SIAM J. SCI. COMPUT.
Vol. 41, No. 4, pp. B669–B693

ADAPTIVE PBDW APPROACH TO STATE ESTIMATION: NOISY OBSERVATIONS; USER-DEFINED UPDATE SPACES*

YVON MADAY† AND TOMMASO TADDEI‡

Abstract. We provide a number of extensions and further interpretations of the parameterized-background data-weak (PBDW) formulation, a real-time and in situ data assimilation framework for physical systems modeled by parameterized PDEs, proposed in [Y. Maday et al., *Int. J. Numer. Methods Engrg.*, 102 (2015), pp. 933–965]. Given M noisy measurements of the state, PBDW seeks an approximation of the form $u^\star = z^\star + \eta^\star$, where the *background* $z^\star$ belongs to an N-dimensional *background space* informed by a parameterized mathematical model and the *update* $\eta^\star$ belongs to an M-dimensional *update space* informed by the experimental observations. The contributions of the present work are threefold. First, we extend the adaptive formulation proposed in [T. Taddei, *ESAIM Math. Model. Numer. Anal.*, 51 (2017), pp. 1827–1858] to general linear observation functionals to effectively deal with noisy observations; second, we consider a user-defined choice of the update space to improve convergence with respect to the number of measurements. Third, we propose an a priori error analysis for general linear functionals in the presence of noise to identify the different sources of state estimation error and ultimately motivate the adaptive procedure. We present results for two synthetic model problems in acoustics, to illustrate the elements of the methodology and to prove its effectiveness. We further present results for a synthetic problem in fluid mechanics to demonstrate the applicability of the approach to vector-valued fields.

Key words. variational data assimilation, parameterized partial differential equations, model order reduction, design of experiment

AMS subject classifications. 62-07, 93E24

DOI. 10.1137/18M116544X

1. Introduction.

Data assimilation (DA) refers to the process of integrating information coming from a (possibly parameterized) mathematical model with experimental observations for prediction. State estimation is a particular DA task in which the quantity of interest is the state u^{true} of a physical system over a domain of interest $\Omega \subset \mathbb{R}^d$. In this work, we propose a number of extensions and further interpretations of the parameterized-background data-weak (PBDW) approach to state estimation, first presented in [12].

As in [18], we denote by $\{y_m\}_{m=1}^M$ the set of experimental measurements, and we denote by $u^{\text{bk}}(\mu) \in \mathcal{X}$ the solution to the parameterized best-knowledge (bk) mathematical model for the parameter value $\mu \in \mathcal{P}^{\text{bk}}$, $G^{\text{bk},\mu}(u^{\text{bk}}(\mu)) = 0$. Here, the space $\mathcal{X} = \mathcal{X}(\Omega)$ is a suitable Hilbert space defined over Ω, $G^{\text{bk},\mu}(\cdot)$ denotes the parameterized bk mathematical model associated with the physical system, and $\mathcal{P}^{\text{bk}} \subset \mathbb{R}^P$ reflects the uncertainty in the value of the parameters of the model. We here consider measurements of the form $y_m = \ell_m^o(u^{\text{true}}) + \epsilon_m$, where ℓ_m^o is a local average of the state about the location x_m^{obs} in Ω and ϵ_m reflects the observational noise. On the other hand, the uncertainty in the parameters of the model leads

*Submitted to the journal's Computational Methods in Science and Engineering section January 16, 2018; accepted for publication (in revised form) May 8, 2019; published electronically July 2, 2019.

https://doi.org/10.1137/18M116544X

†Sorbonne Universités, UPMC Univ. Paris 06, Laboratoire Jacques-Louis Lions, Paris, F 75005, France, and Brown University, Division of Applied Mathematics, Providence RI, 02906 (maday@ann.jussieu.fr and yvon_jean_maday@brown.edu).

‡Institut de Mathématiques de Bordeaux (UMR 5251), Team MEMPHIS, Inria Bordeaux–Sud Ouest (tommaso.taddei@inria.fr).

to the definition of the bk manifold $\mathcal{M}^{\rm bk} := \{u^{\rm bk}(\mu)|_\Omega : \mu \in \mathcal{P}^{\rm bk}\} \subset \mathcal{X}$, which collects the solution to the parameterized bk model for all values of the parameter in $\mathcal{P}^{\rm bk}$, restricted to the domain of interest Ω. Since in practice the model is affected by nonparameterized (unanticipated) uncertainty, the true field does not in general belong to $\mathcal{M}^{\rm bk}$.

The key idea of the PBDW formulation is to seek an approximation $u^\star = z^\star + \eta^\star$ to the true field $u^{\rm true}$ employing projection-by-data. The first contribution to $u^\star$, $z^\star \in \mathcal{Z}_N$, is the *deduced background estimate*. The linear N-dimensional space $\mathcal{Z}_N \subset \mathcal{X}$ is informed by the bk manifold $\mathcal{M}^{\rm bk}$: It thus encodes—in a mathematically convenient way—the available prior knowledge about the system coming from the model. The second contribution to $u^\star$, $\eta^\star \in \mathcal{U}_M$, is the *update estimate*. The linear M-dimensional space $\mathcal{U}_M$ is the span of the Riesz representers of the M observation functionals $\{\ell^o_m\}_{m=1}^M$: $\mathcal{U}_M$ improves the approximation properties of the search space associated with the state estimation procedure. From a modeling perspective, the deduced background $z^\star$ addresses the parameterized uncertainty in the bk model, while the update addresses the nonparametric or unanticipated uncertainty. In [12], for perfect measurements ($\epsilon_m = 0$ for $m = 1, \ldots, M$), the pair $(z^\star, \eta^\star) \in \mathcal{Z}_N \times \mathcal{U}_M$ is computed by searching for $\eta^\star$ of minimum norm subject to the observation constraints $\ell^o_m(z^\star + \eta^\star) = y_m$ for $m = 1, \ldots, M$. In [18], for pointwise noisy measurements, the pair $(z^\star, \eta^\star)$ is computed by solving a suitable Tikhonov regularization of the original constrained minimization statement proposed in [12].

PBDW provides some new contributions. First, the variational formulation facilitates the construction of a priori error estimates, which might guide the optimal choice of the experimental observations. Second, the background space $\mathcal{Z}_N$ accommodates anticipated uncertainty associated with the parameters of the model in a computationally convenient way; the construction of $\mathcal{Z}_N$ based on $\mathcal{M}^{\rm bk}$ relies on the application of parametric model order reduction techniques. Third, unlike standard least-squares methods, PBDW provides a mechanism—the update $\eta^\star$—to correct the deficiencies of the bk model. Finally, projection-by-data, as opposed to projection-by-model, simplifies the treatment of uncertainty in boundary conditions, particularly when the mathematical model is defined over a domain $\Omega^{\rm bk}$ that strictly contains Ω ([20]). We remark that several of these ingredients have appeared in different contexts: We refer the reader to [12, 13, 18] for a thorough overview of the links between the PBDW formulation and other DA techniques and to [1, 5] for a thorough introduction to DA.

In this work, we propose three contributions to the original PBDW formulation:

1. We extend the adaptive formulation proposed in [18] for pointwise measurements to general linear observation functionals. The approach reads as a Tikhonov regularization of the constrained minimization statement in [12] and can also be interpreted as a convex relaxation of the partial spline model (PSM) [24, Chapter 9].
2. In [12], the update space $\mathcal{U}_M$ is induced by the particular inner product chosen for $\mathcal{X}$ (*variational update*); in this work, we propose to first choose the space $\mathcal{U}_M$ and then recover the variational formulation through the definition of a suitable inner product (*user-defined update*). We demonstrate that our choice reduces the offline costs and improves the approximation properties of the update space for smooth fields. We emphasize that the idea of selecting the approximation space before introducing the inner products is widely used in the kernel methods literature (see, e.g., [25]) for pointwise measurements and has already been exploited in [18].

3. We propose an a priori error analysis for general linear functionals in the presence of noise. First, we present a bound for the deterministic error between u^{true} and the PBDW solution u^{opt} fed with perfect measurements. Then we present a bound for the stochastic error between the possibly regularized PBDW solution $u^\star$ fed with imperfect (noisy) measurements and the PBDW solution u^{opt} fed with perfect measurements. The idea of identifying different sources of error and deriving distinct bounds for each source is the same as that exploited in [21] for the estimate of the state estimation error based on local measurements of the state.

The outline of the paper is as follows. In section 2, we present the methodology: We review the derivation of the adaptive PBDW formulation for noisy measurements as discussed in [19], we introduce the user-defined update space and discuss its variational interpretation, we discuss the selection of the observation functionals, we present an adaptive procedure for the selection of the hyperparameter associated with the Tikhonov regularizer, and we discuss the extension to vector-valued problems. In section 3, we present an a priori error analysis for noisy measurements for the state estimation error. In section 4, we present results for two synthetic model problems in acoustics to illustrate the elements of the methodology and to prove its effectiveness. Finally, in section 5, we consider a synthetic fluid mechanics model problem to demonstrate the applicability of PBDW to vector-valued problems.

2. Formulation.

2.1. Preliminaries. Given the Lipschitz domain $\Omega \subset \mathbb{R}^d$, we introduce the Hilbert space $\mathcal{X}$ defined over Ω; we endow $\mathcal{X}$ with the inner product $(\cdot,\cdot)$ and the induced norm $\|\cdot\| = \sqrt{(\cdot,\cdot)}$. We further denote by $\mathcal{X}'$ the dual space of $\mathcal{X}$. For any closed linear subspace $\mathcal{Q} \subset \mathcal{X}$, we denote by $\Pi_{\mathcal{Q}} : \mathcal{X} \to \mathcal{Q}$ the orthogonal projection operator onto $\mathcal{Q}$, and we denote by $\mathcal{Q}^\perp$ its orthogonal complement. Given the linear functional $\ell \in \mathcal{X}'$, we denote by $R_{\mathcal{X}}\ell \in \mathcal{X}$ the corresponding Riesz element, $(R_{\mathcal{X}}\ell, f) = \ell(f)$ for all $f \in \mathcal{X}$.

Given a random variable X, we denote by $\mathbb{E}[X]$ and by $\mathbb{V}[X]$ the mean and the variance, where $\mathbb{E}$ denotes expectation. We denote by $X \sim \mathcal{N}(m, \sigma^2)$ a Gaussian random variable with mean m and variance σ^2. Similarly, we denote by $X \sim \text{Uniform}(\Omega)$ a uniform random variable over Ω. Furthermore, we refer to an arbitrary random variable ε such that $\mathbb{E}[\varepsilon] = 0$ and $\mathbb{V}[\varepsilon] = \sigma^2$ using the notation $\varepsilon \sim (0, \sigma^2)$.

We state up front that in this section, we limit ourselves to real-valued problems; however, the discussion can be trivially extended to the complex-valued case. Furthermore, in sections 2.2–2.7, we consider scalar fields; then, in section 2.8, we discuss the extension to vector-valued problems.

2.2. Derivation of the PBDW statement. As explained in the introduction, we aim to estimate the deterministic state $u^{\text{true}} \in \mathcal{X}$ over the domain of interest $\Omega \subset \mathbb{R}^d$. We shall afford ourselves two sources of information: a bk mathematical model,

$$G^{\text{bk},\mu}(u^{\text{bk}}(\mu)) = 0, \quad \mu \in \mathcal{P}^{\text{bk}},$$

defined over a domain Ω^{bk} that contains Ω, and M experimental observations $y_1, \dots, y_M$ such that

$$y_m = \ell_m^o(u^{\text{true}}) + \epsilon_m, \qquad m = 1, \dots, M,$$

where $\ell_1^o, \dots, \ell_M^o \in \mathcal{X}'$ are suitable observation functionals and $\{\epsilon_m\}_{m=1}^M$ are unknown disturbances caused by either systematic error in the data acquisition system or experimental random noise. Here, $\mathcal{P}^{\text{bk}} \subset \mathbb{R}^P$ is a confidence region for the

true values of the parameters of the model. We further introduce the bk manifold $\mathcal{M}^{\rm bk} = \{u^{\rm bk}(\mu)|_\Omega : \mu \in \mathcal{P}^{\rm bk}\}$ associated with the solution to the parameterized model for all values of the parameters in $\mathcal{P}^{\rm bk}$, restricted to the domain of interest Ω.

In order to combine the parameterized model with the experimental observations, Wahba proposed the following generalization of the 3D-VAR [10] statement for parameterized background, known as PSM [24, Chapter 9]: Find the state estimate $u^\star_\xi = u^{\rm bk}(\mu^\star_\xi) + \eta^\star_\xi$ such that

$$(\mu^\star_\xi, \eta^\star_\xi) := \arg \min_{(\mu,\eta)\in\mathcal{P}^{\rm bk}\times\mathcal{U}} \xi\|\eta\|^2 + V_M\left(\mathcal{L}_M(u^{\rm bk}(\mu)+\eta) - \mathbf{y}\right), \tag{1a}$$

where $\xi > 0$ is a tuning hyperparameter, $\mathbf{y} = [y_1, \dots, y_M]$ is the vector of experimental observations, and $\mathcal{L}_M(u) = [\ell^o_1(u), \dots, \ell^o_M(u)]$. The loss $V_M : \mathbb{R}^M \to \mathbb{R}_+$ is a strictly convex loss function with minimum in $\mathbf{0}$. The choice of the loss function V_M should be based on the expected experimental disturbances: In this work we consider the Euclidean loss

$$V_M(\cdot) := \|\cdot\|_2^2. \tag{1b}$$

This choice is appropriate for homoscedastic noise (i.e., $\epsilon_1, \dots, \epsilon_M$ are independent realizations of the random variable $\epsilon \sim (0, \sigma^2)$). Finally, $\mathcal{U}$ is the *search space*: $\mathcal{U}$ is a closed linear subspace contained in $\mathcal{X}$ that will be specified in the next section.

