3D-VAR for parameterized partial differential equations: a certified reduced basis approach

Nicole Aretz-Nellesen¹ · Martin A. Grepl² · Karen Veroy¹,3

Received: 1 October 2018 / Accepted: 8 July 2019 /
Published online: 25 July 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract
In this paper, we propose a reduced order approach for 3D variational data assimilation governed by parameterized partial differential equations. In contrast to the classical 3D-V AR formulation that penalizes the measurement error directly, we present a modified formulation that penalizes the experimentally observable misfit in the measurement space. Furthermore, we include a model correction term that allows to obtain an improved state estimate. We begin by discussing the influence of the measurement space on the amplification of noise and prove a necessary and sufficient condition for the identification of a “good” measurement space. We then propose a certified reduced basis (RB) method for the estimation of the model correction, the state prediction, the adjoint solution, and the observable misfit with respect to the true state for real-time and many-query applications. A posteriori bounds are proposed for the error in each of these approximations. Finally, we introduce different approaches for the generation of the reduced basis spaces and the stability-based selection of measurement functionals. The 3D-V AR method and the associated certified reduced basis approximation are tested in a parameter and state estimation problem for a steady-state thermal conduction problem with unknown parameters and unknown Neumann boundary conditions.

Keywords  Variational data assimilation · 3D-V AR · Model correction · Reduced basis method · State estimation · Parameter estimation · A posteriori error estimation

Mathematics Subject Classification (2010) 49K20 · 65K10 · 65M15 · 65M60

Communicated by: Anthony Nouy

Martin A. Grepl
grepl@igpm.rwth-aachen.de

Extended author information available on the last page of the article.
1 Introduction

In numerical simulations, mathematical models—such as ordinary or partial differential equations (PDEs)—are widely used to predict the state or behavior of a physical system. The goal of variational data assimilation is to improve state predictions through the incorporation of measurement data, e.g., experimental observations, into the mathematical model. Variational data assimilation is prevalent in meteorology [22–24] and oceanography [3, 4], for example, in weather forecasting and ocean circulation modeling. Prominent examples include the 3D- and 4D-V AR methods, which weigh deviations from a prior best-knowledge (initial) state against differences with respect to measurement data; see the recent texts [21, 32] and references therein for a discussion of variational data assimilation.

Whereas 4D-V AR considers dynamical systems (i.e., three space dimensions plus time) and usually aims to estimate the (unknown) initial condition of the system, 3D-V AR considers the stationary case. In this paper, we propose a certified reduced order approach for a modified 3D-V AR method for parameterized, coercive PDEs. Compared with the classical 3D-V AR method, our modified formulation penalizes the experimentally observable misfit in the measurement space instead of the difference in the measurements. Furthermore, we account for an imperfect model by introducing a model bias in the formulation, similar to the weak-constraint 4D-V AR approach [38].

The proposed method makes data-informed modifications to a best-knowledge background model to generate a compromise between the original model and the observed measurements. The method thereby accounts for the physical integrity of the state prediction and can be used to estimate unknown model properties and parameters along with the state. The 3D-V AR problem—and variational data assimilation in general—is usually cast as an optimization problem and has very close connections to optimal control theory [39]. In addition, the optimality condition has a saddle-point structure which can be analyzed using standard functional analysis arguments. Based on these observations, we can identify necessary and sufficient properties of the measurement space that increase the stability of the 3D-V AR formulation and provide a practical procedure for the generation of a measurement space meeting these conditions.

This paper builds upon the model-data weak approach presented in [41], the parameterized-background data-weak (PBDW) approach to variational data assimilation introduced in [27, 28], and the certified reduced order approach for 4D-V AR in [18]. The model-data weak formulation in [41] finds a state estimate by minimizing the distance between state and observed state while penalizing model corrections. This method indirectly accepts any kind of model modification in the dual space and can thereby account for unpredicted behavior. In the PBDW approach to variational data assimilation introduced in [27, 28], a state estimate is obtained by projecting an observation from the measurement space onto a space featuring model properties, such as an RB space that approximates the solution manifold of a parameterized PDE. The PBDW framework has recently been extended with adaptive [29, 36] and localization [35] strategies and has been recast in a multispace setting [5]. In addition to the discussions in the PBDW literature, the optimal selection of measurements
with greedy orthogonal matching pursuit (OMP) algorithms in state estimation has been analyzed recently in [6]. We build upon this work in the construction of our measurement space.

