cs)$  with<sup>3</sup>  $c = 0.7$  (cf. line 7); then, we enforce that  $g(s) = 0$  for  $s \in \{0, 1\}$  and  $g \geq 0$  (cf. line 8); finally, in line 9, we ensure that  $\text{Im}[g] \subset [0, \bar{u}_{\max}]$ .

In the numerical experiments, we provide samples of the boundary conditions for various values of  $\alpha \in \mathbb{R}_+$  and we investigate performance for the model problem considered. In particular, we discuss the impact of the choice of  $\alpha$ . Note that the sampling strategy proposed in this section depends on several parameters— $N_f, \alpha, \bar{u}_{\max}$  in Algorithm 5.2 and  $p_{\text{src}}$  in (5.6)—that might be difficult to tune. This observation justifies the use of a few global reduced solves at the training stage to improve performance of the CB-ROM.

**6. Basis enrichment based on reduced global solves.** In several contexts, it is possible to identify at the training stage a class of global configurations of interest. To provide a concrete reference for the model problem of section 2, we might be interested in solving the global PDE for (i) any choice of  $n_{\text{dd}} \in \{n_{\text{dd},\text{LB}}, \dots, n_{\text{dd},\text{UB}}\}$  with  $n_{\text{dd},\text{LB}}, n_{\text{dd},\text{UB}} \in \mathbb{N}$ , (ii) any  $\mu^{(i)} \in \widehat{\mathcal{P}}$ , (iii) up to  $n_{\text{src}}$  distinct sources. The aim of this section is to devise a localized training procedure with adaptive global enrichment that exploits prior knowledge about the global system to enrich the local spaces. In subsection 6.1, we present a residual-based error estimator that will be used to drive the enrichment strategy; in subsection 6.2, we present the training procedure; in subsection 6.3, we present an a priori convergence result for linear coercive problems. As in section 5, we assume that the system is described by a single archetype component to shorten notation.

**6.1. Residual-based error estimation.** Exploiting notation introduced in subsection 4.3, given  $i \in \{1, \dots, N_{\text{dd}}\}$ , and  $u \in H^1(\Omega)$ , we define the local Riesz elements  $\psi_\mu[u] \in \mathcal{X}_{i,0}$  as

$$(6.1a) \quad (\psi_\mu[u], v)_{1, \omega_i} = \int_{\omega_i} \widehat{\eta}_\mu^{(i)}(x; u, v) \, dx \quad \forall v \in \mathcal{X}_{i,0},$$

and the dual residual

$$(6.1b) \quad \mathbf{r}_\mu^{(i)}[u] := \|\psi_\mu[u]\|_{1, \omega_i}.$$

Next, Lemma 6.1 provides an upper bound for the global dual residual in terms of the localized dual residuals  $\{\mathbf{r}_\mu^{(i)}[\cdot]\}_i$ . The proof of Lemma 6.1 can be found in [10, Proposition 5.1], and is also provided for the sake of completeness in section SM1.

**LEMMA 6.1.** *Let  $\{\phi_i\}_i$  be a PoU that satisfies (4.3). Then, given  $u \in \mathcal{X}_{\text{pum}}$ , we have*

$$(6.2) \quad \|\mathfrak{R}_\mu(u, \cdot)\|_{-1, \Omega} \leq \sqrt{M} \left( \max_{i=1, \dots, N_{\text{dd}}} C_i^r \right) \sqrt{\sum_{i=1}^{N_{\text{dd}}} \left( \mathbf{r}_\mu^{(i)}[u] \right)^2}$$

with  $C_i^r := \sqrt{\max\{C_i + C_i^2 + 1, 2\}}$ .

We will employ the local residuals (6.1b) to mark instantiated components of the partitions where the error is large; see subsection 6.2. Let us also note that as the infinite-dimensional analogon of  $\mathfrak{R}_\mu$  as a map from  $H_0^1(\Omega)$  to  $H^{-1}(\Omega)$  is not in  $C^1$ , one cannot expect that the  $\|\cdot\|_{-1, \Omega}$ -norm of the residual (see (4.11) for the definitions)

<sup>3</sup>The choice  $c = 0.7$  is not crucial for the methodology.

stays bounded if the mesh size goes to zero. As a remedy one may consider  $\mathfrak{R}_\mu$  as a mapping from  $W_0^{1,p}(\Omega)$  to  $W^{-1,p}(\Omega)$ ,  $p > 2$ ; see [14, 56, 64]. As this significantly complicates the calculations of the dual norms, we opt here for assuming that the dimension of the HF space is fixed and consider the  $\|\cdot\|_{-1,\Omega}$ -norm. We may then define the error indicator

$$(6.3) \quad \Delta_\mu = \sqrt{\sum_{i=1}^{N_{\text{dd},\mu}} (\mathfrak{r}_\mu^i)^2}.$$

For linear problems it is straightforward to derive a rigorous a posteriori bound based on  $\mathfrak{R}_\mu[\cdot]$  (see, e.g., [4, 11]). Here, we combine Lemma 6.1 with the Brezzi–Rappaz–Raviart (BRR) theory [8, 14] to derive a rigorous residual-based error bound for the global error; see, in particular, [15, 72] for the application of the BRR theory in the context of model order reduction. To that end, if we denote by  $\mathfrak{R}'_\mu(\hat{u}_\mu)$  the Fréchet derivative of  $\mathfrak{R}_\mu$  at  $\hat{u}_\mu$ , we require that

$$(6.4) \quad 0 < \beta_{2,p} := \inf_{\substack{w \in \mathcal{X}_{\text{pum}} \\ |w|_{1,\Omega} \neq 0}} \sup_{\substack{v \in \mathcal{X}_{\text{pum}} \\ |v|_{1,\Omega} \neq 0}} \frac{\langle \mathfrak{R}'_\mu(\hat{u}_\mu)w, v \rangle}{|w|_{1,\Omega}|v|_{1,\Omega}},$$

and that there exist constants  $\gamma_{2,p}$  and  $L_{2,p}$  such that

$$(6.5) \quad \langle \mathfrak{R}'_\mu(\hat{u}_\mu)w, v \rangle \leq \gamma_{2,p}|w|_{W^{1,p}(\Omega)}|v|_{1,\Omega},$$

$$(6.6) \quad \|\mathfrak{R}'_\mu(\hat{u}_\mu) - \mathfrak{R}'_\mu(w)\| \leq L_{2,p}|\hat{u}_\mu - w|_{W^{1,p}(\Omega)}$$

for  $w \in B(\hat{u}_\mu, R) \subset \mathcal{X}_{\text{pum}}$  and  $v \in \mathcal{X}_{\text{pum}}$ . Here,  $R$  is supposed to be sufficiently large and  $|w|_{1,\Omega} := \|\nabla w\|_{L^2(\Omega)}$  and  $|w|_{W^{1,p}(\Omega)} := \|\nabla w\|_{L^p(\Omega)}$ . We note that conditions (6.4)–(6.6) are satisfied for the considered model problem (2.2) in the infinite-dimensional setting albeit potentially with different norms [62, Theorem 3.4], while the inf-sup condition (6.4) can be verified a posteriori. To obtain a proximity indicator [72, 15], which is based on localized and easily computable residuals via the Riesz representation, we employ, as in [64], the finite dimensionality of  $\mathcal{X}_{\text{pum}}$  and define  $c_h := \sup_{v \in \mathcal{X}_{\text{pum}}} (|v|_{W^{1,p}(\Omega)})/|v|_{1,\Omega}$  and

$$(6.7) \quad \tau_{\mu,p} := \frac{2L_{2,p}c_h}{\beta_{2,p}^2} \sqrt{M} \left( \max_{i=1,\dots,N_{\text{dd}}} C_i^{\text{r}} \right) \sqrt{\sum_{i=1}^{N_{\text{dd}}} (\mathfrak{r}_\mu^{(i)}[u])^2}.$$

The proximity indicator  $\tau_{\mu,p}$  will be used to validate whether  $\hat{u}_\mu$  is close enough to  $u_\mu$  within the adaptive Algorithm 6.1. We obtain the following result.

**PROPOSITION 6.2** (global a posteriori error bound). *Let  $\tau_{\mu,p} < 1$  and (6.4), (6.5), and (6.6) be fulfilled. Then there exists a unique solution  $u_\mu \in B(\hat{u}_\mu, \frac{\beta_{2,p}}{L_{2,p}c_h}) \subset \mathcal{X}_{\text{pum}}$  of (4.12) and the error estimator*

$$(6.8) \quad \Delta_{\mu,p} := \frac{\beta_{2,p}}{L_{2,p}c_h} (1 - \sqrt{1 - \tau_{\mu,p}})$$

satisfies

$$(6.9) \quad \|\hat{u}_\mu - u_\mu\|_{1,\Omega} \leq \Delta_{\mu,p}.$$

*Proof.* Lemma 6.1 and  $\tau_{\mu,p} < 1$  imply that  $\tilde{\tau}_{\mu,p} := \frac{2L_{2,p}c_h}{\beta_{2,p}^2} \|\mathfrak{R}_\mu(\hat{u}_\mu, \cdot)\|_{-1,\Omega} < 1$ . The existence of a unique solution  $u_\mu \in B(\hat{u}_\mu, \frac{\beta_{2,p}}{L_{2,p}c_h})$  of (4.12) and

$$(6.10) \quad |\hat{u}_\mu - u_\mu|_{1,\Omega} \leq \frac{\beta_{2,p}}{L_{2,p}c_h} (1 - \sqrt{1 - \tilde{\tau}_{\mu,p}})$$

then follows using standard arguments in the BRR theory (see [15, 72, 14, 56] and for this particular PDE [62]). As the function  $t(x) := 1 - \sqrt{1-x}$  is strictly increasing on  $(0,1)$ , applying Lemma 6.1 to the right side of (6.10) concludes the proof.  $\square$

*Remark 6.3* (discussion of the result). It is well known that for nonlinear PDEs the dual norm of the residual can only be used as an a posteriori error estimator if the approximation is already close to the HF solution (see, e.g., [14, 71]). Relying solely on the dual norm of the residual can therefore be problematic as it may seem that the approximation error is acceptable even though that might not be the case. The proximity indicator  $\tau_{\mu,p}$  (6.7), which only relies on computable constants, can be used to assess, whether indeed the approximation  $\hat{u}_\mu$  is close enough to  $u_\mu$  such that the error estimation (6.9) is valid. While the proximity indicator  $\tau_{\mu,p}$  (6.7) and thus the a posteriori error estimator (6.8) solely rely on the dual norms of local residuals that can be computed on the components and therefore do not require any global solutions, the constants  $L_{2,p}$  and  $\beta_{2,p}$  are global constants. We will discuss some strategies on how to estimate these constants in Remark 6.4. To the best of our knowledge even for linear elliptic PDEs there are no results in the conforming setting that solely rely on local constants (the a posteriori error estimators in [10, 63], e.g., both contain the global coercivity constant). A fully localizable a posteriori error estimator for nonlinear nonmonotone PDEs would therefore be at least a paper on its own and is thus beyond the scope of this paper.