The first term in (1) penalizes the distance of $u^\star_\xi$ from the bk manifold $\mathcal{M}^{\rm bk}$, while the second term penalizes the data misfit; finally, the hyperparameter $\xi > 0$ weights the importance of the background compared to the experimental data and should be selected adaptively. We observe that (1) is nonconvex in μ; furthermore, evaluations of the map $\mu \mapsto u^{\rm bk}(\mu)$ involve the solution to the bk model. Therefore, (1) is not suitable for real-time computations.

If we introduce the rank-N approximation ([6]) of the bk field $u^{\rm bk}(\mu)$,

$$u^{\rm bk}_N|_\Omega(x,\mu) = \sum_{n=1}^N z_n(\mu)\,\zeta_n(x), \qquad x \in \Omega,\ \ \mu \in \mathcal{P}^{\rm bk},$$

we can approximate statement (1) as

$$(\mu^\star_\xi, \eta^\star_\xi) := \arg \min_{(\mu,\eta)\in\mathcal{P}^{\rm bk}\times\mathcal{U}} \xi\|\eta\|^2 + V_M\left(\mathcal{L}_M\left(\sum_{n=1}^N z_n(\mu)\zeta_n + \eta\right) - \mathbf{y}\right). \tag{2}$$

Provided that we are not interested in the estimate of the parameter $\mu^\star_\xi$, (2) is equivalent to

$$(\boldsymbol{z}^\star_\xi, \eta^\star_\xi) = \arg \min_{(\boldsymbol{z},\eta)\in\Phi_N\times\mathcal{U}} \xi\|\eta\|^2 + V_M\left(\mathcal{L}_M\left(\sum_{n=1}^N z_n\zeta_n + \eta\right) - \mathbf{y}\right), \tag{3}$$

where $\Phi_N = \{[z_1(\mu), \dots, z_N(\mu)] : \mu \in \mathcal{P}^{\rm bk}\} \subset \mathbb{R}^N$ is the image of the parameter-dependent coefficients and the corresponding state estimate is given by $u^\star_\xi = \sum_{n=1}^N (\mathbf{z}^\star_\xi)_n\zeta_n + \eta^\star_\xi$.

The set Φ_N in (3) is nonconvex; therefore, we convexify (3) by minimizing the objective over the convex approximation of Φ_N, $\widetilde{\Phi}_N$:

$$(\boldsymbol{z}^\star_\xi, \eta^\star_\xi) = \arg \min_{(\boldsymbol{z},\eta)\in\widetilde{\Phi}_N\times\mathcal{U}} \xi\|\eta\|^2 + V_M\left(\mathcal{L}_M\left(\sum_{n=1}^N z_n\zeta_n + \eta\right) - \mathbf{y}\right). \tag{4}$$

We refer to (4) as to the PBDW formulation: The formulation generalizes the state-

ments introduced in [12, 13, 18]. The objective in (4) penalizes the distance of the state estimate $u_\xi^\star = z_\xi^\star + \eta_\xi^\star$ from the set $\mathcal{K}_N = \{\sum_n z_n\zeta_n : \mathbf{z} \in \widetilde{\Phi}_N\}$. If $\widetilde{\Phi}_N$ is convex, (4) is considerably less expensive to solve than the PSM (1). To properly incorporate the prior knowledge associated with the mathematical model, the set of functions $\{\zeta_n\}_n$ and the admissible set $\widetilde{\Phi}_N$ should be chosen so that $\mathcal{K}_N$ is an accurate approximation of the manifold $\mathcal{M}^{\text{bk}}$. Towards this end, we resort to model reduction techniques to build the expansion $\zeta_1, \ldots, \zeta_N$: More in detail, for the examples in section 4 we resort to the weak-greedy algorithm (see, e.g., [17]) to build $\{\zeta_n\}_n$, while in section 5 we resort to proper orthogonal decomposition (POD) (see, e.g., [23]). On the other hand, we here consider $\widetilde{\Phi}_N = \mathbb{R}^N$; we refer to a future work for the analysis of tighter relaxations $\widetilde{\Phi}_N$ of Φ_N.

For $\widetilde{\Phi}_N = \mathbb{R}^N$, (4) can be restated as

$$(z_\xi^\star, \eta_\xi^\star) := \arg \inf_{(z,\eta)\in\mathcal{Z}_N\times\mathcal{U}} J_\xi(z,\eta) := \xi\|\eta\|^2 + V_M\left(\mathcal{L}_M\left(z+\eta\right) - \mathbf{y}\right), \tag{5}$$

where the minimization is performed over elements of the background space $\mathcal{Z}_N := \text{span}\{\zeta_n\}_{n=1}^N \subset \mathcal{X}$ rather than N-dimensional vectors of coefficients associated with the basis $\{\zeta_n\}_{n=1}^N$. In the limit $\xi \to 0^+$, the formulation reduces to ([19, Proposition 2.5.1])

$$(z_\xi^\star, \eta_\xi^\star) := \arg \inf_{(z,\eta)\in\mathcal{Z}_N\times\mathcal{U}} \|\eta\| \quad \text{subject to } \mathcal{L}_M\left(z+\eta\right) = \mathbf{y}, \tag{6}$$

which corresponds to the original formulation proposed in [12]. In the remainder of this paper, we consider (5)–(6) with $V_M(\cdot) = \|\cdot\|_2^2$.

2.3. Choice of $\mathcal{U}$: Variational and user-defined update spaces. We consider two choices for the search space $\mathcal{U}$: (i) $\mathcal{U} = \mathcal{X}$ and (ii) $\mathcal{U} = \text{span}\{\psi_m\}_{m=1}^M$, where $\psi_1, \ldots, \psi_M \in \mathcal{X}$ are linearly independent user-defined functions satisfying the unisolvency condition:

$$\psi \in \mathcal{U}_M,\ \ell_m^o(\psi) = 0, \quad m = 1, \ldots, M \quad \Leftrightarrow \quad \psi \equiv 0. \tag{7}$$

The first choice is considered in the original PBDW papers, while the second choice is proposed here for the first time.

As shown in [19, Proposition 2.2.3] for (5), for $\mathcal{U} = \mathcal{X}$, the update $\eta_\xi^\star$ belongs to the update space $\mathcal{U}_M = \text{span}\{R_\mathcal{X}\ell_m^o\}_{m=1}^M$; we refer to this choice of the update space $\mathcal{U}_M$ (or equivalently of $\mathcal{U}$) as the *variational update*. If $\ell_1^o, \ldots, \ell_M^o$ are linearly independent in $\mathcal{X}'$, it is easy to verify that the space $\mathcal{U}_M$ satisfies (7). Computation of the update space $\mathcal{U}_M$ requires the solution to M variational problems and depends on the choice of the inner product $(\cdot,\cdot)$ of $\mathcal{X}$ and on the observation functionals $\ell_1^o, \ldots, \ell_M^o$.

We refer to the second choice $\mathcal{U}_M = \mathcal{U} = \text{span}\{\psi_m\}_{m=1}^M$ as the *user-defined update*. The functions $\psi_1, \ldots, \psi_M$ are chosen based on approximation considerations to guarantee fast convergence with respect to the number of measurements. We observe that $\mathcal{U}_M$ mildly depends on the observation functionals—it should satisfy (7)—but it does not depend on the choice of the inner product $(\cdot,\cdot)$ of $\mathcal{X}$. In particular, computation of the user-defined update does not in principle require the solution to any variational problem: We might thus rely on user-defined updates to simplify the implementation of the method. It is possible to recover a variational interpretation for this update space for a suitable inner product of $\mathcal{X}$; we discuss this point in section 2.4.

Practical choice of the update space. As explained in the introduction, we are interested in experimental observations that can be modeled by linear functionals of the form

$$\ell_m^o(u) = \ell\left(u, x_m^{\rm obs}, r_{\rm w}\right) = C(x_m^{\rm obs}) \int_\Omega \omega\left(\frac{1}{r_{\rm w}}\|x - x_m^{\rm obs}\|_2\right) u(x)\, dx, \tag{8}$$

where $r_{\rm w}$ reflects the filter width of the transducer, $x_m^{\rm obs}$ denotes the transducer location, and ω describes the local averaging process performed by the experimental device. For this class of observation functionals, we consider updates of the form

$$\mathcal{U}_M = \text{span}\{\phi(\cdot, x_m^{\rm obs})\}_{m=1}^M,$$

where ϕ is here named *update generator*. Note that the variational update corresponds to the choice $\phi^{\rm var}(\cdot, x_m^{\rm obs}) = R_{\mathcal{X}}\ell(\cdot, x_m^{\rm obs}, r_{\rm w})$: We refer to $\phi^{\rm var}$ as the *variational generator*.

If $\mathcal{X}$ is the free space $H^s(\Omega)$ for some $s \geq 1$, we might consider the generator

$$\phi(\cdot, x) = \Phi\left(\|\cdot - x\|_2\right), \tag{9}$$

where $\Phi : \mathbb{R}_+ \to \mathbb{R}_+$ is a positive definite kernel Φ (see, e.g., [25]). Computation of (9) does not require either the offline solution to M variational problems or the storage of the update functions in the grid points associated with the high-fidelity mesh. For this reason, it is considerably less expensive to evaluate and store than its variational counterpart. Furthermore, for small values of $r_{\rm w}$ it leads to superior approximation properties compared to the standard H^1 variational generator, as we will demonstrate in the numerical results.

Another potential choice is to consider

$$\phi(\cdot, x) = R_{\mathcal{X}}\ell(\cdot, x, r_{\rm w}^{\rm ud}) \tag{10}$$

for some $r_{\rm w}^{\rm ud} > r_{\rm w}$. As shown in the numerical results, increasing $r_{\rm w}^{\rm ud}$ smoothens the Riesz representers and improves their approximation properties for smooth fields. Generator (10) requires the solution and the storage to M variational problems during the offline stage (as opposed to (9)); therefore, it has similar costs to the variational update. On the other hand, unlike (10), it permits an easy imposition of homogeneous Dirichlet boundary conditions.[1]

2.4. Variational interpretation of the user-defined space and well-posedness analysis. Given the update space $\mathcal{U}_M \subset \mathcal{X}$, we introduce the interpolation operator $\mathcal{I}_M : \mathcal{X} \to \mathcal{U}_M$ s.t. $\ell_m^o(\mathcal{I}_M(u)) = \ell_m^o(u)$ for all $u \in \mathcal{X}$, $m = 1, \ldots, M$. Then we define the symmetric bilinear form

$$((u, v)) = (u - \mathcal{I}_M(u), v - \mathcal{I}_M(v)) + (\mathcal{I}_M(u), \mathcal{I}_M(v)). \tag{11}$$

The bilinear form (11) and the induced (semi-)norm $|||\cdot||| = \sqrt{((\cdot,\cdot))}$. satisfy the following properties.

LEMMA 2.1. *Let $\mathcal{U}_M = \text{span}\{\psi_m\}_{m=1}^M$ be an M-dimensional space satisfying* (7)*, where $\psi_1, \ldots, \psi_M$ are orthonormal. Then the following hold:*

[1] Homogeneous Dirichlet boundary conditions could be imposed using the generator (9) by adding measurements on the Dirichlet boundary: This approach increases the dimensionality M of the update space, and thus it increases the online cost associated with the computation of the state estimate.

1. $|||u||| = \|u\| \quad \forall\, u \in \mathcal{U}_M$.
2. *Given the orthonormal basis* $\psi_1, \dots, \psi_M$ *of* $\mathcal{U}_M$*, we define* $\mathbb{L}_\eta \in \mathbb{R}^{M,M}$ *such that* $(\mathbb{L}_\eta)_{m,m'} = \ell^o_m(\psi_{m'})$*. Then we can rewrite* (11) *as follows:*

$$((u,v)) = (u - \mathcal{I}_M(u), v - \mathcal{I}_M(v)) + \sum_{m,m'=1}^{M} \ell^o_m(u)\, \ell^o_{m'}(v)\, \mathbb{W}_{m,m'}, \tag{12}$$

   where $\mathbb{W} := \mathbb{L}_\eta^{-T}\mathbb{L}_\eta^{-1}$.
3. *The bilinear form* $((\cdot,\cdot))$ *induces a norm over* $\mathcal{X}$*. More precisely, the following estimate holds:*

$$\frac{1}{2}\|u\|^2 \le |||u|||^2 \le \left(2 + 3\|\mathcal{I}_M\|^2_{\mathcal{L}(\mathcal{X})}\right)\|u\|^2, \quad \|\mathcal{I}_M\|_{\mathcal{L}(\mathcal{X})} = \sup_{v\in\mathcal{X}} \frac{\|\mathcal{I}_M(v)\|}{\|v\|}. \tag{13}$$

4. *Let* $\phi_m = \sum_{p=1}^M (\mathbb{L}_\eta)_{m,p}\psi_p$*. Then* $((\phi_m, v)) = \ell^o_m(v)$ *for all* $v \in \mathcal{X}$.
5. *If* $\mathcal{U}_M = \operatorname{span}\{R_\mathcal{X}\ell^o_m\}_{m=1}^M$*, then* $((\cdot,\cdot)) = (\cdot,\cdot)$.
6. $\Pi^{|||\cdot|||}_{\mathcal{U}_M} u = \mathcal{I}_M(u)$ *for all* $u \in \mathcal{X}$.

Proof. If $u \in \mathcal{U}_M$, $\mathcal{I}_M(u) = u$. Therefore,

$$|||u|||^2 = (u - \mathcal{I}_M(u), u - \mathcal{I}_M(u)) + (\mathcal{I}_M(u), \mathcal{I}_M(u)) = (u,u) = \|u\|^2,$$

which proves the first statement.