The connection and distinctions between 3D-VAR and the PBDW formulation has already been discussed in [27, 36]. We also note that the PBDW approach is related to the generalized empirical interpolation method [25, 26] and gappy-POD [40] (see [27, 36] for a thorough discussion). We will show that—under certain assumptions on the spaces and in the limit of the regularization parameter going to infinity—our modified 3D-VAR formulation is equivalent to the PBDW formulation. However, there are also decisive differences between the two formulations. We also note that [30] considers the proper generalized decomposition (PGD) method for real-time parameter estimation in the best-knowledge model. Although the method allows for deviations from the best-knowledge states if desired, they are restricted to the reduced basis directions only and hence do not account for more general physically motivated model modifications.

After a brief discussion of the mathematical background in Section 2, we present the following main contributions:

– We introduce a data-weak reformulation of the classical 3D-VAR method in Section 3. We present a detailed analysis of the method’s stability with respect to the properties of the original best-knowledge model, the model bias, and the choice of the measurement space. We show that our method is stable in the sense of Hadamard independently of the choice of measurement functionals and uniformly over the parameter domain. Furthermore, we identify a necessary and sufficient property of the measurement space that further increases the stability and restricts the amplification of measurement noise when the 3D-VAR method emphasizes closeness to the data over the best-knowledge model.

– In Section 4, we present the RB method for the 3D-VAR problem and develop online computationally efficient approximations and a posteriori error bounds for the state estimate, the adjoint solution, the model correction, and the misfit. This significant reduction in computational complexity makes the 3D-VAR method feasible for repeated computations with measurement noise and varying parameters. Furthermore, the RB approximation aids in the efficient selection of the measurements.

– In Section 5, we integrate the theoretical results from both the stability and error analyses to propose an iterative selection of the measurement functionals. More specifically, we propose an algorithm that employs the OMP measurement selection from [6] in a greedy manner over the parameter domain to generate the measurement space. The space can be tailored specifically to the stable estimation of parameters and model properties of the 3D-VAR method from measurement data. We also discuss a construction of the reduced basis spaces which does not require the measurements to be known a priori.

In Section 6 we present numerical results for a steady-state thermal conduction problem with uncertain parameters and unknown Neumann boundary condition and investigate the influence of measurement noise upon the estimation of the uncertain parameters and boundary condition.
2 Preliminaries

In this section, we specify the mathematical framework within which we cast the 3D-VAR method in the following section. We introduce the required spaces and forms and list all our assumptions upon them.

We start with the state space $\mathcal{Y}$, which we take to be a Hilbert space with inner product $(\cdot, \cdot)_{\mathcal{Y}}$ and induced norm $\| \cdot \|_{\mathcal{Y}}$. Likewise, we consider a Hilbert space $\mathcal{U}$ for the regularizing model correction, with norm $\| \cdot \|_{\mathcal{U}}$ induced by the inner product $(\cdot, \cdot)_{\mathcal{U}}$. In the numerical implementation, these spaces are typically finite-dimensional subspaces—in the RB literature usually referred to as “truth” spaces—of Sobolev or $L^2$ spaces, e.g., finite element spaces. For notational convenience, we omit the explicit distinction between the infinite-dimensional and finite-dimensional settings in the following since the results presented hold for both cases (with the appropriate definitions). The RB approximation in Section 4 utilizes closed subspaces $\mathcal{Y}_R \subset \mathcal{Y}$ and $\mathcal{U}_R \subset \mathcal{U}$ that represent the most dominant model dynamics.

For the incorporation of measurement data, we additionally consider a nontrivial, finite-dimensional subspace $\mathcal{T} \subset \mathcal{Y}$ as measurement space and let the operator $\Pi_1^{\mathcal{T}} : \mathcal{Y} \rightarrow \mathcal{T}$ denote the orthogonal projection onto this space. We assume $\dim \mathcal{T} < \infty$, but note that most of our analysis holds for a generic closed subspace $\mathcal{T} \subset \mathcal{Y}$.