*Remark 6.4* (estimation of constants). Regarding the estimation of the constant  $c_h$  in the inverse inequality, we refer to classical results, e.g., in [23] noting that the global inverse inequality only requires the measure of  $\Omega$ . Estimating the constant  $L_{2,p}$  relies on estimates of the constant in the Poincaré inequality for  $L^p$ ,  $W^{1,p}$  and the Sobolev embedding inequality  $\|v\|_{C^0(\Omega)} \leq c_E \|v\|_{W^{1,p}(\Omega)}$  (see, e.g., [62, subsection 3.1.2]). The estimation of  $c_E$  can be easily localized. An estimate of the constant in the Poincaré inequality involving the measure of  $\Omega$  can be found in [26, (7.44)] for functions that are zero on  $\partial\Omega$ . We hope that if the local reduced bases contain the constant function it is maybe possible to obtain localized and more precise estimates of the Poincaré constant. Finally, similarly to [63], we propose to use a localized model order reduction approximation of  $\beta_{2,p}$ . In detail, we suggest using the following heuristic and hierarchical estimator  $\beta_{2,p}^{\text{app}} := \inf_{w \in \tilde{\mathcal{Z}}_{\text{gfem}}} \sup_{v \in \tilde{\mathcal{Z}}_{\text{gfem}}} \frac{\langle \mathfrak{R}'_\mu(\hat{u}_\mu)_w, v \rangle}{|w|_{1,\Omega} |v|_{1,\Omega}}$ , where  $\tilde{\mathcal{Z}}_{\text{gfem}} \subsetneq \tilde{\mathcal{Z}}_{\text{gfem}} \subset \mathcal{X}_{\text{pum}}$ . We conjecture that using a certain number of additional local basis functions per component might already yield an acceptable estimate of  $\beta_{2,p}$ .

**6.2. Adaptive algorithm.** We introduce the pdfs  $\{p_\mu^\bullet, p_{\text{bc}}^\bullet : \bullet \in \mathfrak{L}\}$  for localized sampling and the pdf  $p_\mu^{\text{glo}}$  that is used to generate global problems. In the numerical examples, we consider  $n_{\text{dd}} \sim \text{Uniform}(\{4, \dots, 12\})$ ,  $\mu^{(i)} \stackrel{\text{iid}}{\sim} \text{Uniform}(\hat{\mathcal{P}})$ , and we assume that exactly one source term is active in  $\Omega$  (that is,  $n_{\text{src}} = 1$ ). Given the partition  $\{\omega_i\}_i$ , we define the local solution operators

$$(6.11) \quad T_\mu^{(i)} : \mathcal{X}_i \rightarrow \mathcal{X}_{i,0} \quad \text{s.t.} \quad \mathfrak{R}_\mu(u + T_\mu^{(i)}(u), v) = 0 \quad \forall v \in \mathcal{X}_{i,0}$$

---

**Algorithm 6.1** Randomized localized training with global enrichment.

---

*Inputs (localized training):*  $n_{\text{train}}^{\text{loc}}$  = number of solves,  $n^{\text{loc}}$  = size of the POD spaces,  $\{p_{\mu}^{\bullet}, p_{\text{bc}}^{\bullet}\}_{\bullet}$  sampling pdfs.

*Inputs (enrichment):*  $n_{\text{train}}^{\text{glo}}$  = number of global simulations per iteration,  $n^{\text{glo}}$  = number of modes added at each iteration, **maxit** = maximum number of outer loop iterations, *tol* = tolerance for termination criterion,  $p_{\mu}^{\text{glo}}$  = global configuration sampler,  $m_r$  = percentage of marked components at each iteration.

*Outputs:*  $\{\mathcal{Z}^{\bullet}\}_{\bullet \in \mathcal{L}}$  local approximation spaces.

**Localized training**

- 1: Apply Algorithm 5.1 to obtain the local spaces  $\{\mathcal{Z}^{\bullet}\}_{\bullet \in \mathcal{L}}$ .

**Enrichment**

- 1: Sample  $n_{\text{train}}^{\text{glo}}$  configurations  $\mu^{(k)} \stackrel{\text{iid}}{\sim} p_{\mu}^{\text{glo}}$ ,  $\mathcal{P}_{\text{train}} := \{\mu^{(k)}\}_k$
  - 2: **for**  $\ell = 1, \dots, \text{maxit}$  **do**
  - 3:   Initialize the datasets  $\mathcal{D}^{\bullet} = \emptyset$  for  $\bullet \in \{\text{co}, \text{ed}, \text{int}\}$ .
  - 4:   **for**  $\mu \in \mathcal{P}_{\text{train}}$  **do**
  - 5:     Compute  $\hat{u}_{\mu}$  using the PUM-CB-ROM (cf. section 4).
  - 6:     Compute local residuals (6.1)  $\mathbf{r}_{\mu}^i = \mathbf{r}_{\mu}^{(i)}[\hat{u}_{\mu}]$  for  $i = 1, \dots, N_{\text{dd}, \mu}$ .
  - 7:     **for**  $\bullet \in \mathcal{L}$  **do**
  - 8:       Mark the  $m_r$  % instantiated components of type  $\bullet$  with the largest residual,  $\{\omega_i\}_{i \in \mathbf{I}_{\text{mark}, \bullet}^{\mu}}$ .
  - 9:       Solve the local problems (6.11) in  $\{\omega_i\}_{i \in \mathbf{I}_{\text{mark}, \bullet}^{\mu}}$ ,  $u_{i, \mu}^{\bullet} = \frac{1}{\phi_i} T_{\mu}^{(i)}(\hat{u}_{\mu} | \omega_i)$ .
  - 10:       Augment the dataset  $\mathcal{D}^{\bullet} = \mathcal{D}^{\bullet} \cup \{u_{i, \mu}^{\bullet} \circ \Phi_i : i \in \mathbf{I}_{\text{mark}, \bullet}^{\mu}\}$ .
  - 11:     **end for**
  - 12:     Compute  $\Delta_{\mu, p}$  (6.8) with approximate constants.
  - 13:   **end for**
  - 14:   Update the POD spaces  $\mathcal{Z}^{\bullet} = \mathcal{Z}^{\bullet} \cup \text{POD}(\{w - \Pi_{\mathcal{Z}^{\bullet}} w : w \in \mathcal{D}^{\bullet}\}, (\cdot, \cdot)_{\bullet}, n^{\text{glo}})$ .
  - 15:   **if**  $\max_{\mu \in \mathcal{P}_{\text{train}}} \Delta_{\mu, p} < \text{tol}$  **then**
  - 16:     **BREAK**
  - 17:   **end if**
  - 18: **end for**
- 

for  $i = 1, \dots, N_{\text{dd}, \mu}$ . The particular choice of the operators  $\{T_{\mu}^{(i)}\}_i$  in (6.11) is motivated by the convergence analysis in subsection 6.3 (cf. line 6 of Algorithm 6.2).

Algorithm 6.1 contains the data compression procedure. First, we initialize the local spaces using Algorithm 5.1. Then, we sample  $n_{\text{train}}^{\text{glo}}$  configurations  $\mathcal{P}_{\text{train}} = \{\mu^{(k)}\}_{k=1}^{n_{\text{train}}^{\text{glo}}}$  with  $\mu^{(k)} \stackrel{\text{iid}}{\sim} p_{\mu}^{\text{glo}}$ , and we proceed with the enrichment iterations. At the  $\ell$ th iteration, for each  $\mu \in \mathcal{P}_{\text{train}}$ , we resort to the CB-ROM proposed in section 4 to estimate the solution  $\hat{u}_{\mu}$ ; we compute the local residuals (6.1)  $\mathbf{r}_{\mu}^i = \mathbf{r}_{\mu}^{(i)}[\hat{u}_{\mu}]$ ,  $i = 1, \dots, N_{\text{dd}, \mu}$ , and, for all  $\bullet \in \mathcal{L}$ , we mark the  $m_r$  % instantiated components of type  $\bullet$  with the largest residual,  $\mathbf{I}_{\text{mark}, \bullet}^{\mu} \subset \{1, \dots, N_{\text{dd}, \mu}\}$ . Then, we solve (6.11) to obtain the local fields  $u_{i, \mu}^{\bullet}$  for all  $i \in \mathbf{I}_{\text{mark}, \bullet}^{\mu}$  (cf. line 9), and we update the dataset of simulations  $\mathcal{D}^{\bullet}$  associated with the marked elements of type  $\bullet$  (cf. line 10). In view of the termination condition, we further compute the error estimator  $\Delta_{\mu, p}$  (6.8) with approximate constants. At the end of the loop over the parameters, we update the local spaces using POD (cf. subsection 5.2) (cf. line 14), and we check if  $\max_{\mu \in \mathcal{P}_{\text{train}}} \Delta_{\mu, p}$  is below a user-defined tolerance.

The solution to (6.11) is performed over the domain  $\omega_i$  (or equivalently  $\widehat{\Omega}^{L_i} \subset \widehat{\Omega}_{\text{ovr}}^{L_i}$ ), and the Newton solver can be initialized with the null solution; for this reason, it is significantly cheaper than the solution to (5.3a). We observe that the local solutions  $u_{i,\mu}^\bullet$  (cf. line 9) are not well-defined on  $\partial\omega_i$  (i.e., division of 0 by 0); however, since we are ultimately interested in the PUM space  $\mathcal{Z}_{\text{gfem}}$  (4.5) and due to the choice of the local norm  $\|\cdot\|_\bullet$  (cf. (4.9)), this issue does not affect our procedure. We further observe that several steps of the algorithm are embarrassingly parallelizable: the loop over the configurations (cf. lines 4 to 13), the computation of the residuals (cf. line 6), the solution to the local problems (cf. line 9).