Observing that $\mathcal{I}_M(u) = \sum_{m=1}^M (\mathbb{L}_\eta^{-1}\mathcal{L}_M(u))_m \psi_m$ and recalling that $\psi_1, \dots, \psi_M$ are orthonormal, we find

$$\begin{aligned}(\mathcal{I}_M(u), \mathcal{I}_M(v)) &= \sum_{m,m'=1}^{M} (\mathbb{L}_\eta^{-1}\mathcal{L}_M(u))_m (\mathbb{L}_\eta^{-1}\mathcal{L}_M(v))_{m'}\, (\psi_m, \psi_{m'}) \\ &= \sum_{m=1}^{M} (\mathbb{L}_\eta^{-1}\mathcal{L}_M(u))_m (\mathbb{L}_\eta^{-1}\mathcal{L}_M(v))_m = \sum_{p,p'=1}^{M} \ell^o_p(u)\, \ell^o_{p'}(v)\, \mathbb{W}_{p,p'},\end{aligned}$$

which proves the second statement.

We now prove (13) (cf. statement 3). First, we find that

$$\|u\|^2 \le 2\left(\|u - \mathcal{I}_M(u)\|^2 + \|\mathcal{I}_M(u)\|^2\right) = 2|||u|||^2,$$

where the first inequality follows from the triangular inequality and $(a+b)^2 \le 2a^2 + 2b^2$, while the second identity follows directly from the definition of $((\cdot,\cdot))$. Second, exploiting the fact that the interpolation operator is continuous in $\mathcal{X}$, we find

$$|||u|||^2 = \|u - \mathcal{I}_M(u)\|^2 + \|\mathcal{I}_M(u)\|^2 \le 2\|u\|^2 + 3\|\mathcal{I}_M(u)\|^2 \le \left(2 + 3\|\mathcal{I}_M\|^2_{\mathcal{L}(\mathcal{X})}\right)\|u\|^2.$$

Combining the latter two estimates, we obtain (13).

We prove the fourth statement. Since $\phi_m \in \mathcal{U}_M$, $\phi_m - \mathcal{I}_M(\phi_m) \equiv 0$. Therefore,

$$((\phi_m, v)) = \sum_{p,p'=1}^{M} \ell^o_p(\phi_m)\, \ell^o_{p'}(v) \mathbb{W}_{p,p'}.$$

By definition of the interpolation operator, we have $\ell^o_{p'}(v) = \ell^o_{p'}(\mathcal{I}_M(v))$ for $p' = 1, \dots, M$, and $\mathcal{I}_M(v) = \sum_{k'=1}^M \left(\mathbb{L}_\eta^{-1}\mathcal{L}_M(v)\right)_{k'} \psi_{k'}$. Then we find

$$\begin{aligned}((\phi_m, v)) &= \sum_{k,k'=1}^M \sum_{p,p'=1}^M (\mathbb{L}_\eta)_{m,k}\, \ell^o_p(\psi_k)\ell^o_{p'}(\psi_{k'})\mathbb{W}_{p,p'} \left(\mathbb{L}_\eta^{-1}\mathcal{L}_M(v)\right)_{k'} \\ &= \sum_{k,q=1}^M (\mathbb{L}_\eta)_{m,k} \left(\mathbb{L}_\eta^{-1}\right)_{k,q} \ell^o_q(v) = \sum_{q=1}^M \left(\mathbb{L}_\eta \mathbb{L}_\eta^{-1}\right)_{m,q} \ell^o_q(v) = \ell^o_m(v)\end{aligned}$$

for all $v \in \mathcal{X}$. The thesis follows.

In order to prove the fifth statement, we observe that for $\mathcal{U}_M = \operatorname{span}\{R_\mathcal{X}\ell^o_m\}_{m=1}^M$, $\mathcal{I}_M(u) = \Pi_{\mathcal{U}_M} u$. Then, exploiting the properties of $((\cdot,\cdot))$ shown so far and the projection theorem, we find

$$|||u|||^2 = \|u - \mathcal{I}_M(u)\|^2 + \|\mathcal{I}_M(u)\|^2 = \|\Pi_{\mathcal{U}_M^\perp}(u)\|^2 + \|\Pi_{\mathcal{U}_M}(u)\|^2 = \|u\|^2.$$

The proof the sixth statement follows directly from the fourth property:

$$((\Pi_{\mathcal{U}_M}^{|||\cdot|||}u, v)) = ((u,v))\ \forall\, v \in \mathcal{U}_M \ \Leftrightarrow\ \ell^o_m(\Pi_{\mathcal{U}_M}^{|||\cdot|||}u) = \ell^o_m(v),\ m = 1, \dots, M.$$

The latter completes the proof. ∎

Remark 2.1. The norm $\|\mathcal{I}_M\|_{\mathcal{L}(\mathcal{X})}$ is the Lebesgue constant in the $\mathcal{X}$ norm, and it is the inverse of the inf-sup constant ([11, Theorem 2.4]):

$$\gamma_M := \inf_{w \in \mathcal{U}_M} \sup_{v \in \mathcal{W}_M} \frac{(w,v)}{|||w|||\,|||v|||}, \quad \mathcal{W}_M = \operatorname{span}\{R_\mathcal{X}\ell^o_m\}_{m=1}^M.$$

The Lebesgue constant depends on the triple $\mathcal{O} = (\{\ell^o_m\}_{m=1}^M, \mathcal{U}_M, (\mathcal{X}, \|\cdot\|))$. For certain choices of the triple $\mathcal{O}$, it is possible to determine the asymptotic behavior of $\|\mathcal{I}_M\|_{\mathcal{L}(\mathcal{X})}$: To provide a concrete example, we refer to [3, Corollary 1.17] for an important result concerning the behavior of $\|\mathcal{I}_M\|_{\mathcal{L}(\mathcal{X})}$ for one-dimensional polynomial interpolation.

Lemma 2.1 can be exploited to recover an infinite-dimensional variational interpretation of the PBDW statement with a user-defined update. The proof is a straightforward consequence of [19, Proposition 2.2.3] and is here omitted.

PROPOSITION 2.2. *Let $\mathcal{U}_M$ be an M-dimensional space satisfying* (7)*. Then the PBDW solution to* (5) *for $\mathcal{U} = \mathcal{U}_M$ solves the following problem:*

$$(z^\star_\xi, \eta^\star_\xi) := \arg \inf_{(z,\eta) \in \mathcal{Z}_N \times \mathcal{X}} \xi|||\eta|||^2 + \frac{1}{M}\|\mathcal{L}_M(z+\eta) - \mathbf{y}\|_2^2, \tag{14}$$

where $|||\cdot|||$ is the norm induced by the inner product $((\cdot,\cdot))$ of $\mathcal{X}$ defined in (11)*.*

If the inf-sup constant ([12])

$$\beta_{N,M} = \inf_{z \in \mathcal{Z}_N} \sup_{q \in \mathcal{U}_M} \frac{((z,q))}{|||z|||\,|||q|||} \tag{15}$$

is strictly positive, then the solution $(z^\star_\xi, \eta^\star_\xi)$ to (14) *is unique and solves the following saddle-point problem:*

$$\begin{cases} 2\xi((\eta^\star_\xi, q)) + \frac{2}{M}\sum_{m=1}^M \left(\ell^o_m(z^\star_\xi + \eta^\star_\xi) - y_m\right)\ell^o_m(q) = 0 & \forall\, q \in \mathcal{U}_M, \\ ((\eta^\star_\xi, p)) = 0 & \forall\, p \in \mathcal{Z}_N. \end{cases} \tag{16}$$

Furthermore, the estimate $u_\xi^\star$ *belongs to the space*

$$u_\xi^\star \in \mathcal{Z}_N \oplus \mathcal{Z}_N^{\perp,|||\cdot|||} \cap \mathcal{U}_M = \{z+\eta : z \in \mathcal{Z}_N,\ \ \eta \in \mathcal{U}_M,\ \ ((\eta, q)) = 0 \quad \forall q \in \mathcal{Z}_N\},$$

where $\mathcal{Z}_N^{\perp,|||\cdot|||}$ *denotes the orthogonal complement of* $\mathcal{Z}_N$ *with respect to the* $|||\cdot|||$ *norm.*

Remark 2.2. In the variational approach, we first choose an inner product $(\cdot,\cdot)$ for $\mathcal{X}$, and then we appeal to a high-fidelity solver to generate the update space; on the other hand, in the user-defined approach, we first choose the space $\mathcal{U}_M$, and then we use it to define the inner product $((\cdot,\cdot))$ (11). We remark that the second approach is used in the kernel methods literature (see, e.g., [25]) for pointwise measurements. Given a positive definite kernel, it is possible to characterize the Sobolev regularity of the resulting ambient (*native*) space: The characterization of the properties of the ambient space can be then exploited to prove a priori error bounds for the state estimation error and estimate the asymptotic convergence rate as $M \to \infty$. On the other hand, the construction presented in this section does not allow us to characterize the smoothness of the space induced by the norm $|||\cdot|||$ in the limit $M \to \infty$ unless $\|\mathcal{I}_M\|_{\mathcal{L}(\mathcal{X})}$ is bounded for $M \to \infty$.

2.5. Algebraic formulation. Given the fields $z \in \mathcal{Z}_N$ and $\eta \in \mathcal{U}_M$, we introduce the vectors $\mathbf{z} \in \mathbb{R}^N$ and $\boldsymbol{\eta} \in \mathbb{R}^M$ such that $z = \sum_{n=1}^N z_n \zeta_n$ and $\eta = \sum_{m=1}^M \eta_m \psi_m$, where $\zeta_1, \dots, \zeta_N$ and $\psi_1, \dots, \psi_M$ are orthonormal bases (with respect to $\|\cdot\|$) of $\mathcal{Z}_N$ and $\mathcal{U}_M$, respectively. We remark up front that the derivation is independent of the particular construction of the update space. Then we can rewrite (5) as a discrete optimization problem for the coefficients,

$$(\mathbf{z}_\xi^\star, \boldsymbol{\eta}_\xi^\star) = \arg \min_{(\mathbf{z},\boldsymbol{\eta}) \in \mathbb{R}^N \times \mathbb{R}^M} \ \xi \|\boldsymbol{\eta}\|_2^2 \ + \ \frac{1}{M} \|\mathbb{L}_\eta \boldsymbol{\eta} + \mathbb{L}_z \mathbf{z} - \mathbf{y}\|_2^2, \tag{17}$$

where $\mathbb{L}_z \in \mathbb{R}^{M,N}$ and $\mathbb{L}_\eta \in \mathbb{R}^{M,M}$ are given by $(\mathbb{L}_z)_{m,n} = \ell_m^o(\zeta_n)$ and $(\mathbb{L}_\eta)_{m,m'} = \ell_m^o(\psi_{m'})$. We observe that (7) implies that $\mathbb{L}_\eta$ is invertible. By differentiating the objective function with respect to $\mathbf{z}$ and $\boldsymbol{\eta}$, we obtain

$$\begin{cases} (\xi M \mathbb{I} + \mathbb{L}_\eta^T \mathbb{L}_\eta) \boldsymbol{\eta}_\xi^\star + \mathbb{L}_\eta^T \mathbb{L}_z \boldsymbol{z}_\xi^\star = \mathbb{L}_\eta^T \boldsymbol{y}, \\ \qquad\qquad\qquad \mathbb{L}_z^T \mathbb{L}_\eta \boldsymbol{\eta}_\xi^\star + \mathbb{L}_z^T \mathbb{L}_z \boldsymbol{z}_\xi^\star = \mathbb{L}_z^T \boldsymbol{y}. \end{cases} \tag{18}$$

If we multiply $(18)_1$ by $\mathbb{L}_z^T \mathbb{L}_\eta^{-T}$ and exploit $(18)_2$, we obtain

$$\xi M \mathbb{L}_\eta^T \mathbb{L}_\eta^{-T} \boldsymbol{\eta}_\xi^\star + \mathbb{L}_z^T \mathbb{L}_\eta \boldsymbol{\eta}_\xi^\star + \mathbb{L}_z^T \mathbb{L}_z \boldsymbol{z}_\xi^\star = \mathbb{L}_z^T \boldsymbol{y} = \mathbb{L}_z^T \mathbb{L}_\eta \boldsymbol{\eta}_\xi^\star + \mathbb{L}_z^T \mathbb{L}_z \boldsymbol{z}_\xi^\star \Rightarrow \mathbb{L}_z^T \mathbb{L}_\eta^{-T} \boldsymbol{\eta}_\xi^\star = \mathbf{0}.$$

Finally, if we introduce $\tilde{\boldsymbol{\eta}}_\xi^\star = \mathbb{L}_\eta^{-T} \boldsymbol{\eta}_\xi^\star$ and multiply $(18)_1$ by $\mathbb{L}_\eta^{-T}$, we obtain

$$\begin{cases} \left(\xi M \mathbb{I} + \mathbb{L}_\eta \mathbb{L}_\eta^T\right) \tilde{\boldsymbol{\eta}}_\xi^\star + \mathbb{L}_z \boldsymbol{z}_\xi^\star = \boldsymbol{y}, \\ \qquad\qquad\qquad \mathbb{L}_z^T \tilde{\boldsymbol{\eta}}_\xi^\star = \mathbf{0}, \end{cases} \qquad \boldsymbol{\eta}_\xi^\star = \mathbb{L}_\eta^T \tilde{\boldsymbol{\eta}}_\xi^\star. \tag{19}$$

System (19) can be used to efficiently compute the PBDW solution.

Remark 2.3 (spectrum of the PBDW system). Following the argument in [2, section 3.4], we observe that the saddle-point system (19) is congruent to the block-diagonal matrix

$$\begin{bmatrix} \mathbb{L}_\eta \mathbb{L}_\eta^T + \xi M \mathbb{I} & 0 \\ 0 & -\mathbb{L}_z^T (\mathbb{L}_\eta \mathbb{L}_\eta^T + \xi M \mathbb{I})^{-1} \mathbb{L}_z \end{bmatrix}.$$

If the update satisfies (7), $\mathbb{L}_\eta$ is invertible, and thus $\mathbb{L}_\eta \mathbb{L}_\eta^T + \xi M \mathbb{I}$ is positive definite for all $\xi \geq 0$.