Remark 1 We briefly explain how the space $\mathcal{T}$ can be linked to physical measurements: Suppose we are given $L < \infty$ linearly independent measurement functionals $g_l \in \mathcal{Y}'$, $l \in \{1, \ldots, L\}$. We can then choose the hierarchical space $\mathcal{T}$ as the span of the Riesz representations of the measurement functions (see [2, 27]) via

$$\mathcal{T} = \text{span}\{ \tau_l \in \mathcal{Y} : 1 \leq l \leq L \text{ and } (\tau_l, y)_{\mathcal{Y}} = g_l(y) \forall y \in \mathcal{Y} \}. \quad (1)$$

For any state $y \in \mathcal{Y}$, the projection $\Pi_1^{\mathcal{T}} y$ is then the only state in $\mathcal{T}$ that yields the same measurements as $y$, i.e., $g_l(\Pi_1^{\mathcal{T}} y) = g_l(y) \in \mathbb{R}$ for $l = 1, \ldots, L$. In this context, we can consider $\Pi_1^{\mathcal{T}} y$ to be the experimentally observed part of $y$. Due to the linear independence of the measurement functionals, any set $m = (m_l)_{l=1}^L \in \mathbb{R}^L$ is obtained as the measurement data of exactly one state $y_m(\mathbf{m})$ in $\mathcal{T}$ and—due to the projection theorem—of all states in $\mathcal{Y}_R^{\perp} + \mathcal{T}^{\perp}$. More specifically, two states in $\mathcal{Y}$ yield the same measurements if and only if their difference lies in $\mathcal{T}^{\perp}$. In practice, the number $L$ of measurements should be kept small to limit experimental expenses.

We let $\mathcal{C} \subset \mathbb{R}^d$ be a compact set of all admissible parameters. For this set, we introduce three (possibly) parameter-dependent forms, namely, $f_{bk, \mu} \in \mathcal{Y}'$ and the bilinear forms $a_{\mu} : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}$ and $b_{\mu} : \mathcal{U} \times \mathcal{Y} \rightarrow \mathbb{R}$. Throughout this paper, the index $bk$ stands for “best-knowledge,” whereas the index $\mu$ signifies the dependence on a parameter $\mu \in \mathcal{C}$. The first two forms represent the best-knowledge model dynamics.

$$\text{Find } y_{bk, \mu} \in \mathcal{Y} \text{ such that } a_{\mu}(y_{bk, \mu}, \psi) = f_{bk, \mu}(\psi) \quad \forall \psi \in \mathcal{Y}, \quad (2)$$

whereas for $u \in \mathcal{U}$, $b_{\mu}(u, \cdot) \in \mathcal{Y}'$ denotes the induced model modification.
We make the assumptions that \( a_\mu \) is uniformly coercive over \( C \),

\[
\exists \alpha_a > 0 \text{ s.t. } \alpha_a(\mu) := \inf_{y \in Y} \frac{a_\mu(y, y)}{|y||y|_Y^2} \geq \alpha_a \forall \mu \in C,
\]

and that both bilinear forms \( a_\mu \) and \( b_\mu \) are uniformly bounded, i.e.,

\[
\exists \gamma_a > 0 \text{ s.t. } \gamma_a(\mu) := \sup_{y \in Y} \sup_{z \in Y} \frac{a_\mu(y, z)}{|y||z|_Y} \leq \gamma_a < \infty \forall \mu \in C,
\]

\[
\exists \gamma_b > 0 \text{ s.t. } \gamma_b(\mu) := \sup_{u \in U} \sup_{y \in Y} \frac{b_\mu(u, y)}{|u||y|_Y} \leq \gamma_b < \infty \forall \mu \in C.
\]

The conditions (3) and (4) guarantee that the best-knowledge model (2) and the 3D-VAR formulation introduced in the next section are uniformly stable and that the error of the RB approximation is quasi-optimal with uniformly bounded constants over \( C \). In anticipation of the a posteriori error estimation procedure, we additionally presuppose that we can compute a lower bound \( \alpha_{a}^{\text{LB}}(\mu) \leq \alpha_a(\mu) \) and an upper bound \( \gamma_{b}^{\text{UB}}(\mu) \geq \gamma_b(\mu) \) at reasonably low cost for all \( \mu \in C \), e.g., through the min-\( \theta \)-approach or the successive constraint method \([17, 34]\).