We observe that the performance of Algorithm 6.1 depends on the choice of several hyperparameters and in particular on the number of modes  $n^{\text{glo}}$  added at each iteration. In the POD literature, the size of the POD space is typically chosen based on an energy criterion (see, e.g., [57, eq. (6.12)]). A thorough investigation of the sensitivity of Algorithm 6.1 with respect to  $n^{\text{glo}}$  is beyond the scope of the present paper. Choosing small values of  $n^{\text{glo}}$  requires more outer-loop iterations for any prescribed accuracy; however, since the CB-ROM is more and more accurate as the iteration count  $\ell$  increases, we envision that reducing the size of  $n^{\text{glo}}$  might lead to more accurate reduced spaces for any fixed dimension; therefore, the choice of  $n^{\text{glo}}$  ultimately reflects a trade-off between offline costs and online efficiency.

*Remark 6.5* (computational complexity of Algorithm 6.1). The computational costs of Algorithm 6.1 are dominated by the solutions of the PDE and the computations of the local residuals on the subdomains or oversampling domains. In detail, conducting the randomized local training in Algorithm 5.1 requires  $n_{\text{train}}$  solutions of the PDE on the oversampling domain for each archetype component, which can however be performed in an embarrassingly parallel manner. Similarly we can use a randomized SVD [28] to compute the POD basis that is also amenable to parallelization.

As the reduced basis functions for one subdomain  $\omega_i$  have joint support only with the reduced basis functions of very few other subdomains  $\omega_j$ , we expect that each of the  $\dim(\mathcal{P}_{\text{train}})$  solutions of the global reduced system should scale linearly in the number of subdomains  $N_{\text{dd}}$  (see also (6.1a)) and, at worst quadratically in the dimension of the local reduced spaces  $\{\mathcal{Z}^\bullet\}_{\bullet \in \mathcal{L}}$  if one resorts to a conjugate gradient method within Newton's method. The  $\dim(\mathcal{P}_{\text{train}}) \cdot N_{\text{dd},\mu}$  computations of the dual norms of the local residuals  $\mathbf{r}_\mu^{(i)}$  require the solutions of  $\dim(\mathcal{P}_{\text{train}}) \cdot N_{\text{dd},\mu}$  local linear PDEs (6.1a) to compute the corresponding Riesz representatives. As we solve in line 9 of Algorithm 6.1 the local problems only for the marked subdomains, we do not expect these computational costs to be dominating. The same applies for the POD to be performed in line 14.

**6.3. A priori convergence analysis for coercive linear problems.** We study the in-sample a priori convergence of the enrichment procedure in Algorithm 6.1; we consider the case of linear coercive problems, and we apply the simplified randomized procedure contained in Algorithm 6.2; the proof follows the argument of [9, Theorem 1]. To clarify the presentation and avoid unnecessary notation, we assume  $\mathcal{P}_{\text{train}} = \{\mu\}$ , and we omit dependency on  $\mu$ ; in the supplementary material (KSTT\_supplement.pdf [local/web 224KB]), we discuss the extension to multiple configurations. We also assume that the cardinality of the library of components is equal to one and we omit dependency on  $\bullet$ .

We define the model problem,

$$(6.12) \quad \text{find } u \in \mathcal{X} : \mathfrak{R}(u, v) = f(v) - a(u, v) = 0 \quad \forall v \in \mathcal{X},$$

---

**Algorithm 6.2** Simplified randomized localized training with global enrichment.

---

- 1: Initialize  $\mathcal{Z} = \mathcal{Z}_0$ .
  - 2: Sample  $n_{\text{train}}^{\text{glo}} = 1$  configurations  $\mu \sim p_{\mu}^{\text{glo}}$ ,  $\mathcal{P}_{\text{train}} := \{\mu\}$
  - 3: **for**  $\ell = 0, \dots, \text{maxit}$  **do**
  - 4:   Compute  $\hat{u}_{\ell}$  using the PUM-CB-ROM (cf. section 4).
  - 5:   Find  $k = \arg \max_{i=1, \dots, N_{\text{dd}}} \mathbf{r}^{(i)}[\hat{u}_{\ell}]$ .
  - 6:   Solve the local problem: find  $T^{(k)}(\hat{u}_{\ell}) \in \mathcal{X}_{k,0}$  such that  $\mathfrak{R}(\hat{u}_{\ell} + T^{(k)}(\hat{u}_{\ell}), v) = 0$  for all  $v \in \mathcal{X}_{k,0}$ .
  - 7:   Define  $u^* = \frac{1}{\phi_k} T^{(k)}(\hat{u}_{\ell})$  and update the local space  $\mathcal{Z} = \mathcal{Z} \cup \text{span}\{u^* \circ \Phi_k\}$ .
  - 8: **end for**
- 

where  $H_0^1(\Omega) \subset \mathcal{X} \subset H^1(\Omega)$  is a suitable Hilbert space on  $\Omega$ . We also introduce the energy norm and the associated dual norm:

$$(6.13) \quad \|w\|_a = \sqrt{a(w, w)} \quad \forall w \in \mathcal{X}, \quad \|f\|_{\mathcal{X}'} = \sup_{v \in \mathcal{X}} \frac{f(v)}{\|v\|_a} \quad \forall f \in \mathcal{X}'.$$

Given the partition  $\{\omega_i\}_{i=1}^{N_{\text{dd}}}$ , we further define the associated mappings  $\{\Phi_i\}_{i=1}^{N_{\text{dd}}}$ , the associated PoU  $\{\phi_i\}_{i=1}^{N_{\text{dd}}}$ , and the local spaces  $\mathcal{X}_i = H^1(\omega_i) \cap \mathcal{X}$  and  $\mathcal{X}_{i,0} = H_0^1(\omega_i)$ . Then, we define the local dual residual norms such that

$$(6.14) \quad \mathbf{r}^{(i)}[u] = \sup_{v \in \mathcal{X}_{i,0}} \frac{\mathfrak{R}(u, v)}{\|v\|_a}, \quad i = 1, \dots, N_{\text{dd}}.$$

Finally, we denote by  $c_{\text{pu}}$  the constant such that (see (6.2))

$$(6.15) \quad \|\mathfrak{R}_{\mu}(u, \cdot)\|_{\mathcal{X}'} \leq c_{\text{pu}} \sqrt{\sum_{i=1}^{N_{\text{dd}}} (\mathbf{r}^{(i)}[u])^2}.$$

Proposition 6.6 shows that the reconstruction error decreases exponentially with respect to the iteration count  $\ell$  for any choice of the initial reduced space.

**PROPOSITION 6.6.** *The sequence of PUM-CB-ROM solutions  $\{\hat{u}_{\ell}\}_{\ell=1,2,\dots}$  satisfies  $\|u - \hat{u}_{\ell}\|_a \leq \left(1 - \frac{1}{N_{\text{dd}} c_{\text{pu}}^2}\right)^{\ell/2} \|u - \hat{u}_0\|_a$ .*

Next the lemma summarizes two standard results that will be used in the proof of Proposition 6.6.

**LEMMA 6.7.** *Let  $\mathcal{Z}_{\text{gfem}} \subset \mathcal{X}$  and let  $\hat{u} \in \mathcal{Z}_{\text{gfem}}$  satisfy  $\mathfrak{R}(\hat{u}, v) = 0 \quad \forall v \in \mathcal{Z}_{\text{gfem}}$ . Then, we have*

$$(6.16a) \quad \|\hat{u} - u\|_a = \inf_{\varphi \in \mathcal{Z}_{\text{gfem}}} \|\varphi - u\|_a,$$

$$(6.16b) \quad \|\hat{u} - u\|_a = \|\mathfrak{R}(u, \cdot)\|_{\mathcal{X}'}.$$

Proof (Proposition 6.6). Exploiting (6.15) and then (6.16b), we find

$$(6.17) \quad \left(\mathbf{r}^{(k)}[\hat{u}_{\ell}]\right)^2 \geq \frac{1}{N_{\text{dd}}} \sum_{j=1}^{N_{\text{dd}}} \left(\mathbf{r}^{(j)}[\hat{u}_{\ell}]\right)^2 \geq \frac{1}{N_{\text{dd}} c_{\text{pu}}^2} \|\mathfrak{R}(\hat{u}_{\ell}, \cdot)\|_{\mathcal{X}'}^2 = \frac{1}{N_{\text{dd}} c_{\text{pu}}^2} \|u - \hat{u}_{\ell}\|_a^2.$$

By construction,  $u_k^{\text{loc}} = T^{(k)}(\hat{u}_\ell)$  in Algorithm 6.2 belongs to  $\mathcal{Z}_{\text{gfem}}^{(\ell+1)}$ . As a result, if we consider  $\varphi = \hat{u}_\ell + u_k^{\text{loc}}$  in (6.16a), we find

$$\|u - \hat{u}_{\ell+1}\|_a^2 \leq \|u - \hat{u}_\ell - u_k^{\text{loc}}\|_a^2 = \|u - \hat{u}_\ell\|_a^2 - 2a(u - \hat{u}_\ell, u_k^{\text{loc}}) + \|u_k^{\text{loc}}\|_a^2.$$

Since

$$a(u - \hat{u}_\ell, u_k^{\text{loc}}) = \underbrace{\Re(\hat{u}_\ell + u_k^{\text{loc}}, u_k^{\text{loc}})}_{=0} + \|u_k^{\text{loc}}\|_a^2 = \|u_k^{\text{loc}}\|_a^2$$

and

$$\|u_k^{\text{loc}}\|_a = \sup_{v \in \mathcal{X}_{k,0}} \frac{a(u_k^{\text{loc}}, v)}{\|v\|_a} = \sup_{v \in \mathcal{X}_{k,0}} \frac{\Re(u - \hat{u}_\ell, v)}{\|v\|_a} = \mathbf{r}^{(k)}[\hat{u}_\ell],$$

we obtain

$$(6.18) \quad \|u - \hat{u}_{\ell+1}\|_a^2 \leq \|u - \hat{u}_\ell\|_a^2 - \left(\mathbf{r}^{(k)}[\hat{u}_\ell]\right)^2.$$

By combining (6.17) and (6.18), we obtain

$$\|u - \hat{u}_{\ell+1}\|_a^2 \leq \left(1 - \frac{1}{N_{\text{dd}} c_{\text{pu}}^2}\right) \|u - \hat{u}_\ell\|_a^2,$$

which completes the proof.  $\square$

**7. Numerical results.** In subsection 7.1, we investigate performance of the randomized sampling algorithm for a linear problem; then, in subsection 7.2, we consider the nonlinear diffusion problem introduced in section 2. Numerical simulations are performed in MATLAB 2020b on a commodity laptop.