On the other hand, the second block $-\mathbb{L}_z^T(\mathbb{L}_\eta\mathbb{L}_\eta^T+\xi M\mathbb{I})^{-1}\mathbb{L}_z$ is negative definite if and only if the rank of $\mathbb{L}_z$ is equal to N: This condition is equivalent to the positivity of $\beta_{N,M}$ in Proposition 2.2.

In view of the analysis of the method and of the definition of the greedy procedure for the selection of the transducers' location, the next proposition provides a computable expression for the inf-sup constant $\beta_{N,M}$ defined in (15) and for the norm $\|\mathcal{I}_M\|_{\mathcal{L}(\mathcal{X})}$ in (13).

PROPOSITION 2.3. *Suppose that $\zeta_1,\dots,\zeta_N$ and $\psi_1,\dots,\psi_M$ are orthonormal in $\|\cdot\|$. The inf-sup constant $\beta_{N,M}$ is the square root of the minimum eigenvalue of the eigenproblem*

$$\mathbb{L}_z^T\mathbb{W}\mathbb{L}_z\,\mathbf{z}_n=\nu_n\left(\mathbb{I}+2\mathbb{L}_z^T\mathbb{W}\mathbb{L}_z-2\mathrm{sym}\left(\mathbb{C}^T\mathbb{L}_\eta^{-1}\mathbb{L}_z\right)\right)\mathbf{z}_n,\quad n=1,\dots,N, \tag{20}$$

while the norm $\|\mathcal{I}_M\|_{\mathcal{L}(\mathcal{X})}$ is the inverse of the square root of the minimum eigenvalue of the following eigenproblem:

$$\mathbb{L}_\eta^T\mathbb{K}^{-1}\mathbb{L}_\eta\boldsymbol{\eta}_m=\lambda_m\boldsymbol{\eta}_m,\quad m=1,\dots,M. \tag{21}$$

Here, $\mathbb{W}=\mathbb{L}_\eta^{-T}\mathbb{L}_\eta^{-1}$, $\mathrm{sym}(A)=\frac{1}{2}(A+A^T)$, while the matrices $\mathbb{C}\in\mathbb{R}^{M,N}$ and $\mathbb{K}\in\mathbb{R}^{M,M}$ are given by $\mathbb{C}_{m,n}=(\psi_m,\zeta_n)$ and $\mathbb{K}_{m,m'}=(R_{\mathcal{X}}\ell_m^o,R_{\mathcal{X}}\ell_{m'}^o)$.

Proof. Given $z=\sum_{n=1}^N z_n\zeta_n$, $\eta=\sum_{m=1}^M\eta_m\psi_m$, we can write the interpolation operator as $\mathcal{I}_M(z)=\sum_{k=1}^M(\mathbb{L}_\eta^{-1}\mathbb{L}_z\mathbf{z})_k\;\psi_k$. Then we obtain

$$((z,\eta))=\mathbf{z}^T\mathbb{L}_z^T\mathbb{W}\mathbb{L}_\eta\boldsymbol{\eta},\quad |||\eta|||^2=\boldsymbol{\eta}^T\boldsymbol{\eta},$$

and

$$|||z|||^2=\|z\|^2+\|\mathcal{I}_M(z)\|^2-2(z,\mathcal{I}_M(z))+\mathbf{z}^T\mathbb{L}_z^T\mathbb{W}\mathbb{L}_z\boldsymbol{z}=\mathbf{z}^T\,\mathbb{B}\,\mathbf{z},$$

where $\mathbb{B}=\mathbb{I}+2\mathbb{L}_z^T\mathbb{W}\mathbb{L}_z-2\mathrm{sym}(\mathbb{C}^T\mathbb{L}_\eta^{-1}\mathbb{L}_z)$. Recalling the definition of $\mathbb{W}$, we find

$$\beta_{N,M}^2=\inf_{\mathbf{z}\in\mathbb{R}^N}\sup_{\boldsymbol{\eta}\in\mathbb{R}^M}\frac{\left(\mathbf{z}^T\mathbb{L}_z^T\mathbb{W}\mathbb{L}_\eta\boldsymbol{\eta}\right)^2}{(\boldsymbol{\eta}^T\boldsymbol{\eta})\,\mathbf{z}^T\mathbb{B}\mathbf{z}}=\inf_{\mathbf{z}\in\mathbb{R}^N}\frac{\mathbf{z}^T\mathbb{L}_z^T\mathbb{W}\mathbb{L}_z\mathbf{z}}{\mathbf{z}^T\mathbb{B}\mathbf{z}}.$$

Introducing the Lagrangian multiplier ν, we can write the optimality conditions as

$$\begin{cases}\mathbb{L}_z^T\mathbb{W}\mathbb{L}_z\mathbf{z}=\nu\mathbb{B}\mathbf{z},\\ \mathbf{z}^T\,\mathbb{B}\mathbf{z}=1.\end{cases}$$

The thesis follows.

The proof of (21) exploits Remark 2.1 and the same argument used to prove (20). We omit the details. ∎

2.6. `SGreedy+approximation` algorithm for the selection of the observation centers. Algorithm 1 summarizes the greedy procedure for the selection of the transducers' locations $x_1^{\mathrm{obs}},\dots,x_M^{\mathrm{obs}}$. During the first (stability) stage, we maximize the constant $\beta_{N,M}$ in a greedy manner. During the second (approximation) stage, we minimize the fill distance $h_M:=\sup_{x\in\Omega}\min_{m=1,\dots,M}\|x-x_m^{\mathrm{obs}}\|_2$ in a greedy manner. We switch from the first to the second stage when the inf-sup constant $\beta_{N,M}$ is larger than an user-defined constant *tol*: If $tol=0$, all the centers are selected through the approximation loop; if $tol=1$, all centers are selected through the stability loop. Representative values for *tol* used in the numerical simulations are $tol\in[0.2,0.6]$.

Since $((\cdot,\cdot))$ and the induced norm $|||\cdot|||$ depend in general on the iteration m (unless ϕ is the variational generator), we use here the subscript m.

Algorithm 1 Greedy stability-approximation balancing (`SGreedy+approx.`) algorithm

Input	$\mathcal{Z}_N = \text{span}\{\zeta_n\}_{n=1}^N$	background space
	M	number of sensors
	$tol > 0$	threshold for the stability constant
	$\phi : \Omega \times \Omega \to \mathbb{R}$	update generator
Output	$\mathcal{U}_M$	update space

Stability

1: Compute $x_1^{\text{obs}} := \arg\max_{x\in\bar{\Omega}} |\zeta_1(x)|$, $\mathcal{U}_1 = \text{span}\{\phi(\cdot, x_1^{\text{obs}})\}$, $m = 1$
2: **while** $m \le M$ **do**
3: Compute $\beta_{N,m} = \min_{w\in\mathcal{Z}_N} \max_{v\in\mathcal{U}_m} \frac{((w,v))_m}{|||w|||_m |||v|||_m}$.
4: **if** $\beta_{N,m} \le tol$ **then**
5: Compute $z_{\min,m} := \arg\min_{z\in\mathcal{Z}_N} \max_{v\in\mathcal{U}_m} \frac{((z,v))_m}{|||z|||_m |||v|||_m}$.
6: Compute $x_{m+1}^{\text{obs}} := \arg\max_{x\in\bar{\Omega}} |z_{\min,m}(x) - \mathcal{I}_m(z_{\min,m})(x)|$.
7: Set $\mathcal{U}_{m+1} = \mathcal{U}_m \cup \text{span}\{\phi(\cdot, x_{m+1}^{\text{obs}})\}$, $m = m + 1$.
8: **else**
9: `Break`
10: **end if**
11: **end while**

Approximation

1: **while** $m \le M$ **do**
2: Compute $x_{m+1}^{\text{obs}} := \arg\max_{x\in\bar{\Omega}} \min_{m'=1,\dots,m} \|x - x_{m'}^{\text{obs}}\|_2$.
3: Set $\mathcal{U}_{m+1} = \mathcal{U}_m \cup \text{span}\{\phi(\cdot, x_{m+1}^{\text{obs}})\}$, $m = m + 1$.
4: **end while**
5: Define $((\cdot,\cdot)) := ((\cdot,\cdot))_M$ and $|||\cdot||| := |||\cdot|||_M$.

The greedy procedure was first presented in [19, Algorithm 3.2.1] for the variational update space. Note that the choice of the update (variational or user-defined) enters in the algorithm through the choice of the generator ϕ and thus influences (i) the computation of the new element of the basis (cf. line 7) and (ii) the computation of the inf-sup constant $\beta_{N,M}$ (cf. line 3) based on (20). Provided that M is significantly smaller than the dimension $\mathcal{N}$ of the high-fidelity discrete space, the cost is dominated by the computation of the M Riesz representers. As a result, the use of the RBF generator (9) significantly reduces the cost compared to the variational generator $\phi(\cdot, x) = R_{\mathcal{X}}\ell(\cdot, x, r_{\text{w}})$ and to (10). The stability loop—for variational update—corresponds[2] to the SGreedy algorithm proposed in [12]; on the other hand, the strategy for the approximation step is strongly related to the so-called *farthest-first traversal* approach to the minimax facility location problem (see, e.g., [14]), first

[2]In the SGreedy procedure listed in [12, Algorithm 2], steps 6 and 7 of Algorithm 1 are replaced by $\ell^o_{m+1} = \arg\min_{\ell\in\mathcal{L}} |\ell(z_{\min,m} - \mathcal{I}_m(z_{\min,m}))|$ and $\mathcal{U}_{m+1} = \mathcal{U}_m \cup \text{span}\{R_{\mathcal{X}}\ell^o_{m+1}\}$, where $\mathcal{L} = \{\ell(\cdot, x, r_{\text{w}}) : x \in \overline{\Omega}\}$ is the dictionary of available functionals. The approach presented here is easier to implement and less computationally expensive. Furthermore, the two procedures are asymptotically equivalent for $r_{\text{w}} \to 0^+$ and return similar results if r_{w} is small enough compared to the characteristic length-scale of the elements in $\mathcal{Z}_N$.

proposed by Rosenkrantz, Stearns, and Lewis in [16]. For the variational update, since $\beta_{N,M}$ is a nondecreasing function of M for a fixed value of N (see [12]), the stability constant remains above the threshold during the approximation stage. On the other hand, for the user-defined update, there is in general no guarantee that the inf-sup constant will be above the threshold at the end of the greedy procedure since the inner product $((\cdot,\cdot))$ and the induced norm vary with m. In the numerical experiments we investigate the behavior of the inf-sup constant for large values of M.

2.7. Choice of ξ. The choice of ξ is performed using holdout validation. We consider two mutually exclusive data sets $\mathcal{D}_M^{\text{train}} = \{(\ell_m^o, y_m)\}_{m=1}^M$ and $\mathcal{D}_I^{\text{val}} = \{(\ell_i^o, y_i)\}_{i=1}^I$; given the finite-dimensional search space $\Xi \subset \mathbb{R}_+$, we then select $\xi^\star$ such that

$$\xi^\star = \arg\min_{\xi\in\Xi} \widehat{\text{MSE}}(I) := \frac{1}{I}\sum_{i=1}^{I} \left(y_i - \ell_i^o(u_\xi^\star)\right)^2,$$

where $u_{\bar\xi}^\star$ is the PBDW solution based on the training data set $\mathcal{D}_M^{\text{train}}$, for $\xi = \bar\xi$.

In this work, we consider validation measurements of the form $\{y_i = \ell(u, x_i^{\text{obs}}, r_{\text{w}}) + \epsilon_i\}_{i=1}^I$, where ℓ is introduced in (8) and $x_1^{\text{obs}}, \dots, x_I^{\text{obs}}$ are independent realizations of an uniformly distributed random variable over Ω, $X \sim \text{Uniform}(\Omega)$. If r_{w} is small, for this choice of the observation centers, and assuming that $\epsilon_1, \dots, \epsilon_I$ are independent realizations of $\epsilon \sim (0, \sigma^2)$, it is possible to show that (see [21])

$$\mathbb{E}\left[\widehat{\text{MSE}}(I)\right] = \frac{1}{|\Omega|}\|u^{\text{true}} - u_\xi^\star\|_{L^2(\Omega)}^2 + \sigma^2 + C(r_{\text{w}}, u^{\text{true}} - u_\xi^\star),$$

where $C(r_{\text{w}}, u^{\text{true}} - u_\xi^\star) \to 0$ as $r_{\text{w}} \to 0^+$ if $\nabla(u^{\text{true}} - u_\xi^\star) \in L^q(\Omega)$, $q > d$. Therefore, for sufficiently large values of I, the validation procedure approximately minimizes the L^2 error.

The choice of the number of validation measurements is a trade-off between experimental cost (given by the number of transducers dedicated to validation) and reliability of the validation procedure. κ-fold cross validation (see, e.g., [8, Chapter 7] and [9]) reduces the experimental cost by partitioning the training data set $\mathcal{D}_M$ into κ equal-sized subsamples (folds) $\{\mathcal{D}_M^{(k)}\}_{k=1}^\kappa$. Of the κ folds, a single fold is retained for testing, and the remaining $\kappa - 1$ folds are used for training. The procedure is then repeated κ times with each of the κ folds used once as the validation data set. After having selected the hyperparameter ξ, the entire data set is finally used for training. We emphasize that cross validation relies on the assumption that the pairs (x_m^{obs}, y_m) are independently sampled from the joint distribution of inputs and outputs (*random design*). On the other hand, in our framework the training centers $\{x_m^{\text{obs}}\}_{m=1}^M$ are chosen deterministically (*fixed design*). For this reason, in this work, we simply consider holdout validation with $I = M/2$; we refer to a future work for more advanced cross-validation strategies.