For the efficient computation of the RB solution and a posteriori error bounds in an offline-online procedure, we make the assumption that the \( \mu \)-dependent forms \( a_\mu \), \( b_\mu \), and \( f_{bk,\mu} \) are affine in functions of the parameter, i.e.,

\[
a_\mu = \sum_{\vartheta=1}^{\Theta_a} \theta_a^\vartheta(\mu) a^\vartheta \quad b_\mu = \sum_{\vartheta=1}^{\Theta_b} \theta_b^\vartheta(\mu) b^\vartheta \quad f_{bk,\mu} = \sum_{\vartheta=1}^{\Theta_f} \theta_f^\vartheta(\mu) f_{bk}^\vartheta,
\]

where the coefficient functions \( \theta_a^\vartheta, \theta_b^\vartheta, \theta_f^\vartheta : C \to \mathbb{R} \) are continuous in the parameter \( \mu \), and the bilinear forms \( a^\vartheta \) and \( b^\vartheta \) as well as the linear forms \( f_{bk}^\vartheta \) are parameter-independent. In the nonaffine case, (generalized) empirical interpolation techniques may be used to construct an affine approximation \([1, 14, 25, 26]\).

### 3 3D-VAR formulation

In the following, we introduce the 3D-VAR formulation for a coercive, parameter-dependent PDE. The method aims at finding a weighted compromise between a best-knowledge model and measurement data by making a data-informed perturbation of the model. After a reformulation as a saddle-point problem, a stability analysis shows how the choice of measurement functionals influences the amplification of errors in the measurements and the approximation of the best-knowledge source term. This leads to a practical criterion for the selection of suitable measurement functionals.

---

1 Throughout this paper, we adopt the notational convention that infima and suprema exclude elements with norm 0.
3.1 Problem statement

The main goal in data assimilation is to estimate the (unknown) state \( y_{\text{true}} \in \mathcal{Y} \) of a physical system. To this end, we suppose that a best-knowledge mathematical model exists in terms of the underlying elliptic PDE (2) whose unique solution \( y_{bk,\mu} \) provides a first approximation of \( y_{\text{true}} \). However, any model can only provide an approximation to the underlying physics due to simplifications in the derivation of the governing PDE, and boundary conditions, or due to approximations and uncertainties of the system geometry or loading, \( f_{bk,\mu} \). Since \( y_{bk,\mu} \neq y_{\text{true}} \) in general, we thus introduce a perturbation of the best-knowledge problem (2) in order to allow for a better approximation of the state \( y_{\text{true}} \). Here, we specifically consider model corrections of the form \( b_{\mu}(u, \cdot) \in \mathcal{Y}' \) with \( u \in \mathcal{U} \), leading to the modified problem:

\[
\begin{align*}
\text{For a given model modification } u_0 \in \mathcal{U}, \text{ find } y_{\mu} &= y_{\mu}(u_0) \in \mathcal{Y} \text{ s.t.} \\
a_{\mu}(y_{\mu}, \psi) &= f_{bk,\mu}(\psi) + b_{\mu}(u_0, \psi) \quad \forall \psi \in \mathcal{Y}.
\end{align*}
\]

Within this framework, we now aim to construct an improved approximation of \( y_{\text{true}} \) by finding a “good” model correction \( u \in \mathcal{U} \).

To do this, we incorporate knowledge of the true state \( y_{\text{true}} \) in form of an approximation \( y_d \in \mathcal{T} \) to \( y_{\text{true}} \). In the notation of Remark 1, this approximation might be \( y_d = y_d(m_d) \in \mathcal{T} \), resulting from measurement data \( m_d \in \mathbb{R}^L \), \( (m_d)_l = g_l(y_{\text{true}}) + \epsilon_l \) of the true state with noise \( \epsilon_l \in \mathbb{R} \). To approximate \( y_{\text{true}} \in \Pi_{\mathcal{T}}y_{\text{true}} + \mathcal{T}^\perp \approx y_d + \mathcal{T}^\perp \), the model correction \( u \in \mathcal{U} \) should thus be chosen such that the solution \( y_{\mu}(u) \) of the modified model (6) for \( u_0 = u \) is close to \( y_d + \mathcal{T}^\perp \). By the projection theorem, this corresponds to choosing \( u \) such that the experimentally observable misfit \( d_{\mu}(u) = y_d - \Pi_{\mathcal{T}}y_{\mu}(u) \) between \( y_{\mu}(u) \) and the data state \( y_d \), measured in the \( \mathcal{Y} \) norm, is small. At the same time, however, we want the model correction, measured in \( \|u\|_U \), to be small such that the state \( y_{\mu}(u) \) stays close to our best-knowledge model (2). Note that \( y_{\mu}(0) = y_{bk,\mu} \) for \( u_0 = 0 \in \mathcal{U} \) in (6). The 3D-VAR formulation thus takes the form