**7.1. Performance of randomized training for a linear problem.** We first provide numerical investigations for the linear advection-diffusion-reaction problem

$$\begin{cases} -\nabla \cdot (\mu_1 \kappa \nabla u_{\mu,g} + [\mu_2, \mu_3]^T u_{\mu,g}) + \mu_4 u_{\mu,g} = 0 & \text{in } \Omega_{\text{OVR}} = (0, 0.3)^2, \\ u_{\mu,g} = g & \text{on } \partial\Omega_{\text{OVR}} =: \Gamma_{\text{in}}, \end{cases}$$

where  $\kappa(x) = \frac{1}{1+\|x\|_2^2}$  and  $\mu = [\mu_1, \mu_2, \mu_3, \mu_4] \in \mathcal{P} = [0.2, 1] \times [-1, 1]^2 \times [0, 1]$ . We consider the extracted domain  $\hat{\Omega} = (0.1, 0.2)^2$ . The linear problem allows us to compare our randomized method with a previously developed data compression algorithm. Note that the transfer operator  $T : (\mu, g) \mapsto u_{\mu,g}|_{\hat{\Omega}}$  is nonlinear due to the presence of parameters. We discretize the problem using the finite element method based on cubic (P3) polynomials with  $N_{\text{in}} = 360$  degrees of freedom on the boundary  $\Gamma_{\text{in}}$ .

We compare performance of our randomized algorithm with the approach in [68] (TE + POD): given the training set  $\mathcal{P}_{\text{train}} = \{\mu^k\}_{k=1}^{n_{\text{train}}} \subset \mathcal{P}$ , we first solve  $n_{\text{train}}$  independent transfer eigenproblems [3] for each value of the parameter and then we use POD to combine the resulting spaces. We refer to [68] for further details and analysis, and we refer to [65] for a similar data compression algorithm. In the numerical experiments, we set  $n_{\text{train}} = 100$ : this implies that TE + POD is required to solve  $n_{\text{train}} \cdot N_{\text{in}} = 36000$  PDEs. We envision that the total number of PDE solves can be reduced up to  $\mathcal{O}(n \cdot n_{\text{train}})$  by resorting to Krylov methods to solve the transfer eigenproblem; we refer to the above-mentioned literature for further details.

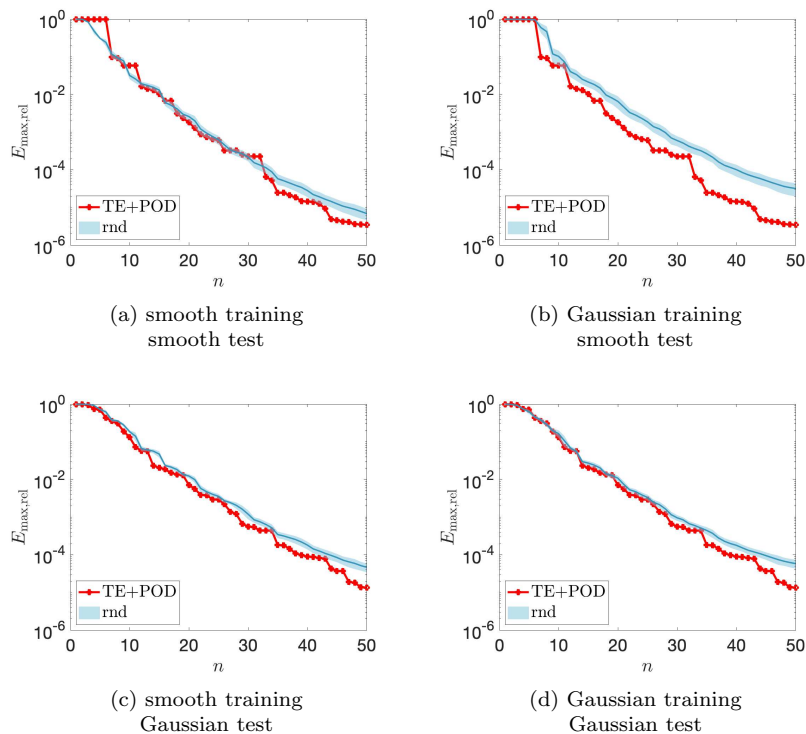


FIG. 5. *Linear problem. Out-of-sample performance; comparison with deterministic training for  $N_r = 100$  choices of the random samples and for two fixed test sets. (a)–(b) Smooth test set. (c)–(d) Gaussian test set.*

We set  $p_\mu = \text{Uniform}(\mathcal{P})$  and we consider samples of the random field  $g = \text{Real}[\tilde{g}(\cdot; \mathbf{c}^{\text{re}}, \mathbf{c}^{\text{im}})]$  (cf. (5.7)) with  $N_f = 20$ . Given the restriction of the finite element Lagrangian basis to the input boundary  $\{\phi_i^{\text{fe}}\}_{i \in \mathcal{I}_{\text{dir}}}$ , we further define the random field

$$(7.1) \quad g(x; \mathbf{c}) := \sum_{i \in \mathcal{I}_{\text{dir}}} c_i \phi_i^{\text{fe}}(x) \quad \text{with } c_i \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1),$$

which is used below for comparison. To assess performance, we compare the maximum relative projection error

$$(7.2) \quad E_{\text{max,rel}}(\mathcal{Z}) := \max_{j=1, \dots, n_{\text{test}}} \frac{\|\Pi_{\mathcal{Z}^\perp} u_{\mu^{(j)}, g^{(j)}}|_{\hat{\Omega}}\|_{H^1(\hat{\Omega})}}{\|u_{\mu^{(j)}, g^{(j)}}|_{\hat{\Omega}}\|_{H^1(\hat{\Omega})}}, \quad \mu^{(j)} \stackrel{\text{iid}}{\sim} \text{Uniform}(\mathcal{P}), \quad g^{(j)} \stackrel{\text{iid}}{\sim} p_{\text{bc}},$$

for the two choices of  $p_{\text{bc}}$ —“smooth” (with  $\alpha = 1$ ) and “Gaussian” (7.1)—and  $n_{\text{test}} = 100$ .

Figure 5 shows the results for smooth and Gaussian training and test sets. Here, we consider training sets of size  $n_{\text{train}} = 50$  in Algorithm 5.1; furthermore, we compare error bar plots based on 100 independent choices of the training set. We observe that our smooth sampling strategy is nearly as effective as TE + POD for  $n \lesssim 40$  for both smooth and Gaussian test sets. This result empirically demonstrates that randomized methods are extremely effective for identifying dominant POD modes even for nonlinear transfer operators. We further observe that Gaussian sampling

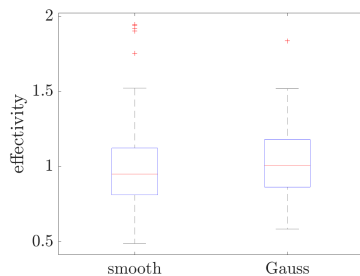


FIG. 6. *Linear problem. Effectivity of the error indicator  $\eta = \frac{\widehat{E}(n_{\text{test}}=10)}{\widehat{E}(n_{\text{test}}=100)}$  for 100 independent runs and for both Gaussian and smooth training.*

is clearly inferior when tested on smooth data, while it performs as accurately as smooth sampling on the Gaussian test set: we conjecture that this behavior is due to the low-pass filtering properties of the differential operator.

In Figure 6, we show the behavior of the error indicator  $\widehat{E}$  in Remark 5.1; more precisely, we show boxplots of the approximate effectivity  $\eta = \frac{\widehat{E}(n_{\text{test}}=10)}{\widehat{E}(n_{\text{test}}=100)}$  for 100 independent runs and for both Gaussian and smooth training. Note that, with very high probability,  $\eta \in [0.5, 1.5]$ . Note also, however, that  $\widehat{E}$  strongly depends on the choice of the sampling distribution, which in practice is largely unknown.

**7.2. Application to the nonlinear diffusion problem.** We consider the application to the nonlinear diffusion problem introduced in section 2. We discretize the problem using a Q3 spectral element method based on a structured grid with 961 degrees of freedom in each subdomain  $\Omega_i$ . We apply Algorithm 5.2 with  $n_{\text{train}} = 200$ ; we set  $\{p_\mu^\bullet\}_\bullet$  as discussed in subsection 5.2 and we consider the smooth sampler described in Algorithm 5.2 for  $N_f = 20$  and various choices of  $\alpha$  and  $\bar{u}_{\text{max}}$ . We further compare performance with randomized training based on the random field

$$g^\bullet(x; \mathbf{c}) := \sum_{i \in \mathbf{I}_{\text{dir}}^\bullet} f(c_i, \bar{u}_{\text{max}}) \phi_i^{\text{fe}, \bullet}(x), \quad c_i \stackrel{\text{iid}}{\sim} \mathcal{N}\left(\frac{\bar{u}_{\text{max}}}{2}, \frac{\bar{u}_{\text{max}}^2}{4}\right),$$

$$(7.3) \quad f(c, u) = \max\{\min\{c, u\}, 0\},$$

where  $\{\mathbf{I}_{\text{dir}}^\bullet\}_\bullet$  denotes the set of indices of the mesh on the patch input boundaries and  $\{\phi_i^{\text{fe}, \bullet}\}_\bullet$  are the Lagrangian bases associated with the HF discretization. We refer to the sampling procedure in Algorithm 5.2 as *smooth sampling*; we refer to the sampling procedure based on (7.3) as *Gaussian sampling*. Nonlinear systems are solved using a standard Newton's method with line search.

We consider a piecewise tensorized bilinear PoU  $\{\phi_{i+(j-1)n_{\text{dd}}}(x) = \phi_i^{1\text{d}}(x_1) \phi_j^{1\text{d}}(x_2)\}_{i,j=1}^{n_{\text{dd}}}$ , where  $\{\phi_i^{1\text{d}}\}_i$  is a PoU subordinate to the cover

$$\{\omega_i^{1\text{d}} = ((i-1)H - \delta_{\text{over}}/2, iH + \delta_{\text{over}}/2)\}_{i=1}^{n_{\text{dd}}}.$$

For this choice of the PoU, we have that  $\|\frac{d}{dx} \phi_i^{1\text{d}}\|_{L^\infty(\Omega)} = \frac{1}{\delta_{\text{over}}}$  and thus the constants  $C_i$  in (4.3b) are given by  $C_i = \frac{\sqrt{2}}{\delta_{\text{over}}}$  for  $i = 1, \dots, N_{\text{dd}}$ . Note that, since we impose Dirichlet conditions on  $\partial\Omega$ , we can consider  $\bigcup_{i=1}^{N_{\text{dd}}} \omega_i = \Omega$ . Note also that the constant  $M$  in (4.3a) is equal to four.