2.8. Extension to vector-valued problems. We can trivially extend the greedy procedure for the selection of the observation centers to vector-valued fields $u^{\text{true}} = [u_1^{\text{true}}, \dots, u_J^{\text{true}}] \in \mathcal{X} = \mathcal{X}(\Omega; \mathbb{R}^J)$. Assuming that each transducer is able to measure all J components of the true field in all d directions, at each iteration of the greedy algorithm, we select x_m^{obs} such that (compare with Algorithm 1, step 6)

$$x_{m+1}^{\text{obs}} = \arg\max_{x\in\overline{\Omega}} \|z_{\min}(x) - \mathcal{I}_m(z_{\min})(x)\|_2.$$

Then we add J modes $\phi_1(\cdot, x_{m+1}^{\text{obs}}), \dots, \phi_J(\cdot, x_{m+1}^{\text{obs}}) \in \mathcal{X}$ in the update; note that the subspace $\mathcal{U}_m$ is $m \times J$ dimensional.

As regards the choice of the J generators, we might consider the generator

$$\phi_i(\cdot,x) = R_{\mathcal{X}}\ell_i(\cdot,x,r_{\mathrm{w}}^{\mathrm{ud}}), \quad \text{where } \ell_i(u,x,r_{\mathrm{w}}^{\mathrm{ud}}) = \ell(u_i,x,r_{\mathrm{w}}^{\mathrm{ud}}), \quad i=1,\dots,J, \tag{22a}$$

and $r_{\mathrm{w}}^{\mathrm{ud}} \geq r_{\mathrm{w}}$. Alternatively, we might consider the update

$$\phi_i(\cdot,x) = \Phi(\|\cdot - x\|_2)\mathbf{e}_i, \quad i=1,\dots,J, \tag{22b}$$

where Φ is a properly defined kernel and $\mathbf{e}_1,\dots,\mathbf{e}_J$ are the canonical vectors in $\mathbb{R}^J$.

Note that in the case of incompressible flows considered in section 5, for a proper choice of the ambient space $\mathcal{X}$, the update space (22a) is divergence-free, while the user-defined update space (22b) is not divergence-free. For this reason, we might resort to matrix-valued divergence-free kernels (see, e.g., [26]). We refer to a future work for the use of matrix-valued divergence-free kernels in the PBDW framework.

3. Error analysis. We present an error analysis for general linear functionals in the presence of noise, based on the identification of two sources: The first source is related to the finite number of measurements, while the second source is related to the presence of noise. For simplicity of notation, we consider scalar problems; however, the analysis can be trivially extended to vector-valued fields. Given the measurements $\{y_m = \ell_m^o(u^{\mathrm{true}}) + \epsilon_m\}_{m=1}^M$, we define the vector $\mathbf{u}^{\mathrm{opt}} = [\tilde{\boldsymbol{\eta}}^{\mathrm{opt}}, \mathbf{z}^{\mathrm{opt}}] \in \mathbb{R}^{M+N}$ corresponding to the solution $u_{\xi=0}^{\mathrm{opt}}$ to (6) fed with perfect observations $\{y_m^{\mathrm{true}} = \ell_m^o(u^{\mathrm{true}})\}_{m=1}^M$ (see (19)). On the other hand, we define $\mathbf{u}_\xi^\star = [\tilde{\boldsymbol{\eta}}_\xi^\star, \mathbf{z}_\xi^\star]$ corresponding to the solution $u_\xi^\star$ to the stabilized formulation (5) fed with imperfect observations.

We first present the result for perfect measurements.

PROPOSITION 3.1. *Let $y_m = \ell_m^o(u^{\mathrm{true}})$, $m=1,\dots,M$; let $\beta_{N,M} > 0$; and let $\mathcal{U}_M$ satisfy* (7)*. We denote by $u_{\xi=0}^{\mathrm{opt}}$ the solution to the PBDW formulation* (6)*. Then the following estimates hold:*

$$\left|\!\left|\!\left| u^{\mathrm{true}} - u_{\xi=0}^{\mathrm{opt}} \right|\!\right|\!\right| \leq \frac{1}{\beta_{N,M}} \inf_{z\in\mathcal{Z}_N} \inf_{q\in\mathcal{U}_M\cap\mathcal{Z}_N^\perp, |||\cdot|||} \left|\!\left|\!\left| u^{\mathrm{true}} - z - q \right|\!\right|\!\right| \tag{23a}$$

$$\|u^{\mathrm{true}} - u_{\xi=0}^{\mathrm{opt}}\| \leq \frac{\sqrt{4+6\|\mathcal{I}_M\|^2_{\mathcal{L}(\mathcal{X})}}}{\beta_{N,M}} \inf_{z\in\mathcal{Z}_N} \inf_{q\in\mathcal{U}_M\cap\mathcal{Z}_N^\perp, |||\cdot|||} \|u^{\mathrm{true}} - z - q\|. \tag{23b}$$

Proof. Thanks to the variational interpretation of the user-defined update (see Proposition 2.2), proof of (23a) follows directly from [4, Remark 2.12]. On the other hand, exploiting (23a) and (13), we find

$$\begin{aligned}\|u^{\mathrm{true}} - u_{\xi=0}^{\mathrm{opt}}\| &\leq \sqrt{2}\left|\!\left|\!\left| u^{\mathrm{true}} - u_{\xi=0}^{\mathrm{opt}} \right|\!\right|\!\right| \leq \frac{\sqrt{2}}{\beta_{N,M}} \inf_{z\in\mathcal{Z}_N} \inf_{q\in\mathcal{U}_M\cap\mathcal{Z}_N^\perp, |||\cdot|||} \left|\!\left|\!\left| u^{\mathrm{true}} - z - q \right|\!\right|\!\right| \\ &\leq \frac{\sqrt{4+6\|\mathcal{I}_M\|^2_{\mathcal{L}(\mathcal{X})}}}{\beta_{N,M}} \inf_{z\in\mathcal{Z}_N} \inf_{q\in\mathcal{U}_M\cap\mathcal{Z}_N^\perp, |||\cdot|||} \|u^{\mathrm{true}} - z - q\|,\end{aligned}$$

which proves (23b). ∎

If we introduce the matrices and vectors

$$\mathbb{A}(\xi) = \begin{bmatrix} \left(\xi M\mathbb{I} + \mathbb{L}_\eta\mathbb{L}_\eta^T\right) & \mathbb{L}_z \\ \mathbb{L}_z^T & 0 \end{bmatrix}; \quad \mathbf{e} = \begin{bmatrix} \boldsymbol{\epsilon} \\ \mathbf{0} \end{bmatrix}; \quad \boldsymbol{\epsilon} = [\epsilon_1,\dots,\epsilon_M]^T,$$

we obtain the following identity that links $\mathbf{u}_\xi^\star$ to $\mathbf{u}^{\rm opt}$:

$$\mathbb{A}(\xi)\mathbf{u}_\xi^\star = \mathbb{A}(0)\mathbf{u}^{\rm opt} + \mathbf{e} = \mathbb{A}(\xi)\mathbf{u}^{\rm opt} - \xi M \begin{bmatrix} \tilde{\boldsymbol{\eta}}^{\rm opt} \\ \mathbf{0} \end{bmatrix} + \mathbf{e}. \tag{24}$$

Identity (24) can be used to estimate the bias of our estimate and the expected error. For simplicity, we present below estimates for the errors in the coefficients, $\mathbf{u}_\xi^\star - \mathbf{u}^{\rm opt}$; the estimates can also be extended to compute the state estimation error in an integral norm of interest. We omit the details.

PROPOSITION 3.2. *Suppose that $\epsilon_1, \ldots, \epsilon_M$ are independent realizations of $\epsilon \sim (0, \sigma^2)$. Then the following hold:*

$$\|\mathbb{E}[\mathbf{u}_\xi^\star] - \mathbf{u}^{\rm opt}\|_2 \le \frac{\xi M}{s_{\min}(\mathbb{A}(\xi))} \|\tilde{\boldsymbol{\eta}}^{\rm opt}\|_2 \tag{25}$$

and

$$\mathbb{E}\left[\|\mathbf{u}_\xi^\star - \mathbf{u}^{\rm opt}\|_2^2\right] \le \left(\frac{\xi M}{s_{\min}(\mathbb{A}(\xi))} \|\tilde{\boldsymbol{\eta}}^{\rm opt}\|_2\right)^2 + \sigma^2 \mathrm{trace}\left(\mathbb{A}(\xi)^{-1}\Sigma\mathbb{A}(\xi)^{-T}\right), \tag{26}$$

where $s_{\min}(\mathbb{A}(\xi))$ denotes the minimum singular value of $\mathbb{A}(\xi)$, $\Sigma = \left[\begin{smallmatrix} \mathbb{I} & 0 \\ 0 & 0 \end{smallmatrix}\right]$.

Proof. Exploiting (24), we find

$$\mathbb{A}(\xi)\mathbf{u}_\xi^\star = \mathbb{A}(\xi)\mathbf{u}^{\rm opt} - \xi M \begin{bmatrix} \tilde{\boldsymbol{\eta}}^{\rm opt} \\ \mathbf{0} \end{bmatrix} + \mathbf{e} \;\Rightarrow\; \|\mathbb{E}[\mathbf{u}_\xi^\star] - \mathbf{u}^{\rm opt}\|_2 \le \frac{\xi M}{s_{\min}(\mathbb{A}(\xi))} \|\tilde{\boldsymbol{\eta}}^{\rm opt}\|_2,$$

which is (25).

In order to show (26), we first observe that (cf. [15, Theorem C, Chapter 14.4])

$$\mathbb{E}\left[\|\mathbb{E}[\mathbf{u}_\xi^\star] - \mathbf{u}_\xi^\star\|_2^2\right] = \sigma^2 \mathrm{trace}\left(\mathbb{A}(\xi)^{-1}\Sigma\mathbb{A}(\xi)^{-T}\right).$$

Then we find

$$\begin{aligned}\mathbb{E}\|\mathbf{u}_\xi^\star - \mathbf{u}^{\rm opt}\|_2^2 &= \|\mathbb{E}[\mathbf{u}_\xi^\star] - \mathbf{u}^{\rm opt}\|_2^2 + \mathbb{E}\left[\|\mathbb{E}[\mathbf{u}_\xi^\star] - \mathbf{u}_\xi^\star\|_2^2\right] \\ &\le \left(\frac{\xi M}{s_{\min}(\mathbb{A}(\xi))} \|\tilde{\boldsymbol{\eta}}^{\rm opt}\|_2\right)^2 + \sigma^2 \mathrm{trace}\left(\mathbb{A}(\xi)^{-1}\Sigma\mathbb{A}(\xi)^{-1}\right),\end{aligned}$$

which is (26). ∎

We observe that the bound for the mean squared error (26) is the sum of two contributions: The former—$\frac{\xi M}{s_{\min}(\mathbb{A}(\xi))} \|\tilde{\boldsymbol{\eta}}^{\rm opt}\|_2$—involves the accuracy of the background space, measured by $\|\tilde{\boldsymbol{\eta}}^{\rm opt}\|_2$ and is monotonically increasing with ξ; the latter—$\sigma^2 \mathrm{trace}(\mathbb{A}(\xi)^{-1}\Sigma\mathbb{A}(\xi)^{-T})$—involves the accuracy of the measurements and is monotonically decreasing with ξ according to our numerical experience. Therefore, the optimal value of ξ depends on the ratio between "measurement inaccuracy" and "model inaccuracy" $\frac{\sigma}{\|\tilde{\boldsymbol{\eta}}^{\rm opt}\|_2}$: If $u^{\rm true} \in \mathcal{Z}_N$, estimate (26) suggests to pick $\xi \to \infty$; if $\sigma = 0$, estimate (26) suggests to pick $\xi \to 0^+$. Since in practice estimates of the ratio $\frac{\sigma}{\|\tilde{\boldsymbol{\eta}}^{\rm opt}\|_2}$ are rarely available, our analysis suggests that the value of ξ should be chosen adaptively using (cross-)validation techniques such as the one discussed in section 2.7. These observations are in agreement with [18, Remark 3.3].

Remark 3.1 (extension to complex-valued problems). Proposition 3.2 can be trivially extended to the complex-valued case. Assuming that measurements are of the

form

$$y_m = \ell_m^o(u^{\text{true}}) + \sigma_{\text{re}}\epsilon_m^{\text{re}} + \mathrm{i}\sigma_{\text{im}}\epsilon_m^{\text{im}}, \qquad \epsilon_m^{\text{re}}, \epsilon_m^{\text{im}} \overset{\text{i.i.d.}}{\sim} (0,1), \quad \sigma_{\text{re}}, \sigma_{\text{im}} > 0,$$

we find the same estimate for the bias (25), and we find

$$\begin{aligned}\mathbb{E}[\|\mathbf{u}_\xi^\star - \mathbf{u}^{\text{opt}}\|_2^2] \le \left(\frac{\xi M}{s_{\min}(\tilde{\mathbb{A}}(\xi))}\|\widetilde{\boldsymbol{\eta}}^{\text{opt}}\|_2\right)^2 + \sigma_{\text{re}}^2 \text{trace}\left(\tilde{\mathbb{A}}(\xi)^{-1}\Sigma^{\text{re}}\tilde{\mathbb{A}}(\xi)^{-T}\right) \\ + \sigma_{\text{im}}^2 \text{trace}\left(\tilde{\mathbb{A}}(\xi)^{-1}\Sigma^{\text{im}}\tilde{\mathbb{A}}(\xi)^{-T}\right),\end{aligned} \tag{27}$$

where $\tilde{\mathbb{A}}(\xi) = \begin{bmatrix} \text{Re}[\mathbb{A}(\xi)] & -\text{Im}[\mathbb{A}(\xi)] \\ \text{Im}[\mathbb{A}(\xi)] & \text{Re}[\mathbb{A}(\xi)] \end{bmatrix}$, $\Sigma^{\text{re}} = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}$, $\Sigma^{\text{im}} = \begin{bmatrix} 0 & 0 \\ 0 & \Sigma \end{bmatrix}$.