\[
\begin{align*}
\min_{(u, y, d) \in \mathcal{U} \times \mathcal{Y} \times \mathcal{T}} & \quad \frac{1}{2} \|u\|_U^2 + \frac{\lambda}{2} \|d\|_{\mathcal{Y}}^2 \\
\text{s.t.} & \quad a_{\mu}(y, \psi) = f_{bk,\mu}(\psi) + b_{\mu}(u, \psi) \quad \forall \psi \in \mathcal{Y}, \\
& \quad (y + d, \tau) = (y_d, \tau) \quad \forall \tau \in \mathcal{T},
\end{align*}
\]

where the regularization parameter \( \lambda > 0 \) quantitatively expresses the trust in the original model (2) over the validity of the measurements: a small factor \( \lambda \) prioritizes proximity to the original model, whereas a large \( \lambda \) favors closeness to the data state \( y_d \). By specifically considering model corrections and the distance in the measurement space \( \mathcal{T} \) (instead of the difference in measurements), we follow the model-data weak approach to variational data assimilation introduced in [41].

Denoting the optimal solution by \( (u_{\mu}^*, y_{\mu}^*, d_{\mu}^*) \in \mathcal{U} \times \mathcal{Y} \times \mathcal{T} \), then \( y_{\mu}^* = y_{\mu}(u_{\mu}^*) \) is the solution of the perturbed model (6) with model correction \( u_0 = u_{\mu}^* \), and \( d_{\mu}^* = y_d - \Pi_{\mathcal{T}}y_{\mu}^* \) is the experimentally observable misfit between the data state \( y_d \) and the 3D-VAR solution state \( y_{\mu}^* \). The minimization over \( \|u\|_U \) in the cost function enforces
that the model correction $u^*_\mu$ contains only components that actively decrease the misfit, i.e., $u^*_\mu$ is perpendicular to the closed subspace

$$U_0(\mu) := \{ u \in \mathcal{U} : \Pi_\mathcal{T} y_{\mu} = 0 \text{ where } a_{\mu}(y_{\mu}, \psi) = b_{\mu}(u, \psi) \forall \psi \in \mathcal{Y} \}. \quad (8)$$

If the measurements can be reproduced in the modified model (6), that is, if $y_d = \Pi_\mathcal{T} y_{\mu}(u_{\text{true}})$ for a solution $y_{\mu}(u_{\text{true}})$ of (6) with $u_0 = u_{\text{true}} \in \mathcal{U}$, then $\frac{1}{2}||u_{\text{true}}||^2_U$ is a natural, $\lambda$-independent upper bound for the cost in (7). Therefore, $||d^*_\mu||_Y \in O(\lambda^{-1/2})$, and $||u^*_\mu||_U \leq ||u_{\text{true}}||_U$ is bounded independently of $\lambda$.

**Remark 2** Formally, the 3D-VAR problem (7) is a Tikhonov regularization

$$\min_{u \in \mathcal{U}} \frac{1}{\lambda} ||u||^2_U + ||Q_{\mu} u - y_d + \Pi_\mathcal{T} y_{bk,\mu}||_Y^2 \quad (9)$$

with regularization parameter $1/\lambda$ and bounded linear operator $Q_{\mu} : \mathcal{U} \to \mathcal{T}$, $Q_{\mu} u_0 := \Pi_\mathcal{T} y_{\mu}(u_0)$, where $y_{\mu}(u_0) \in \mathcal{Y}$ is the unique solution of (6) with source term $f_{bk,\mu} = 0$. With standard theory for Tikhonov regularizations (see, e.g., [37]), there exists a unique solution $u^*_\mu = u^*_\mu(\lambda, y_d) \in \mathcal{U}$ to (9) and hence to the 3D-VAR problem (7) for all data states $y_d \in \mathcal{T}$ and all $\lambda > 0$. In particular, the solution depends continuously on the data with $||u^*_\mu(\lambda, y_d)||_U \leq \sqrt{\lambda} ||y_d - \Pi_\mathcal{T} y_{bk,\mu}||_Y$. Thus, the 3D-VAR problem (7) is well-posed.