It is possible to verify that, for the proposed PoU, there exist  $\widehat{\phi}^{\text{int}}, \widehat{\phi}^{\text{co}}, \widehat{\phi}^{\text{ed}}$  such that

$$\phi_i = \widehat{\phi}^{\text{Li}} \circ \Phi_i^{-1} : i = 1, \dots, N_{\text{dd}},$$

for any choice of  $n_{\text{dd}} \in \mathbb{N}$ ; this is due to the particular choice of the mappings  $\{\Phi_i\}_i$  and of the archetype components.

**7.2.1. Localized training.** We compute  $n_{\text{test}} = 30$  global solutions for  $n_{\text{dd}} = 10$  ( $N_{\text{dd}} = 100$ ) components; then, we define the test datasets  $\{\mathcal{D}^\bullet\}_{\bullet \in \{\text{co}, \text{ed}, \text{int}\}}$  by extracting the solution in each element of  $\mathbb{V}$  —  $\text{card}(\mathcal{D}^\bullet) = 1920$  (resp., 120, 960) for the internal (resp., corner, edge) component. Finally, we introduce the localized error indicators

$$(7.4) \quad E_{\text{avg,rel}}^\bullet(\mathcal{Z}^\bullet) = \frac{1}{\text{card}(\mathcal{D}^\bullet)} \sum_{w \in \mathcal{D}^\bullet} \frac{\|w - \Pi_{\mathcal{Z}^\bullet} w\|_\bullet}{\|w\|_\bullet}, \quad \bullet \in \mathfrak{L},$$

which are used to assess performance.

Figure 7 shows random samples of the boundary conditions on  $\Gamma_{\text{in}}$  for internal and edge components as provided by Algorithm 5.2 for various choices of  $\alpha$  and  $N_f = 20$  and  $\bar{u}_{\text{max}} = 0.5$ . As for the linear case, the value of  $\alpha$  encodes the spatial smoothness of the samples. We further observe that Algorithm 5.2 automatically enforces the proper condition at the extrema  $s = 0$  and  $s = 1$  —  $g(0) = g(1) = 0$  for  $\bullet \in \{\text{co}, \text{ed}\}$ ,  $g(0) = g(1)$  for  $\bullet = \text{int}$ .

Figure 8 shows the behavior of the relative errors (7.4) for the three components for smooth sampling for three choices of  $\alpha$  ( $N_f = 20$ ,  $\bar{u}_{\text{max}} = 0.5$ ), and for Gaussian sampling (7.3). To provide a reference, we also show the performance of the POD spaces based on the datasets  $\{\mathcal{D}_{\text{test}}^\bullet\}_{\bullet \in \{\text{co}, \text{ed}, \text{int}\}}$  (“bench”) generated using 30 additional global simulations with  $N_{\text{dd}} = 100$  components. We observe that smooth sampling outperforms Gaussian sampling for the boundary components; we believe that this is

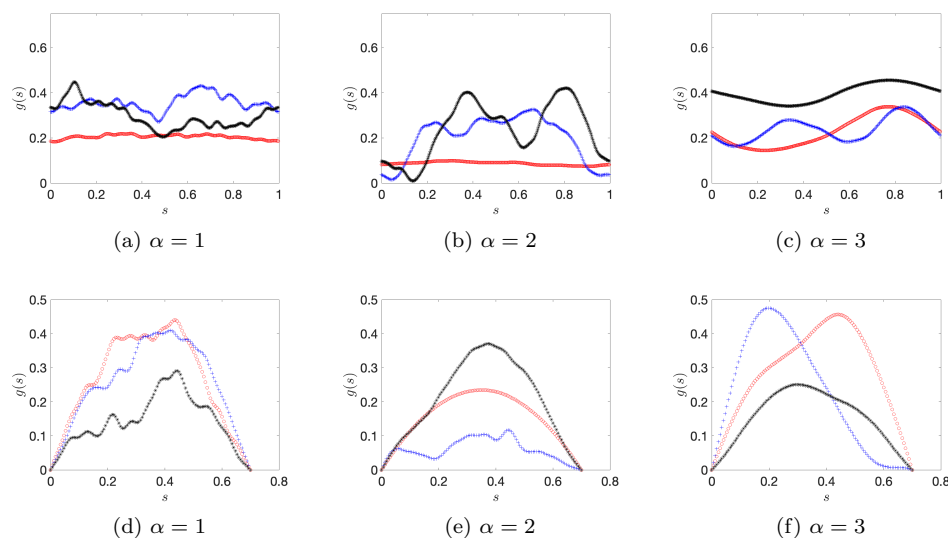


FIG. 7. *Nonlinear problem. Samples of random boundary conditions for three choices of  $\alpha$  ( $N_f = 20$ ,  $\bar{u}_{\text{max}} = 0.5$ ). (a)–(b)–(c) Internal component. (d)–(e)–(f) Edge component.*

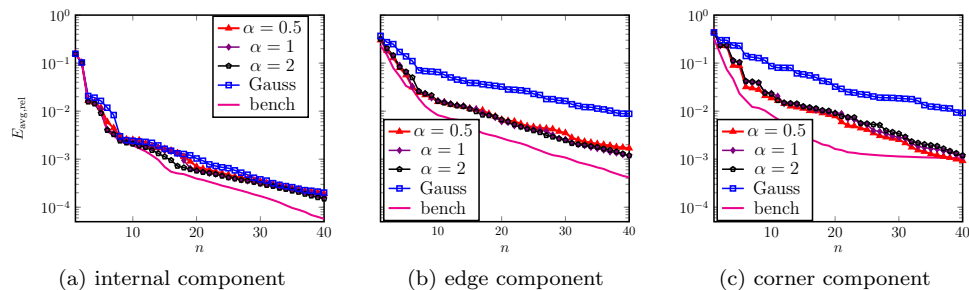


FIG. 8. *Nonlinear problem. Local approximation errors (7.4) for three choices of  $\alpha$  ( $N_f = 20$ ,  $\bar{u}_{max} = 0.5$ ), and for Gaussian sampling (7.3). Comparison with POD spaces based on the datasets  $\{\mathcal{D}_{test}^\bullet\}_{\bullet \in \{co, ed, int\}}$  generated using 30 additional global simulations with  $N_{dd} = 100$  components (bench).*

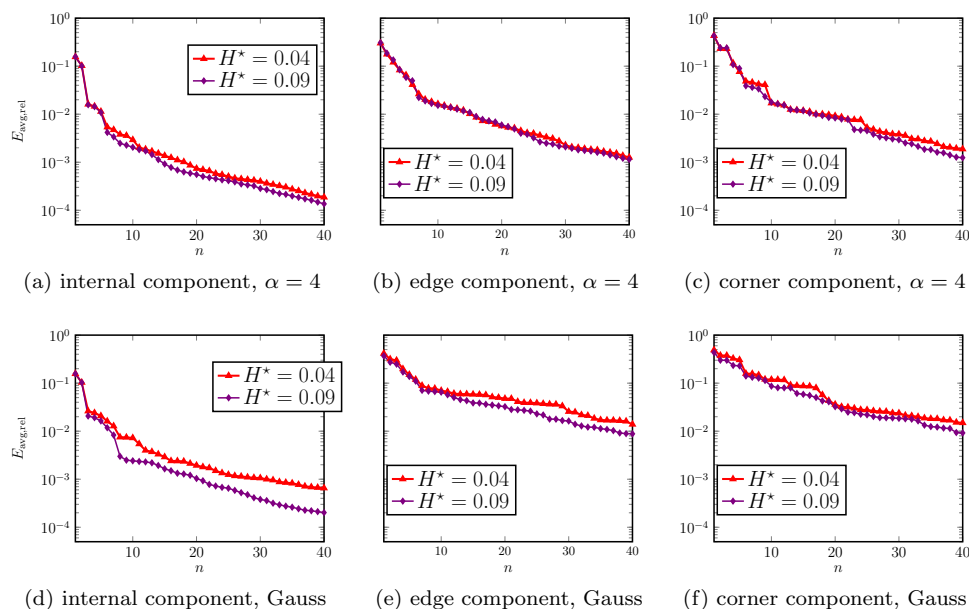


FIG. 9. *Nonlinear problem. Local approximation errors (7.4) for  $\alpha = 4$  ( $N_f = 20$ ,  $\bar{u}_{max} = 0.5$ ), and for Gaussian sampling (7.3), for two choices of  $H^*$ .*

due to the presence of strong Dirichlet conditions on  $\partial\hat{\Omega}_{ovr}^\bullet \setminus \hat{\Gamma}_{in}^\bullet$ . We further observe that results weakly depend on the choice of  $\alpha$ .

In Figure 9, we investigate the choice of the oversampling size. As described in Figure 3, in all numerical tests we consider the oversampling domains

$$\hat{\Omega}_{ovr}^\bullet = \left\{ x \in \mathbb{R}^2 : \inf_{y \in \hat{\Omega}^\bullet} \|x - y\|_\infty < H^* := H - \delta_{ovr}, \right\},$$

where  $H = 0.1$  is the size of the internal domains (cf. (4.6)) and  $\delta_{ovr} = 0.01$  and thus  $H^* = 0.09$ . In Figure 9, we reproduce the same results as Figure 8 with  $H^* = 0.04$ , for the three components, for  $\alpha = 4$  and for Gaussian sampling. We observe that for this particular problem we can consider a significantly smaller oversampling domain without significant deterioration of performance; we also note that Gaussian

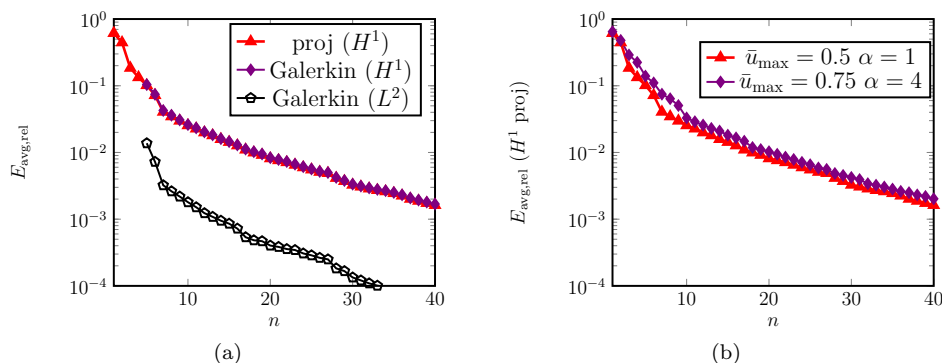


FIG. 10. Nonlinear problem. Performance of the PUM CB-ROM on  $n_{\text{test}} = 30$  global solutions for  $N_{\text{dd}} = 100$ . (a) Galerkin error versus projection error. (b) Projection error for two choices of the parameters in Algorithm 5.2.

sampling is more sensitive to the oversampling size. It thus seems that for this test case choosing boundary conditions that interpolate with high probability a function of higher smoothness (here  $H^4(\Gamma_{\text{in}})$ ) and thus excluding higher frequencies makes enlarging the oversampling domain unnecessary.