4. Application to acoustics. We illustrate the behavior of the PBDW formulation presented in this paper through the vehicle of two acoustic Helmholtz problems. Since the objective of this work is to propose user-defined update spaces that improve convergence with respect to the number of observations (M-convergence), we do not study the effect of the primary approximation provided by the background $\mathcal{Z}_N$: We refer the reader to the PBDW literature for further results concerning the effect of the primary approximation.

4.1. A two-dimensional model problem. The model problem is the same considered in [12, section 3], and [18, section 5].

4.1.1. Problem definition. Given the domain $\Omega = (0,1)^2$, we define the acoustic model problem

$$\begin{cases} -(1+\mathrm{i}\epsilon\mu)\,\Delta u_g(\mu) - \mu^2 u_g(\mu) = \mu\left(2x_1^2 + e^{x_2}\right) + \mu g & \text{in } \Omega, \\ \partial_n\, u_g(\mu) = 0 & \text{on } \partial\Omega, \end{cases} \tag{28}$$

where i is the imaginary unit, $\mu > 0$ is the wave number, $\epsilon = 10^{-2}$ is a fixed dissipation, and $g \in L^2(\Omega)$ is a bias term that will be specified later. Here, the parameter $\mu > 0$ constitutes the anticipated, parametric uncertainty in the system, which might model our uncertainty in the speed of sound, while the function g constitutes the unanticipated (nonparametric) uncertainty in the system.

To assess the performance of the PBDW formulation for various configurations, we define the true field u^{true} as the solution to (28) for some $\mu^{\text{true}} \in \mathcal{P}^{\text{bk}}$ and for the following choice of the *bias* g:

$$g := 0.5(e^{-x_1} + 1.3\cos(1.3\pi x_2)). \tag{29a}$$

On the other hand, we define the bk manifold as

$$\mathcal{M}^{\text{bk}} := \{u_{g=0}(\mu) : \mu \in \mathcal{P}^{\text{bk}}\}. \tag{29b}$$

To measure performance, we introduce the relative L^2 and H^1 errors averaged over $|\mathcal{P}_{\text{test}}^{\text{bk}}| = n_{\text{test}}$ fields associated with different choices of the parameter μ:

$$E_{\text{avg}}^{\text{rel}}(n_{\text{test}}) := \frac{1}{n_{\text{test}}} \sum_{\mu \in \mathcal{P}_{\text{test}}^{\text{bk}}} \frac{\|u^{\text{true}}(\mu) - u_\xi^\star(\mu)\|_\star}{\|u^{\text{true}}(\mu)\|_\star}, \qquad \|\cdot\|_\star = \|\cdot\|_{L^2(\Omega)} \text{ or } \|\cdot\|_{H^1(\Omega)}. \tag{30}$$

In all our tests, we consider $n_{\text{test}} = 10$ equispaced parameters in $\mathcal{P}^{\text{bk}}$.

We model the (synthetic) observations by a Gaussian convolution with standard deviation $r_{\rm w}$,

(31a)
$$\ell_m^o(v) = \text{Gauss}(v, x_m^{\rm obs}; r_{\rm w}) = C(x_m^{\rm obs}) \int_\Omega e^{-\frac{1}{2r_{\rm w}^2}\|x - x_m^{\rm obs}\|_2^2} v(x)\, dx, \quad m = 1, \dots, M,$$

where $C(x_m^{\rm obs})$ is a normalization constant such that $\ell_m^o(1) = 1$ and $x_m^{\rm obs}$ denotes the transducer location. In order to simulate noisy measurements, we corrupt the synthetic measurements with homoscedastic random noise,

(31b)
$$y_\ell = \ell_m^o(u^{\rm true}) + \sigma^{\rm re}\epsilon_\ell^{\rm re} + \mathrm{i}\sigma^{\rm im}\epsilon_\ell^{\rm im}, \qquad \epsilon_\ell^{\rm re}, \epsilon_\ell^{\rm im} \overset{\rm iid}{\sim} \mathcal{N}(0,1),$$

where

(31c)
$$\sigma^{\rm re} = \text{SNR} \times \text{std}\left(\{\text{Re}[\ell_m^o(u^{\rm true})]\}_{m=1}^M\right), \quad \sigma^{\rm im} = \text{SNR} \times \text{std}\left(\{\text{Im}[\ell_m^o(u^{\rm true})]\}_{m=1}^M\right),$$

and SNR denotes the signal-to-noise ratio in the measurements.

4.1.2. PBDW spaces. We introduce the ambient space $\mathcal{X} = H^1(\Omega)$ endowed with the inner product:

(32)
$$(u, v) = \int_\Omega u\bar{v} + \nabla u \cdot \nabla \bar{v}\, dx.$$

Note that $\bar{(\cdot)}$ denotes the complex conjugate of $(\cdot)$. The background space $\mathcal{Z}_N$ is built using the weak-greedy algorithm: We refer the reader to [12] for further details.

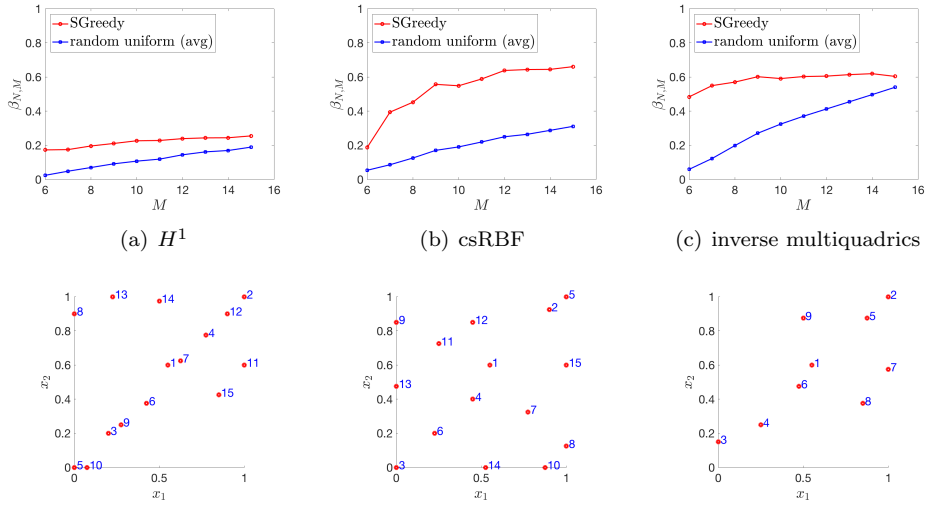
As regards the update space, we consider the variational update associated with the inner product (32) and the user-defined updates $\mathcal{U}_M = \text{span}\{\phi(\lambda\|\cdot - x_m^{\rm obs}\|_2)\}_{m=1}^M$ for

$$\begin{cases} \phi(r) = \dfrac{1}{(1+r^2)^2} & \text{inverse multiquadrics;} \\ \phi(r) = (1-r)_+^4(4r+1) & \text{csRBF.} \end{cases}$$

The observation centers $\{x_m^{\rm obs}\}_{m=1}^M$ are chosen according to Algorithm 1. As regards the choice of the kernel scale, we here set $\lambda = 1$ for inverse multiquadrics and $\lambda = 2$ for csRBF. We remark that the optimal value for the kernel scale parameter λ strongly depends on the number of measurements and also on the characteristic length-scale of the field $u^{\rm true} - z_\xi^\star$; therefore, adaptation of the parameter λ might improve performance, particularly for large values of M.

4.1.3. Numerical results. In Figures 1(a), (b), and (c), we show the behavior of the inf-sup constant $\beta_{N,M}$ with M for the three different choices of the update space (and thus for the three choices of the norm $|||\cdot|||$), $r_{\rm w} = 0.01$ and $N = 6$, and for centers $\{x_m^{\rm obs}\}_{m=1}^M$ selected by the SGreedy algorithm with $tol = 0.6$; to measure performance, we compare the results with the ones obtained using randomly generated centers. For the second choice, we average over 35 different random choices of the M centers. We observe that the SGreedy selection of the sensor locations leads to a more stable formulation compared to randomly generated points. In Figures 1(d), (e), and (f), we show the centers $\{x_m^{\rm obs}\}_{m=1}^M$ selected by the SGreedy algorithm for the three different choices of the update space: For all the three choices of the update space the Greedy procedure selects most points along the principal diagonal $(0,0) \to (1,1)$.

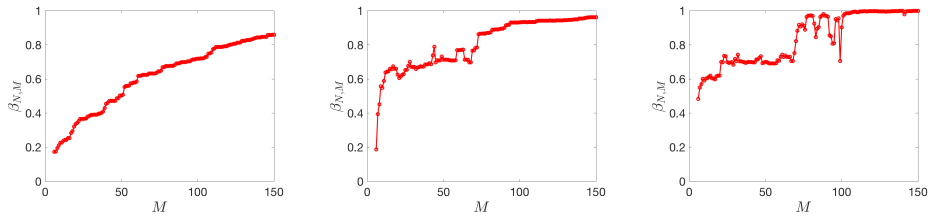
In Figure 2, we show the behavior of $\beta_{N=6,M}$ for three choices of the update space considered, for $M = 6, \dots, 150$. We observe that, for the user-defined update, $\beta_{N,M}$ is not monotonic increasing with M. Nevertheless, we do not observe pathologic behaviors of the inf-sup constant as M increases.



(a) H^1 (b) csRBF (c) inverse multiquadrics

(d) H^1 (e) csRBF (f) inverse multiquadrics

FIG. 1. *Application to a two-dimensional acoustic problem: application of the SGreedy algorithm* ($N = 6, r_{\mathrm{w}} = 0.01, tol = 0.6$). *Results for random centers are averaged over* 35 *random trials.*



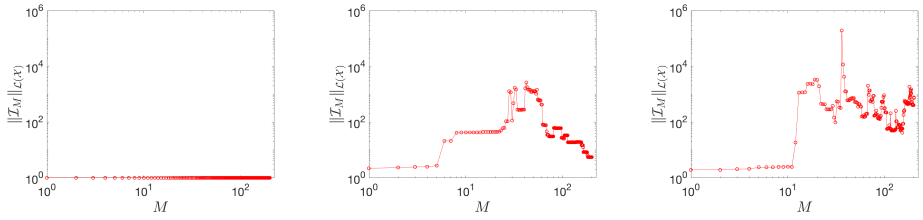
(a) H^1 (b) csRBF (c) inverse multiquadrics

FIG. 2. *Application to a two-dimensional acoustic problem: behavior of* $\beta_{N,M}$ *for* $M = 6, \ldots, 150$, *for three choices of the update* ($N = 6, r_{\mathrm{w}} = 0.01, tol = 0.6$).

In Figure 3, we show the behavior of the Lebesgue constant $\|\mathcal{I}_M\|_{\mathcal{L}(\mathcal{X})}$ defined in (13) for the three choices of the update. As expected for the variational update space, the Lebesgue constant is equal to one. We also observe that $\|\mathcal{I}_M\|_{\mathcal{L}(\mathcal{X})}$ is significantly larger for inverse multiquadrics than for csRBF: Recalling estimate (23b), the use of inverse multiquadrics (and more in general of smooth kernels) is appropriate only for smooth fields.

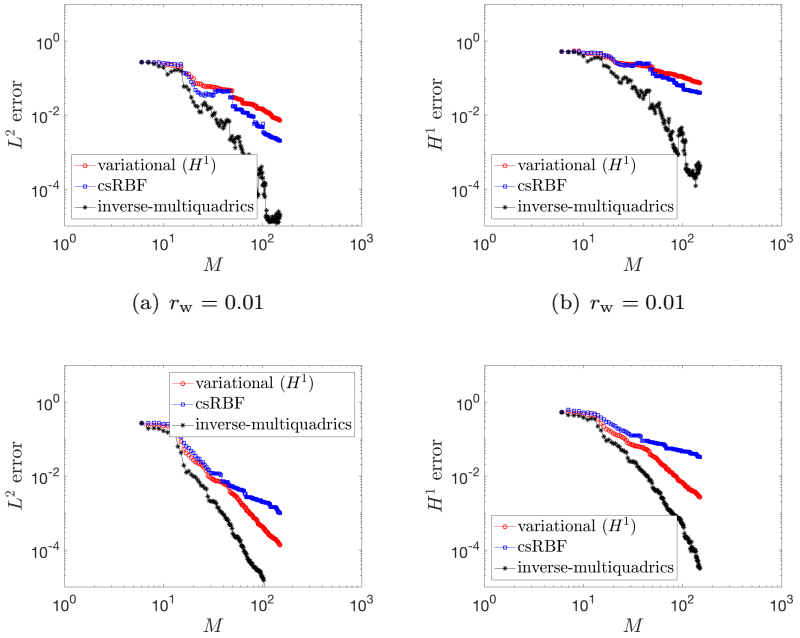
Figure 4 shows the behavior of the relative error $E_{\mathrm{avg}}^{\mathrm{rel}}$ (30) with M for perfect observations, $N = 6$, and two values of r_{w} in (31). We observe that the user-defined update based on inverse multiquadrics leads to more accurate state estimates compared to the standard H^1 PBDW, particularly for $r_{\mathrm{w}} = 0.01$. We further observe that inverse multiquadrics outperform csRBF in all tests considered: Since the true state is smooth in Ω, the inverse multiquadric kernel—which is C^∞—exhibits superior approximation properties compared to the C^2 csRBF kernel (see [25, Chapter 11]).