**Remark 3** Since range$(Q_{\mu}) \subset \mathcal{T}$ and dim $\mathcal{T} < \infty$, we can exploit the connection between Tikhonov regularizations and generalized inverses [15] to obtain that for $\lambda \to \infty$, the solution $u^*_\mu(\lambda, y_d)$ converges to the generalized inverse $Q_{\mu}^+(y_d - \Pi_\mathcal{T} y_{bk,\mu})$, which is the unique solution to

$$\min_{u \in \mathcal{U}} ||u||_U \quad \text{s.t.} \quad u \in \arg \min_{v \in \mathcal{U}} ||Q_{\mu} v - y_d + \Pi_\mathcal{T} y_{bk,\mu}||_Y.$$ 

If the true state $y_{\text{true}}$ can be completely described by (6) for a model modification $u_0 = u_{\text{true}}$, and if in addition $y_d = \Pi_\mathcal{T} y_{\text{true}}$ is an unbiased observation (e.g., from noise-free measurements), then $||Q_{\mu} u - y_d + \Pi_\mathcal{T} y_{bk,\mu}||_Y = 0$ if and only if $u \in u_{\text{true}} + U_0(\mu)$. The 3D-VAR model correction $u^*_\mu$ then converges to $\Pi_{U_0^\perp(\mu)}^U u_{\text{true}}$.

### 3.2 Saddle-point formulation

We next recast the 3D-VAR minimization (7) as a saddle-point problem. Although we already know that the minimization is well-posed as a Tikhonov regularization for any fixed parameter $\mu$, the saddle-point structure has several benefits: First, the analytic properties known for saddle-point problems—such as Brezzi’s theorem—allow us to quantify the structural influences of the original best-knowledge model (2), the model perturbation $b_{\mu}$, and the measurement space $\mathcal{T}$ on the amplification of errors in $y_d$ and $f_{bk,\mu}$. Second, the analysis shows that the 3D-VAR problem is uniformly well-posed over the parameter domain, enabling the use of the method in parameter estimation. Third, in the discretized setting, we obtain a linear system with a saddle-point structure for which various numerical methods exist.
We note that for the solution \((u^*_\mu, y^*_\mu, d^*_\mu) \in \mathcal{U} \times \mathcal{Y} \times \mathcal{T}\) of (7), the experimentally observable misfit, \(d^*_\mu = y_d - \Pi_T y^*_\mu\), between the 3D-V AR solution \(y^*_\mu\) and the observed state \(y_d\) can be computed a posteriori from \(u^*_\mu\) and \(y^*_\mu\). We therefore omit the explicit dependence on \(d^*_\mu\) in the saddle point formulation which not only results in a smaller system but also improved stability constants. For an approach which explicitly includes \(d^*_\mu\), we refer to [31].

We next define the Hilbert space \(\mathcal{H} := \mathcal{U} \times \mathcal{Y}\) with induced inner product \(((u, y), (\phi, \psi))_\mathcal{H} := (u, \phi)_\mathcal{U} + (y, \psi)_\mathcal{Y}\) and induced norm \(||\cdot||_\mathcal{H} := \sqrt{\cdot, \cdot}_\mathcal{H}\). Furthermore, we define the bilinear and linear forms

\[
\begin{align*}
A : \mathcal{H} \times \mathcal{H} &\to \mathbb{R} : A((u, y), (\phi, \psi)) := (u, \phi)_\mathcal{U} + \lambda (\Pi_T y, \Pi_T \psi)_\mathcal{Y} \\
B_\mu : \mathcal{H} \times \mathcal{Y} &\to \mathbb{R} : B_\mu((u, y), \psi) := a_\mu(y, \psi) - b_\mu(u, \psi) \\
F : \mathcal{H} &\to \mathbb{R} : F(\psi) := \lambda (y_d, \psi)_\mathcal{Y} \\
G_\mu : \mathcal{Y} &\to \mathbb{R} : G_\mu(\psi) := f_{b_{k,\mu}}(\psi).
\end{align*}
\]