Figure 10 shows the performance of the CB-ROM based on the PUM. In Figure 10(a), we show the average global  $L^2$  and  $H^1$  relative errors over the test set of  $n_{\text{test}} = 30$  global simulations and we also compare these with the  $H^1$  relative projection error,  $\frac{1}{n_{\text{test}}} \sum_{i=1}^{n_{\text{test}}} \frac{\|u^{(i)} - \Pi_{\mathbb{Z}_{\text{gfem}}} u^{(i)}\|_{1,\Omega}}{\|u^{(i)}\|_{1,\Omega}}$ . We here consider  $\bar{u}_{\max} = 0.5$ ,  $N_f = 20$ , and  $\alpha = 1$ . We observe that Galerkin projection is nearly optimal for all choices of  $n$ ; we further observe exponential convergence with respect to  $n$ . In Figure 10(b), we compare the  $H^1$  relative projection error for  $\bar{u}_{\max} = 0.5$ ,  $N_f = 20$ , and  $\alpha = 1$  with the results obtained for  $\bar{u}_{\max} = 0.75$ ,  $N_f = 20$ , and  $\alpha = 4$ ; we observe that results weakly depend on the choice of these two hyperparameters.

**7.2.2. Adaptive enrichment.** We apply Algorithm 6.1 with error indicator  $\Delta_\mu$  (6.3) to the model problem of subsection 7.2. We consider  $n_{\text{train}}^{\text{loc}} = 30$ ,  $n^{\text{loc}} = 20$ , we set  $\{p_\mu^\bullet\}_\bullet$  as discussed in subsection 5.2, and we generate random samples of boundary conditions at input ports based on (i) Algorithm 5.2 with  $N_f = 20$ ,  $\bar{u}_{\max} = 0.5$ ,  $\alpha = 1$  or (ii) on iid realizations of (7.3). We further consider  $n_{\text{train}}^{\text{glo}} = 50$ ,  $n^{\text{glo}} = 10$ ,  $\text{maxit} = 3$ , and we generate global configurations using the strategy outlined in subsection 6.2. We assess performance based on  $n_{\text{test}} = 20$  out-of-sample randomly chosen configurations.

Figures 11(a) and (b) show boxplots of the relative  $H^1$  error after each iteration of the training algorithm—iteration 0 corresponds to the performance of the CB-ROM without global enrichment. Iteration 0 corresponds to a reduced space of size  $n = 20$ ; iterations  $it = 1, 2, 3$  correspond to reduced spaces of size  $n = 20 + 10 \cdot it$ . We observe that the enrichment iterations significantly improve performance of the CB-ROM and reduce the impact of the initial sampling distribution. Figure 11(c) shows the correlation between the residual indicator (6.3) and the relative  $H^1$  error on the test set for all iterations of the enrichment algorithm for smooth sampling; Figure 11(d) shows the effectivity of the error indicator  $\eta = \Delta_\mu / E_{\text{rel}}$  for smooth sampling. We observe that the residual indicator is strongly correlated with the global error.

Figure 12 shows the maximum relative in-sample and out-of-sample  $H^1$  error for the same sampling strategies considered in Figure 11. To facilitate the interpretation,

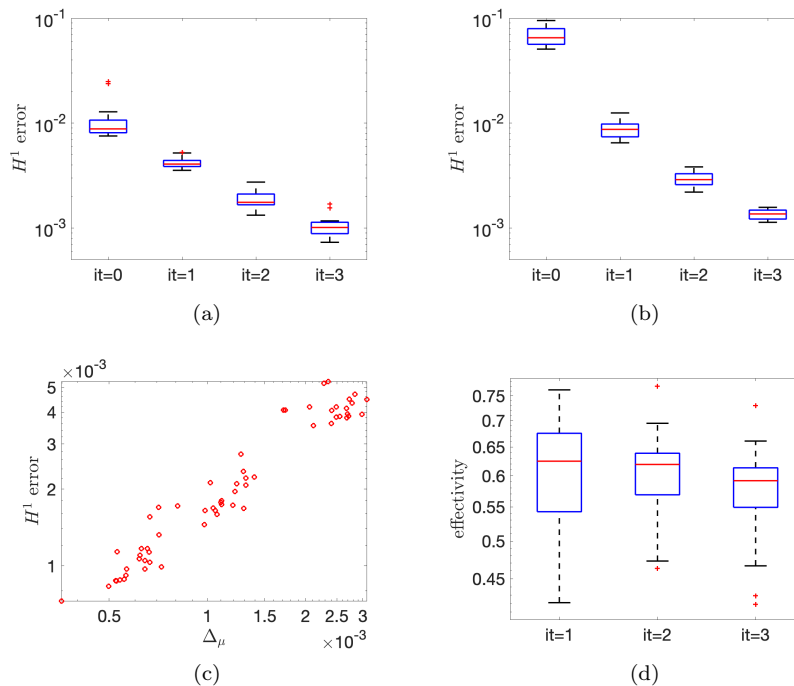


FIG. 11. Nonlinear problem; Adaptive enrichment. (a)–(b) box plots of the relative  $H^1$  error on the test set for smooth and Gaussian sampling of localized BCs. (c) Correlation between  $\Delta_\mu$  and relative  $H^1$  error (smooth sampling). (d) Out-of-sample effectivity of the error indicator  $\Delta_\mu/E_{rel}$  (6.3) (smooth sampling).

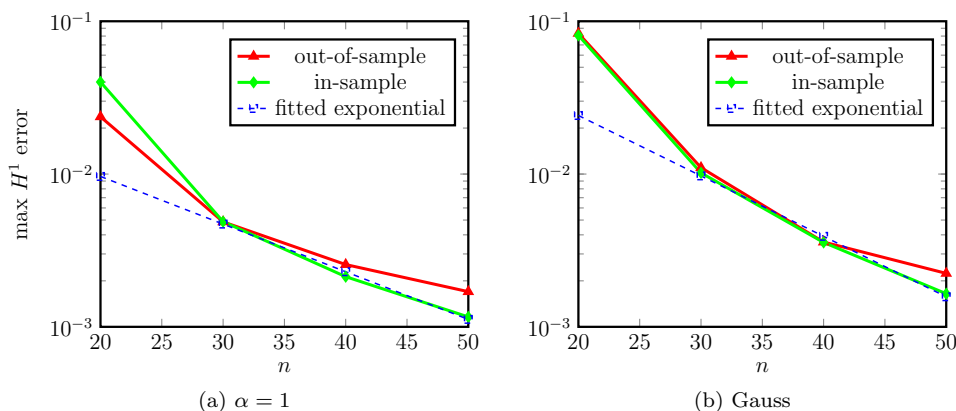


FIG. 12. Nonlinear problem; adaptive enrichment. Behavior of the maximum relative error on training (in-sample) and test (out-of-sample) sets for smooth ( $\alpha=1$ ) and Gaussian sampling.

we also provide the fitted exponential curve  $\hat{E} = \exp(\alpha n + \beta)$  obtained by discarding the first datapoint. For this model problem, numerical results suggest nearly exponential in-sample convergence of the adaptive enrichment strategy.

**8. Conclusions and perspectives.** We presented a CB-pMOR method for parameterized elliptic nonlinear PDEs. The approach relies on the definition of several

archetype components and associated local ROBs and ROMs. CB-pMOR rely on two building blocks: (i) a localized training strategy for the construction of the local approximation spaces, and (ii) a DD strategy for online global predictions. In this paper, we proposed a localized data compression procedure based on oversampling and randomized sampling of boundary conditions of controlled smoothness, and we relied on the PUM to devise global approximation spaces and on Galerkin projection to estimate the global state. Finally, we proposed an adaptive enrichment procedure that exploits global CB-ROM solves to improve approximation properties of the local reduced spaces.

Numerical results for a nonlinear diffusion problem show the impact of the sampling distribution on performance: given a class of nonlinear PDEs, it is thus necessary to devise an effective sampler that is informed by the problem of interest. The approach presented in this work (cf. Algorithm 5.2) is simple to implement, and incorporates relevant features of the problem of interest—lower and upper bounds for the solution, Sobolev regularity, Dirichlet boundary conditions. However, it depends on several hyperparameters that might be difficult to set a priori. In this respect, we numerically showed that the proposed enrichment strategy reduces the impact of the initial sampling distribution.

In the future, we wish to extend the approach in several directions. First, we wish to devise specialized hyperreduction strategies for CB-pMOR methods based on the PUM: hyperreduction is key to reduce efficient online memory and computational costs. Second, we wish to develop rigorous a posteriori error estimators for nonlinear PDEs for online certification. Third, we wish to analyze performance of randomized algorithms for nonlinear operators; this analysis is key to providing mathematical foundations for randomized methods for nonlinear problems and also informing the choice of the sampling distribution.

### Appendix A. Notation.

*Quantities associated with the archetype component*  $\bullet \in \mathfrak{L} = \{\text{co}, \text{ed}, \text{int}\}$ :  $\widehat{\Omega}^\bullet$  reference domain,  $\widehat{\Gamma}_{\text{dir}}^\bullet \subset \partial\widehat{\Omega}^\bullet$  Dirichlet boundary,  $\mathcal{Y}^\bullet \subset H_{0, \widehat{\Gamma}_{\text{dir}}^\bullet}^1(\widehat{\Omega}^\bullet)$  HF discretization,  $\widehat{\phi} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  reference PoU (cf. subsection 4.2),  $\|\cdot\|_\bullet$  seminorm used for POD (cf. (4.9)),  $\mathcal{Z}^\bullet = \text{span}\{\zeta_i^\bullet\}_{i=1}^n \subset \mathcal{Y}^\bullet$  reduced space.