Figure 5 shows results for the behavior of the L^2 relative error $E_{\mathrm{avg}}^{\mathrm{rel}}$ with M for imperfect observations for the user-defined (inverse multiquadrics) update and for the variational H^1 update. The error $E_{\mathrm{avg}}^{\mathrm{rel}}$ is averaged over 25 realizations of the



(a) H^1 (b) csRBF (c) inverse multiquadrics

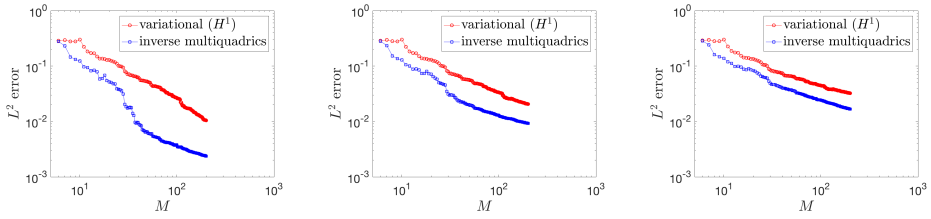
FIG. 3. *Application to a two-dimensional acoustic problem: behavior of* $\|\mathcal{I}_M\|_{\mathcal{L}(\mathcal{X})}$ ($N = 6, r_{\mathrm{w}} = 0.01, tol = 0.6$).



(a) $r_{\mathrm{w}} = 0.01$ (b) $r_{\mathrm{w}} = 0.01$

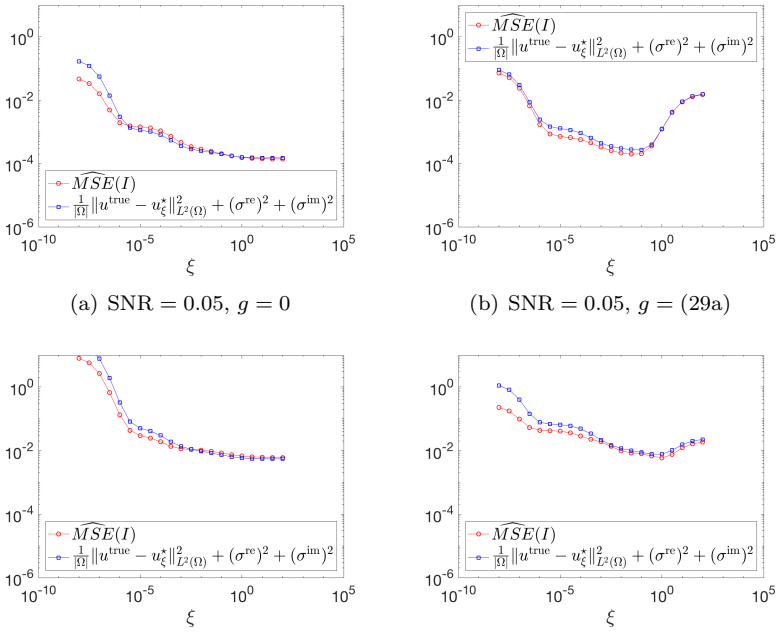
(c) $r_{\mathrm{w}} = 0.05$ (d) $r_{\mathrm{w}} = 0.05$

FIG. 4. *Application to a two-dimensional acoustic problem: M convergence for three different choices of the update space for perfect observations* ($N = 6, \mathrm{SNR} = 0$).



(a) SNR = 0.01 (b) SNR = 0.05 (c) SNR = 0.1

FIG. 5. *Application to a two-dimensional acoustic problem: M convergence for three different choices of the update space for noisy observations* ($N = 4, r_{\mathrm{w}} = 0.01$). *Results are averaged over* 25 *realizations of the random disturbances for each value of M.*



(a) SNR $= 0.05$, $g = 0$ (b) SNR $= 0.05$, $g =$ (29a)

(c) SNR $= 0.3$, $g = 0$ (d) SNR $= 0.3$, $g =$ (29a)

FIG. 6. *Application to a two-dimensional acoustic problem: interpretation of ξ. Behavior of $\widehat{\text{MSE}}(I)$ and of the error $\frac{1}{|\Omega|}\|u^{\text{true}} - u_\xi^\star\|^2_{L^2(\Omega)} + (\sigma^{\text{re}})^2 + (\sigma^{\text{im}})^2$ for two different values of the bias g and for two different values of* SNR. *($M = 100, I = 50, N = 5, r_{\text{w}} = 0.01$, empirical update based on inverse multiquadrics.)*

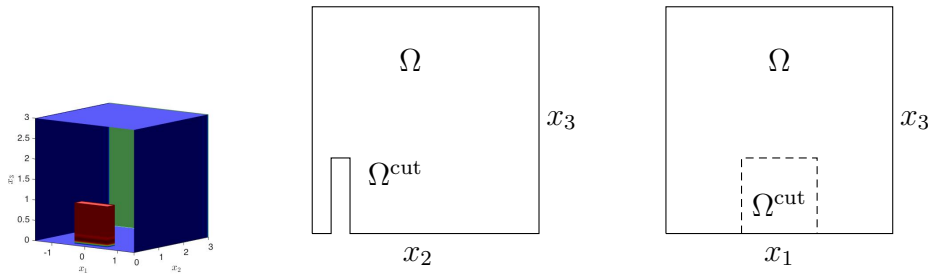
homoscedastic random noise. More in detail, we here consider three different noise levels SNR; we set $N = 4$, and we consider $r_{\text{w}} = 0.01$ in (31). Furthermore, in order to select the value of ξ, we resort to the holdout validation procedure outlined in section 2.7 based on additional $I = M/2$ measurements associated with uniformly generated points in Ω. The rate of convergence with M seems to weakly depend on the choice of the update; nevertheless, also for noisy measurements, the use of inverse multiquadrics strongly improves performance.

In Figure 6 we discuss the interpretation of the regularization parameter ξ. Towards this end, we show the behavior of the mean squared error $\widehat{\text{MSE}}(I)$ introduced in section 2.7 for two different choices of the true field ($u^{\text{true}} = u_g(\mu = 5.8)$ and $u^{\text{true}} = u_{g=0}(\mu = 5.8)$), two different noise levels, and $N = 5$, $M = 100$, $I = 50$. Note that for $g = 0$, u^{true} belongs to the bk manifold. We observe that the optimal value of ξ increases as the noise increases and decreases as the best-fit error increases: This is in good agreement with the discussion in section 3 and also with the results in [18, Figure 4] for pointwise measurements. We further observe that, as anticipated in section 2.7, the error estimator is in good qualitative agreement with the true error $\|u^{\text{true}} - u_\xi^\star\|^2_{L^2(\Omega)}$; therefore, it can be used to select the optimal value of ξ.

4.2. A three-dimensional acoustic problem.

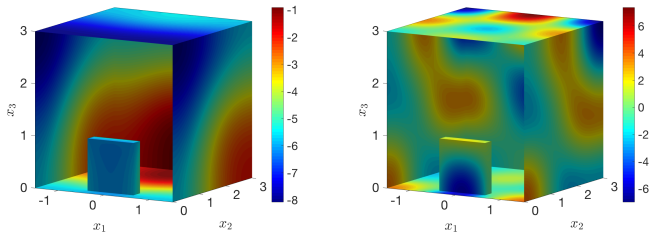
4.2.1. Problem definition. We consider the three-dimensional model problem

$$\begin{cases} -(1+\epsilon i)\Delta u_g(\mu) - (2\pi\mu)^2 u_g(\mu) = g & \text{in } \Omega, \\ \partial_n u_g(\mu) = 0 & \text{on } \partial\Omega, \end{cases} \tag{33}$$

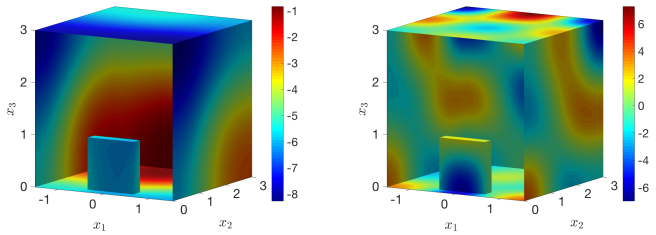


(a) (b) (c)

FIG. 7. *Application to a three-dimensional acoustic problem: computational domain.*



(a) $\text{Re}(u^{\text{bk}})$ $\mu = 0.1$ (b) $\text{Re}(u^{\text{bk}})$ $\mu = 0.5$



(c) $\text{Re}(u^{\text{true}})$ $\mu = 0.1$ (d) $\text{Re}(u^{\text{true}})$ $\mu = 0.5$

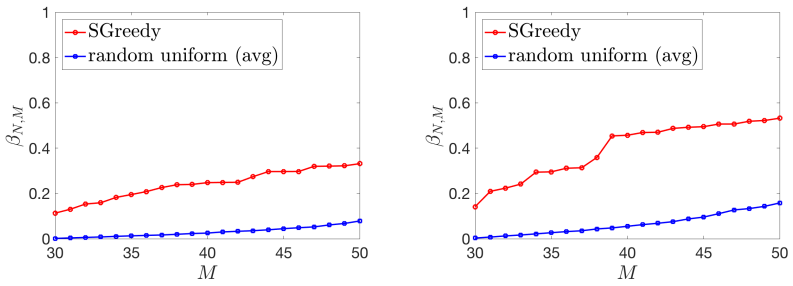
FIG. 8. *Application to a three-dimensional acoustic problem: visualization of bk and true fields.*

where $\epsilon = 10^{-2}$, $\Omega = (-1.5, 1.5) \times (0, 3) \times (0, 3) \setminus \Omega^{\text{cut}}$, and $\Omega^{\text{cut}} = (-0.5, 0.5) \times (0.25, 0.5) \times (0, 1)$. Figure 7 shows the geometry. In this example, we consider the bk manifold $\mathcal{M}^{\text{bk}} = \{u^{\text{bk}}(\mu) = u_{g^{\text{bk}}}(\mu) : \mu \in \mathcal{P}^{\text{bk}} = [0.1, 0.5]\}$, and we define the true field as the solution to (33) for some $\mu^{\text{true}} \in \mathcal{P}^{\text{bk}}$ and $g = g^{\text{true}}$, where

$$g^{\text{bk}}(x) = 10\, e^{-\|x - p^{\text{bk}}\|_2^2}, \quad g^{\text{true}}(x) = 10\, e^{-\|x - p^{\text{true}}\|_2^2},$$

and $p^{\text{bk}} = [0, 2, 1]$, $p^{\text{true}} = [-0.1, 2, 1]$. Parameterized uncertainty in the system models uncertainty in the input frequency μ; nonparametric or unanticipated uncertainty is here associated with the incorrect location of the acoustic source ($p^{\text{bk}} \neq p^{\text{true}}$). Computations are based on a P2 FE discretization with roughly $\mathcal{N} = 16000$ degrees of freedom in Ω. Figure 8 shows the bk and true solutions for two values of μ.

As in the previous example, we model the synthetic observations by a Gaussian convolution with standard deviation r_{w}; see (31). For simplicity, in the tests below, we only consider perfect measurements. Furthermore, we measure performance by computing the relative L^2 and H^1 errors $E^{\text{rel}}_{\text{avg}}$ (30) for $n_{\text{test}} = 10$ different choices of the parameter μ in $\mathcal{P}^{\text{bk}}$.



(a) H^1 (b) inverse multiquadrics

FIG. 9. *Application to a three-dimensional acoustic problem: application of the SGreedy algorithm* ($N = 30$, $r_{\mathrm{w}} = 0.02$, *tol* $= 0.4$). *Results for random centers are averaged over* 35 *random trials.*

4.2.2. PBDW spaces. We consider the ambient space $\mathcal{X} = H^1(\Omega)$ endowed with the inner product $(\cdot,\cdot)$ (32). Furthermore, the background space $\mathcal{Z}_N$ is built using the weak-greedy algorithm based on the residual.

As regards the update space, we consider the variational update associated with the inner product (32) and the user-defined updates $\mathcal{U}_M = \mathrm{span}\{\phi(\lambda\|\cdot - x_m^{\mathrm{obs}}\|_2)\}_{m=1}^M$ for $\phi(r) = \frac{1}{1+r^2}$ (inverse multiquadrics) and $\lambda = 1$. The observation centers $\{x_m^{\mathrm{obs}}\}_{m=1}^M$ are chosen according to Algorithm 1.

4.2.3. Numerical results. In Figure 9, we perform the same test of Figure 1. We compare the behavior of the inf-sup constant $\beta_{N=30,M}$ with M using the SGreedy procedure and using randomly generated centers. For the second choice, we average over 35 different random choices of the M centers. As in the previous example, the greedy procedure improves the stability of the PBDW formulation.

Figure 10 shows the behavior of the relative error $E_{\mathrm{avg}}^{\mathrm{rel}}$ (30) with M for perfect observations, $N = 30$, and $r_{\mathrm{w}} = 0.02$ in (31). As in the previous case, the user-defined update based on inverse multiquadrics leads to more accurate state estimates compared to the standard H^1 PBDW.

5. Application to fluid mechanics.

5.1. Problem statement. We consider the following model problem:[3]

(34)
$$\begin{cases} -\dfrac{1}{\mathrm{Re}}\nabla\cdot\left(\nabla u_g + \nabla u_g^T\right) + (u_g\cdot\nabla)u_g + \nabla p_g = 0 & \text{in } \Omega \\ \nabla\cdot u_g = 0 & \text{in } \Omega \\ p_g\mathbf{n} - \frac{1}{\mathrm{Re}}(\nabla u_g + \nabla u_g^T)\mathbf{n} = 0 & \text{on } \Gamma_{\mathrm{out},1}\cup\Gamma_{\mathrm{out},2} \\ u_g = g\mathbf{e}_1 & \text{on } \Gamma_{\mathrm{in}} \\ u_g = 0 & \text{on } \partial\Gamma_{\mathrm{hom}} = \Omega\setminus(\Gamma_{\mathrm{in}}\cup\Gamma_{\mathrm{out}}), \end{cases}$$

where the domain Ω is depicted in Figure 11(a). We then define the bk and true manifolds as

$$\mathcal{M}^{\mathrm{bk}} = \{u_g(\mathrm{Re}) : \mathrm{Re}\in[50,350],\ \ g(x_2) = 4(1-x_2)\,x_2\}$$

[3]The configuration is similar to the one considered in [7, section 5]; however, while here we consider a two-dimensional domain, in [7] the authors consider a more realistic axi-symmetric model.

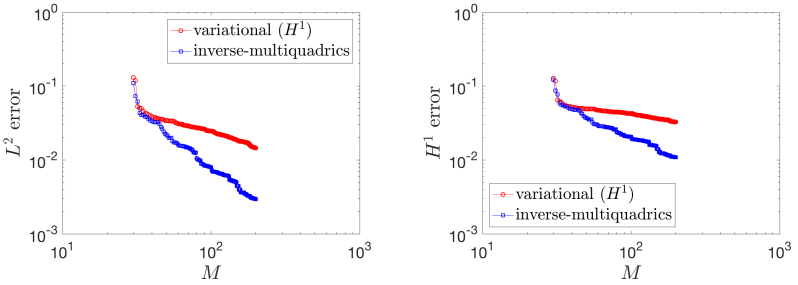


FIG. 10. *Application to a three-dimensional acoustic problem: M convergence for two different choices of the update space for perfect observations* ($N = 30$, $r_{\mathrm{w}} = 0.02$, $tol = 0.4$).

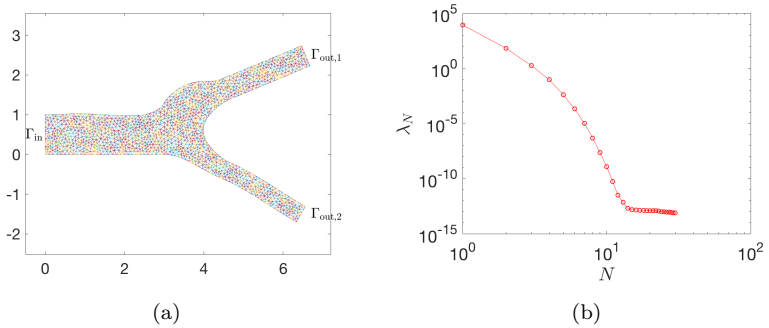


FIG. 11. *Application to fluid dynamics.* (a): *computational domain;* (b): *behavior of the POD eigenvalues* ($n_{\mathrm{train}} = 100$).

and

$$\mathcal{M}^{\mathrm{true}} = \{u_g(\mathrm{Re}) : \mathrm{Re} \in [50, 350], \quad g(x_2) = 4(1 - x_2)x_2\,(1 + 0.1\sin(2\pi x_2))\}.$$

Here, uncertainty in $\mu = \mathrm{Re}$ constitutes the anticipated (parametric) uncertainty in the system, while uncertainty in the inflow condition is the unanticipated (nonparametric) uncertainty.

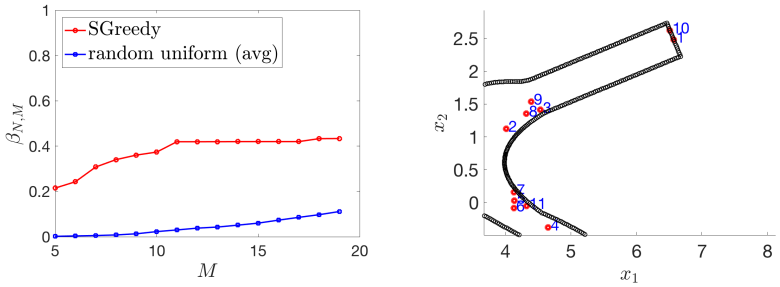
As in the previous examples, we model synthetic observations by a Gaussian convolution with standard deviation $r_{\mathrm{w}} = 0.01$. Furthermore, we resort to a conforming Taylor–Hood P3-P2 finite element discretization with $\mathcal{N}_{\mathrm{u}} = 24038$ degrees of freedom.

5.2. PBDW spaces. We consider the ambient space $\mathcal{X} = \{v \in [H_{0,\Gamma_{\mathrm{hom}}}(\Omega)]^2 : \nabla \cdot v = 0\}$ endowed with the inner product:

$$(u, v) = \int_\Omega \nabla u : \nabla v + u \cdot v \, dx.$$

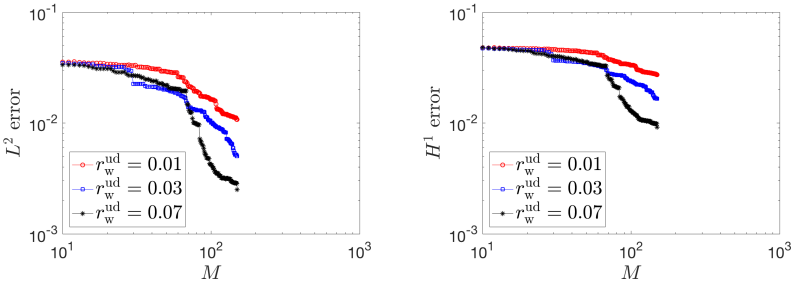
The background space is built using POD based on the $(\cdot, \cdot)$ inner product. Figure 11(b) shows the behavior of the POD eigenvalues. On the other hand, we consider the update generators,

$$\phi_i(\cdot, x) = R_{\mathcal{X}}\, \ell_i(\cdot, x, r_{\mathrm{w}}^{\mathrm{ud}}), \quad \ell_i(u, x, r_{\mathrm{w}}^{\mathrm{ud}}) = \ell(u_i, x, r_{\mathrm{w}}^{\mathrm{ud}}), \qquad i = 1, 2, \tag{35}$$



(a) (b)

FIG. 12. *Application of the SGreedy + approx algorithm.* (a): *behavior of the inf-sup constant* $\beta_{N,M}$ *with* M *for the user-defined update with* $r_{\mathrm{w}}^{\mathrm{ud}} = 0.05$*;* (b): *centers* $\{x_m^{\mathrm{obs}}\}_m$ *selected by the SGreedy algorithm* ($N = 5$, $r_{\mathrm{w}} = 0.01$, $tol = 0.3$).



(a) (b)

FIG. 13. M*-convergence. Behavior of the relative averaged error* $E_{\mathrm{avg}}^{\mathrm{rel}}$ *over* $n_{\mathrm{test}} = 10$ *different true fields in* $\mathcal{M}^{\mathrm{true}}$ ($N = 5$, $r_{\mathrm{w}} = 0.01$, $tol = 0.3$).

for several values of $r_{\mathrm{w}}^{\mathrm{ud}} \geq r_{\mathrm{w}}$. We observe that the update space is divergence-free; furthermore, it satisfies the no-slip boundary conditions on Γ_{hom}.

5.3. Numerical results. In Figure 12, we show the performance of the SGreedy procedure for $N = 5$ and $r_{\mathrm{w}}^{\mathrm{ud}} = 0.05$ ($tol = 0.3$). Figure 12(a) shows the behavior of the inf-sup constant $\beta_{N,M}$ with M for the user-defined update with $r_{\mathrm{w}}^{\mathrm{ud}} = 0.05$: We observe that the SGreedy procedure outperforms on average the random uniform selector. Figure 12(b) shows the location of the points selected by the SGreedy (stabilization stage) algorithm. We observe that SGreedy selects points in the boundary layer that develops near the junction: This can be explained by recalling that the Reynolds number—the parameter associated with the solution manifold—deeply affects the thickness of the boundary layer.

Figure 13 shows the behavior of the relative error $E_{\mathrm{avg}}^{\mathrm{rel}}$ (30) with M for perfect measurements, $N = 5$ and three values of $r_{\mathrm{w}}^{\mathrm{ud}}$ in (35). We observe that for $r_{\mathrm{w}}^{\mathrm{ud}} = 0.01$ the user-defined update corresponds to the variational update. As for the previous test cases, the use of an user-defined update strongly improves performance, particularly for moderate to large values of M.

6. Conclusions. In this paper, we presented a number of extensions to the PBDW formulation for state estimation. First, we proposed a Tikhonov regularization

of the original PBDW statement for general linear functionals, which relies on holdout validation, to systematically deal with noisy measurements. Second, we proposed user-defined update spaces, which guarantee rapid convergence with respect to the number of measurements M and also might not require the solution to M Riesz problems. Third, we presented an a priori error analysis that provides insights into the role of the regularization hyperparameter ξ associated with the penalized formulation.

Numerical results suggested that user-defined update spaces increase the flexibility of the methodology and can be exploited to improve performance. We considered two classes of generators—(9) and (10)—to build the update for local measurements. The kernel-based generator (9) guarantees fast convergence with M for smooth fields and is inexpensive to compute. Generator (10) is beneficial in terms of approximation and is well suited to deal with strong boundary conditions, but it does not reduce offline costs compared to the H^1 variational generator.

We identify a number of future research directions to improve the PBDW formulation and extend its range of applications. First, we wish to extend the formulation to different models for the experimental noise and also to probabilistic background. Towards this end, we wish to rely on the connection between PBDW and PSM established in [18] and recapped in section 2.2. Second, we wish to extend the formulation to time-dependent problems. This might be accomplished by exploiting a space-time variational formulation [22] to incorporate the space-time structure of the evolution equations in the empirical expansion.

REFERENCES

[1] A. F. Bennett, *Inverse Modeling of the Ocean and Atmosphere*, Cambridge University Press, Cambridge, 2002.

[2] M. Benzi, G. H. Golub, and J. Liesen, *Numerical solution of saddle point problems*, Acta Numer., 14 (2005), pp. 1–137.

[3] C. Bernardi, Y. Maday, and F. Rapetti, *Discrétisations Variationnelles de Problemes aux Limites Elliptiques*, Vol. 45, Springer, Berlin, 2004.

[4] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk, *Data assimilation in reduced modeling*, SIAM/ASA J. Uncertain. Quantif., 5 (2017), pp. 1–29.

[5] D. G. Cacuci, I. M. Navon, and M. Ionescu-Bujor, *Computational Methods for Data Evaluation and Assimilation*, CRC Press, Boca Raton, FL, 2013.

[6] A. Cohen and R. DeVore, *Approximation of high-dimensional parametric PDEs*, Acta Numer., 24 (2015), pp. 1–159.

[7] M. D'Elia, M. Perego, and A. Veneziani, *A variational data assimilation procedure for the incompressible Navier-Stokes equations in hemodynamics*, J. Sci. Comput., 52 (2012), pp. 340–359.

[8] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Vol. 2, Springer, Berlin, 2009.

[9] R. Kohavi et al., *A study of cross-validation and bootstrap for accuracy estimation and model selection*, in International Joint Conference on Artificial Intelligence (IJCAI), Vol. 14, 1995, pp. 1137–1145.

[10] A. C. Lorenc, *Analysis methods for numerical weather prediction*, Roy. Meteorol. Soc. Quart. J., 112 (1986), pp. 1177–1194.

[11] Y. Maday, O. Mula, A. Patera, and M. Yano, *The generalized empirical interpolation method: Stability theory on Hilbert spaces with an application to the Stokes equation*, Comput. Methods Appl. Mech. Engrg., 287 (2015), pp. 310–334.

[12] Y. Maday, A. T. Patera, J. D. Penn, and M. Yano, *A parameterized-background data-weak approach to variational data assimilation: Formulation, analysis, and application to acoustics*, Internat. J. Numer. Methods Engrg., 102 (2015), pp. 933–965.

[13] Y. Maday, A. T. Patera, J. D. Penn, and M. Yano, *PBDW state estimation: Noisy observations; configuration-adaptive background spaces; physical interpretations*, ESAIM Proc. Surveys, 50 (2015), pp. 144–168.

[14] S. H. Owen and M. S. Daskin, *Strategic facility location: A review*, Eur. J. Oper. Res., 111 (1998), pp. 423–447.
[15] J. A. Rice, *Mathematical Statistics and Data Analysis*, Duxbury Press, Pacific Grove (2007).
[16] D. J. Rosenkrantz, R. E. Stearns, and P. M. Lewis, II, *An analysis of several heuristics for the traveling salesman problem*, SIAM J. Comput., 6 (1977), pp. 563–581.
[17] G. Rozza, D. P. Huynh, and A. T. Patera, *Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations*, Arch. Comput. Methods Eng., 15 (2008), pp. 229–275.
[18] T. Taddei, *An adaptive parametrized-background data-weak approach to variational data assimilation*, ESAIM Math. Model. Numer. Anal., 51 (2017), pp. 1827–1858.
[19] 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.
[20] 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.
[21] T. Taddei, J. D. Penn, and A. T. Patera, *Validation by Monte Carlo sampling of experimental observation functionals*, Internat. J. Numer. Methods Engrg., (2017).
[22] K. Urban and A. T. Patera, *An improved error bound for reduced basis approximation of linear parabolic problems*, Math. Comp., 83 (2014), pp. 1599–1615.
[23] S. Volkwein, *Model reduction using proper orthogonal decomposition*, Lecture notes, Institute of Mathematics and Scientific Computing, University of Graz, http://www.uni-graz.at/imawww/volkwein/POD.pdf (2011).
[24] G. Wahba, *Spline Models for Observational Data*, Vol. 59, SIAM, Philadelphia, 1990.
[25] H. Wendland, *Scattered Data Approximation*, Vol. 17, Cambridge University Press, Cambridge, 2004.
[26] H. Wendland, *Divergence-free kernel methods for approximating the Stokes problem*, SIAM J. Numer. Anal., 47 (2009), pp. 3158–3179.