*Oversampling.*  $\widehat{\Omega}_{\text{ovr}}^\bullet \supset \widehat{\Omega}^\bullet$  oversampling domain (or patch),  $\widehat{\Gamma}_{\text{in}}^\bullet \subset \partial\widehat{\Omega}_{\text{ovr}}^\bullet$  input boundary,  $\mathcal{G}^\bullet \subset H^{1/2}(\widehat{\Gamma}_{\text{in}}^\bullet)$  space of admissible boundary conditions for the local solution operator,  $\mathcal{P}^\bullet$  active parameter domain in  $\widehat{\Omega}_{\text{ovr}}^\bullet$ ,  $T^\bullet : \mathcal{G}^\bullet \times \mathcal{P}^\bullet \rightarrow \mathcal{Y}^\bullet$  transfer operator,  $p_\mu^\bullet$  and  $p_{\text{bc}}^\bullet$  pdfs of the distributions over  $\mathcal{P}^\bullet$  and  $\mathcal{G}^\bullet$  used for localized training (cf. Algorithm 5.1).

*Instantiated system.*  $\{\omega_i\}_{i=1}^{N_{\text{dd}}}$  instantiated components,  $\text{Neigh}_i = \{j : \omega_i \cap \omega_j \neq \emptyset\}$  index of the neighboring elements of  $\omega_i$ ,  $\mathcal{P}_{\text{glo}}$  global parameter domain,  $\mathbf{L} : \{1, \dots, N_{\text{dd}}\} \rightarrow \mathfrak{L}$  function that returns the label of each instantiated component,  $\Phi_i : \widehat{\Omega}^{\mathbf{L}_i} \rightarrow \omega_i$  geometric mappings,  $\{\phi_i\}_{i=1}^{N_{\text{dd}}}$  instantiated PoU (cf. (4.8)),  $\mathcal{X}_i := \{\zeta \circ \Phi_i^{-1} : \zeta \in \mathcal{Y}^{\mathbf{L}_i}\}$ ,  $\mathcal{X}_{i,0} := \{\phi_i \zeta \circ \Phi_i^{-1} : \zeta \in \mathcal{Y}^{\mathbf{L}_i}\}$  local HF spaces,  $\mathcal{X}_{\text{pum}}$  PUM space (cf. (4.4)),  $\mathcal{Z}_{\text{gfem}}$  global reduced space (cf. (4.5)),  $\mathfrak{R} : \mathcal{X}_{\text{pum}} \times \mathcal{X}_{\text{pum}} \times \mathcal{P}_{\text{glo}} \rightarrow \mathbb{R}$  global variational form (cf. (4.12)),  $u_\mu$  solution to (4.12) for  $\mu \in \mathcal{P}_{\text{glo}}$ ,  $\widehat{u}_\mu$  solution to the Galerkin ROM (4.13) for  $\mu \in \mathcal{P}_{\text{glo}}$ ,  $\{\zeta_{i,j} : \zeta_{i,j} = \zeta_i^{\mathbf{L}_j} \circ \Phi_j^{-1} : i = 1, \dots, n, j = 1, \dots, N_{\text{dd}}\}$  global ROB (cf. (4.14a)).

*Enrichment.*  $\mathbf{r}_\mu^{(i)} : \mathcal{X}_{\text{pum}} \times \rightarrow \mathbb{R}_+$  local residual (4.16b) for the component  $\omega_i$ ,  $\Delta_\mu : \mathcal{X}_{\text{pum}} \times \rightarrow \mathbb{R}_+$  global error indicator (6.3),  $T_\mu^{(i)} : \mathcal{X}_i \rightarrow \mathcal{X}_{i,0}$  local solution correction operators (3.1).

## REFERENCES

- [1] F. ALBRECHT, B. HAASDONK, M. OHLBERGER, AND S. KAULMANN, *The localized reduced basis multiscale method*, in Proceedings of Algorithm 2012, Conference on Scientific Computing, Vysoke Tatry, Slovakia, 2012, pp. 393–403.
- [2] P. F. ANTONIETTI, P. PACCARINI, AND A. QUARTERONI, *A discontinuous Galerkin reduced basis element method for elliptic problems*, ESAIM Math. Model. Numer. Anal., 50 (2016), pp. 337–360.
- [3] I. BABUSKA AND R. LIPTON, *Optimal local approximation spaces for generalized finite element methods with application to multiscale problems*, Multiscale Model. Simul., 9 (2011), pp. 373–406.
- [4] I. BABUŠKA AND J. M. MELENK, *The partition of unity method*, Internat. J. Numer. Methods Engrg., 40 (1997), pp. 727–758.
- [5] J. BAIGES, R. CODINA, AND S. IDELSOHN, *A domain decomposition strategy for reduced order models. Application to the incompressible Navier-Stokes equations*, Comput. Methods Appl. Mech. Engrg., 267 (2013), pp. 23–42, <https://doi.org/10.1016/j.cma.2013.08.001>.
- [6] M. BARRAULT, Y. MADAY, N. NGUYEN, AND A. PATERA, *An 'empirical interpolation' method: Application to efficient reduced-basis discretization of partial differential equations*, C. R. Math. Acad. Sci. Paris Ser. I, 339 (2004), pp. 667–672.
- [7] M. BERGMANN, A. FERRERO, A. IOLLO, E. LOMBARDI, A. SCARDIGLI, AND H. TELIB, *A zonal Galerkin-free POD model for incompressible flows*, J. Comput. Phys., 352 (2018), pp. 301–325, <https://doi.org/10.1016/j.jcp.2017.10.001>.
- [8] F. BREZZI, J. RAPPAPAZ, AND P.-A. RAVIART, *Finite-dimensional approximation of nonlinear problems. I. Branches of nonsingular solutions*, Numer. Math., 36 (1980/81), pp. 1–25, <https://doi.org/10.1007/BF01395985>.
- [9] A. BUHR, *Exponential convergence of online enrichment in localized reduced basis methods*, IFAC-PapersOnLine, 51 (2018), pp. 302–306.
- [10] A. BUHR, C. ENGWER, M. OHLBERGER, AND S. RAVE, *ArbiLoMod, a simulation technique designed for arbitrary local modifications*, SIAM J. Sci. Comput., 39 (2017), pp. A1435–A1465.
- [11] A. BUHR, L. IAPICHINO, M. OHLBERGER, S. RAVE, F. SCHINDLER, AND K. SMETANA, *Localized model reduction for parameterized problems*, in Model Order Reduction: Vol. 2: Snapshot-Based Methods and Algorithms, De Gruyter, Berlin, 2020, pp. 245–306, <https://doi.org/doi:10.1515/9783110671490-006>.
- [12] A. BUHR AND K. SMETANA, *Randomized local model order reduction*, SIAM J. Sci. Comput., 40 (2018), pp. A2120–A2151.
- [13] V. M. CALO, Y. EFENDIEV, J. GALVIS, AND G. LI, *Randomized oversampling for generalized multiscale finite element methods*, Multiscale Model. Simul., 14 (2016), pp. 482–501.
- [14] G. CALOZ AND J. RAPPAPAZ, *Numerical analysis for nonlinear and bifurcation problems*, in Handbook of Numerical Analysis, Vol. V, North-Holland, Amsterdam, 1997, pp. 487–637, [https://doi.org/10.1016/S1570-8659\(97\)80004-X](https://doi.org/10.1016/S1570-8659(97)80004-X).
- [15] C. CANUTO, T. TONN, AND K. URBAN, *A posteriori error analysis of the reduced basis method for nonaffine parametrized nonlinear PDEs*, SIAM J. Numer. Anal., 47 (2009), pp. 2001–2022, <https://doi.org/10.1137/080724812>.
- [16] S. CHATURANTABUT AND D. C. SORESENSEN, *Nonlinear model reduction via discrete empirical interpolation*, SIAM J. Sci. Comp., 32 (2010), pp. 2737–2764.
- [17] K. CHEN, Q. LI, J. LU, AND S. J. WRIGHT, *Randomized sampling for basis function construction in generalized finite element methods*, Multiscale Model. Simul., 18 (2020), pp. 1153–1177.
- [18] S. CHEN, Q. LI, J. LU, AND S. J. WRIGHT, *Manifold learning and nonlinear homogenization*, Multiscale Model. Simul., 20 (2022), pp. 1093–1126.
- [19] Y. CHEN, B. HOSSEINI, H. OWHADI, AND A. M. STUART, *Solving and learning nonlinear PDEs with Gaussian processes*, J. Comput. Phys., 447 (2021), 110668.
- [20] M. CHIPOT, *Elements of Nonlinear Analysis*, Birkhäuser, Basel, 2000, <https://doi.org/10.1007/978-3-0348-8428-0>.
- [21] P. DRINEAS AND M. W. MAHONEY, *RandNLA: Randomized numerical linear algebra*, Commun. ACM, 59 (2016), pp. 80–90.
- [22] J. L. EFTANG AND A. T. PATERA, *Port reduction in parametrized component static condensation: Approximation and a posteriori error estimation*, Internat. J. Numer. Methods Engrg., 96 (2013), pp. 269–302.
- [23] A. ERN AND J.-L. GUERMOND, *Theory and Practice of Finite Elements*, Appl. Math. Sci. 159, Springer, New York, 2004.
- [24] C. FARHAT, T. CHAPMAN, AND P. AVERY, *Structure-preserving, stability, and accuracy properties of the energy-conserving sampling and weighting method for the hyper reduction of*

- nonlinear finite element dynamic models*, Internat. J. Numer. Methods Engrg., 102 (2015), pp. 1077–1110.
- [25] M. GIAQUINTA AND J. SOUČEK, *Caccioppoli's inequality and Legendre-Hadamard condition*, Math. Ann., 270 (1985), pp. 105–107, <https://doi.org/10.1007/BF01455535>.
  - [26] D. GILBARG AND N. S. TRUDINGER, *Elliptic Partial Differential Equations of Second Order*, Grundlehren Math. Wiss. 224, Springer, Berlin, 1977.
  - [27] B. HAASDONK, *Reduced basis methods for parametrized PDEs – a tutorial*, in Model Reduction and Approximation, P. BENNER, A. COHEN, M. OHLBERGER, AND K. WILLCOX, eds., SIAM, Philadelphia, PA, 2017, pp. 65–136.
  - [28] N. HALKO, P. G. MARTINSSON, AND J. A. TROPP, *Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions*, SIAM Rev., 53 (2011), pp. 217–288, <https://doi.org/10.1137/090771806>.
  - [29] P. HENNING AND D. PETERSEIM, *Oversampling for the multiscale finite element method*, Multiscale Model. Simul., 11 (2013), pp. 1149–1175, <https://doi.org/10.1137/120900332>.
  - [30] J. HESTHAVEN, G. ROZZA, AND B. STAMM, *Certified reduced basis methods for parametrized partial differential equations*, Springer Briefs Math., Springer, Cham, Switzerland, 2016.
  - [31] C. HOANG, Y. CHOI, AND K. CARLBERG, *Domain-decomposition least-squares Petrov-Galerkin (DD-LSPG) nonlinear model reduction*, Comput. Methods Appl. Mech. Engrg., 384 (2021), 113997.
  - [32] 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 (1997), pp. 169–189.
  - [33] D. B. P. HUYNH, D. J. KNEZEVIC, AND A. T. PATERA, *A static condensation reduced basis element method: Approximation and a posteriori error estimation*, ESAIM Math. Model. Numer. Anal., 47 (2013), pp. 213–251.
  - [34] L. IAPICHINO, A. QUARTERONI, AND G. ROZZA, *A reduced basis hybrid method for the coupling of parametrized domains represented by fluidic networks*, Comput. Methods Appl. Mech. Engrg., 221–222 (2012), pp. 63–82.
  - [35] M. KAHLBACHER AND S. VOLKWEIN, *Galerkin proper orthogonal decomposition methods for parameter dependent elliptic systems*, Discuss. Math. Differ. Incl. Control Optim., 27 (2007), pp. 95–117.
  - [36] M. KAPTEYN, D. KNEZEVIC, D. HUYNH, M. TRAN, AND K. WILLCOX, *Data-driven physics-based digital twins via a library of component-based reduced-order models*, Internat. J. Numer. Methods Engrg., 123 (2022), pp. 2986–3003.
  - [37] D. J. KNEZEVIC, H. KANG, P. SHARMA, G. MALINOWSKI, T. T. NGUYEN, *Structural integrity management of offshore structures via RB-FEA and fast full load mapping based digital twins*, in The 28th International Ocean and Polar Engineering Conference, International Society of Offshore and Polar Engineers, Mountain View, CA, 2018, ISOPE-I-18-185.
  - [38] A. KOLMOGOROFF, *Über die beste annäherung von Funktionen einer gegebenen funktionenklasse*, Ann. of Math. (2), 37 (1936), pp. 107–110.
  - [39] C. LE BRIS, F. LEGOLL, AND A. LOZINSKI, *An MsFEM type approach for perforated domains*, Multiscale Model. Simul., 12 (2014), pp. 1046–1077, <https://doi.org/10.1137/130927826>, <https://mathscinet.ams.org/mathscinet-getitem?mr=3231975>.
  - [40] X. LIU, E. CHUNG, AND L. ZHANG, *Iterated numerical homogenization for multiscale elliptic equations with monotone nonlinearity*, Multiscale Model. Simul., 19 (2021), pp. 1601–1632.
  - [41] C. MA, R. SCHEICHL, AND T. DODWELL, *Novel design and analysis of generalized finite element methods based on locally optimal spectral approximations*, SIAM J. Numer. Anal., 60 (2022), pp. 244–273.
  - [42] Y. MADAY AND E. M. RØNQUIST, *A reduced-basis element method*, J. Sci. Comput., 17 (2002), pp. 447–459.
  - [43] M. W. MAHONEY AND P. DRINEAS, *CUR matrix decompositions for improved data analysis*, Proc. Natl. Acad. Sci. USA, 106 (2009), pp. 697–702.
  - [44] A. MÅLQVIST AND D. PETERSEIM, *Localization of elliptic multiscale problems*, Math. Comp., 83 (2014), pp. 2583–2603, <https://doi.org/10.1090/S0025-5718-2014-02868-8>.
  - [45] P.-G. MARTINSSON, V. ROKHLIN, AND M. TYGERT, *A randomized algorithm for the decomposition of matrices*, Appl. Comput. Harmon. Anal., 30 (2011), pp. 47–68.
  - [46] J. M. MELENK AND I. BABUŠKA, *The partition of unity finite element method: Basic theory and applications*, Comput. Methods Appl. Mech. Engrg., 139 (1996), pp. 289–314.
  - [47] A. MICHEL, *A finite volume scheme for two-phase immiscible flow in porous media*, SIAM J. Numer. Anal., 41 (2003), pp. 1301–1317, <https://doi.org/10.1137/S0036142900382739>.
  - [48] S. A. NIEDERER, M. S. SACKS, M. GIROLAMI, AND K. WILLCOX, *Scaling digital twins from the artisanal to the industrial*, Nature Comput. Sci., 1 (2021), pp. 313–320.

- [49] M. OHLBERGER AND F. SCHINDLER, *Error control for the localized reduced basis multiscale method with adaptive on-line enrichment*, SIAM J. Sci. Comput., 37 (2015), pp. A2865–A2895.
- [50] H. OWHADI, *Bayesian numerical homogenization*, Multiscale Model. Simul., 13 (2015), pp. 812–828.
- [51] H. OWHADI, *Multigrid with rough coefficients and multiresolution operator decomposition from hierarchical information games*, SIAM Rev., 59 (2017), pp. 99–149.
- [52] H. OWHADI AND L. ZHANG, *Localized bases for finite-dimensional homogenization approximations with nonseparated scales and high contrast*, Multiscale Model. Simul., 9 (2011), pp. 1373–1398, <https://doi.org/10.1137/100813968>.
- [53] H. OWHADI, L. ZHANG, AND L. BERLYAND, *Polyharmonic homogenization, rough polyharmonic splines and sparse super-localization*, ESAIM Math. Model. Numer. Anal., 48 (2014), pp. 517–552, <https://doi.org/10.1051/m2an/2013118>.
- [54] L. PEGOLOTTI, M. R. PFALLER, A. L. MARSDEN, AND S. DEPARIS, *Model order reduction of flow based on a modular geometrical approximation of blood vessels*, Comput. Methods Appl. Mech. Engrg., 380 (2021), 113762.
- [55] A. PINKUS, *N-Widths in Approximation Theory*, Ergeb. Math. Grenzgeb. (3) 7, Springer, Berlin, 1985.
- [56] J. POUSIN AND J. RAPPAZ, *Consistency, stability, a priori and a posteriori errors for Petrov-Galerkin methods applied to nonlinear problems*, Numer. Math., 69 (1994), pp. 213–231, <https://doi.org/10.1007/s002110050088>.
- [57] A. QUARTERONI, A. MANZONI, AND F. NEGRI, *Reduced Basis Methods for Partial Differential Equations*, Unitem 92, Springer, Cham, Switzerland, 2016.
- [58] A. RASHEED, O. SAN, AND T. KVAMSDAL, *Digital twin: Values, challenges and enablers from a modeling perspective*, IEEE Access, 8 (2020), pp. 21980–22012.
- [59] D. RYCKELYNCK, *A priori hyperreduction method: An adaptive approach*, J. Comput. Phys., 202 (2005), pp. 346–366.
- [60] J. SCHLEUSS AND K. SMETANA, *Optimal local approximation spaces for parabolic problems*, Multiscale Model. Simul., 20 (2022), pp. 551–582.
- [61] J. SCHLEUSS, K. SMETANA, AND L. TER MAAT, *Randomized quasi-optimal local approximation spaces in time*, preprint, arXiv:2203.06276, 2022.
- [62] K. SMETANA, *A Dimensional Reduction Approach Based on the Application of Reduced Basis Methods in the Context of Hierarchical Model Reduction*, Ph.D. thesis, University of Münster, Münster, Germany, 2013.
- [63] K. SMETANA, *A new certification framework for the port reduced static condensation reduced basis element method*, Comput. Methods Appl. Mech. Engrg., 283 (2015), pp. 352–383, <https://doi.org/10.1016/j.cma.2014.09.020>.
- [64] K. SMETANA AND M. OHLBERGER, *Hierarchical model reduction of nonlinear partial differential equations based on the adaptive empirical projection method and reduced basis techniques*, ESAIM Math. Model. Numer. Anal., 51 (2017), pp. 641–677, <https://doi.org/10.1051/m2an/2016031>.
- [65] K. SMETANA AND A. T. PATERA, *Optimal local approximation spaces for component-based static condensation procedures*, SIAM J. Sci. Comput., 38 (2016), pp. A3318–A3356.
- [66] T. STROUBOLIS, I. BABUŠKA, AND K. COPPS, *The design and analysis of the generalized finite element method*, Comput. Methods Appl. Mech. Engrg., 181 (2000), pp. 43–69, [https://doi.org/10.1016/S0045-7825\(99\)00072-9](https://doi.org/10.1016/S0045-7825(99)00072-9).
- [67] T. TADDEI, *Model Order Reduction Methods for Data Assimilation; State Estimation and Structural Health Monitoring*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, 2017.
- [68] T. TADDEI AND A. T. PATERA, *A localization strategy for data assimilation; Application to state estimation and parameter estimation*, SIAM J. Sci. Comput., 40 (2018), pp. B611–B636, <https://doi.org/10.1137/17M1116830>.
- [69] T. TADDEI AND L. ZHANG, *A discretize-then-map approach for the treatment of parameterized geometries in model order reduction*, Comput. Methods Appl. Mech. Engrg., 384 (2021), 113956.
- [70] B. VERFÜRTH, *Numerical homogenization for nonlinear strongly monotone problems*, IMA J. Numer. Anal., 42 (2022), pp. 1313–1338, <https://doi.org/10.1093/imanum/drab004>.
- [71] R. VERFÜRTH, *A posteriori error estimates for nonlinear problems. Finite element discretizations of elliptic equations*, Math. Comp., 62 (1994), pp. 445–475, <https://doi.org/10.2307/2153518>.
- [72] K. VEROY AND A. T. PATERA, *Certified real-time solution of the parametrized steady incompressible Navier-Stokes equations: Rigorous reduced-basis a posteriori error bounds*, Internat. J. Numer. Methods Fluids, 47 (2005), pp. 773–788, <https://doi.org/10.1002/fld.867>.

- [73] M. YANO AND A. T. PATERA, *An LP empirical quadrature procedure for reduced basis treatment of parametrized nonlinear PDEs*, Comput. Methods Appl. Mech. Engrg., 344 (2019), pp. 1104–1123, <https://doi.org/10.1016/j.cma.2018.02.028>.
- [74] D. YU AND S. CHAKRAVORTY, *A randomized proper orthogonal decomposition technique*, in 2015 American Control Conference (ACC), IEEE, Piscataway, NJ, 2015, pp. 1137–1142.