Note that \(F\) and \(G_\mu\) are linear with \(||F||_{\mathcal{H}} = \lambda ||y_d||_\mathcal{Y}\) and \(||G_\mu||_{\mathcal{Y}} = ||f_{b_{k,\mu}}||_{\mathcal{Y}}\). Furthermore, \(A\) is symmetric and positive semi-definite, and both \(A\) and \(B_\mu\) are bilinear with bounded continuity constants

\[
\begin{align*}
\gamma_A := \sup_{h \in \mathcal{H}} \sup_{k \in \mathcal{H}} \frac{A(h, k)}{||h||_\mathcal{H}||k||_\mathcal{H}} = \max\{1, \lambda\} \\
\gamma_B(\mu) := \sup_{h \in \mathcal{H}} \sup_{y \in \mathcal{Y}} \frac{B_\mu(h, y)}{||h||_\mathcal{H}||y||_\mathcal{Y}} \leq \sqrt{\gamma_A^2 + \gamma_B(\mu)^2}.
\end{align*}
\]

With the definitions in (10), the 3D-V AR minimization (7) is equivalent to

\[
\min_{h \in \mathcal{H}} \frac{1}{2} A(h, h) - F(h) \text{ such that } B_\mu(h, \psi) = G_\mu(\psi) \text{ for all } \psi \in \mathcal{Y}.
\]

Employing a Lagrangian approach, we obtain the associated necessary, and in our setting sufficient, first-order optimality conditions [16, 33]: Given \(\mu \in \mathcal{C}\), find \((h^*_\mu, p^*_\mu) = ((u^*_\mu, y^*_\mu), p^*_\mu) \in \mathcal{H} \times \mathcal{Y}\) such that

\[
\begin{align*}
A(h^*_\mu, k) + B_\mu(k, p^*_\mu) &= F(k) \quad \forall k \in \mathcal{H}, \\
B_\mu(h^*_\mu, \psi) &= G_\mu(\psi) \quad \forall \psi \in \mathcal{Y}.
\end{align*}
\]

### 3.3 Stability analysis

The a priori stability of the 3D-V AR formulation is closely linked to the choice of the measurement space \(\mathcal{T}\). The following stability analysis thus provides a criterion for choosing the measurement space \(\mathcal{T}\) in Section 5.1. Furthermore, it allows us to link the properties of the modified model (6) and the choice of the measurement space \(\mathcal{T}\) to the amplification of noise in the 3D-V AR solution. We start with the Ladyzhenskaya-Babuška-Brezzi (LBB) condition, for which we prove the lower bound \(\beta_{\text{LB}}(\mu) := \alpha_a(\mu)\). In the sequel, the superscripts LB and UB shall refer to lower and upper bounds, respectively.

\(\copyright\) Springer
Theorem 1

\[ \beta_B(\mu) := \inf_{y \in Y} \sup_{h \in H} \frac{B_\mu(h, y)}{||h||_H ||y||_Y} \geq \beta_B^{LB}(\mu) := \alpha_a(\mu) \geq \alpha_a > 0. \]

Proof For any \( y \in Y \setminus \{0\} \), we obtain

\[ \sup_{h \in H} \frac{B_\mu(h, y)}{||h||_H ||y||_Y} = \sup_{(\phi, \psi) \in H} \frac{a_\mu(\psi, y) - b_\mu(\phi, y)}{||\phi||_H ||y||_Y} \geq \frac{a_\mu(y, y)}{||y||_Y^2} \geq \alpha_a(\mu), \]

where we have chosen \((\phi, \psi) = (0, y)\) and inserted the coercivity condition (3).

We now turn to verifying the coercivity of \( A \) on the nullspace of \( B_\mu \), given by

\[ H^0(\mu) := \{ h \in H : B_\mu(h, \psi) = 0 \quad \forall \psi \in Y \} \]

\[ = \{ (u, y) \in U \times Y : a_\mu(y, \psi) = b_\mu(u, \psi) \quad \forall \psi \in Y \} \subseteq H. \quad (13) \]

Note that \( H^0(\mu) \) is a closed subspace of \( H \). For a concise notation, we introduce \( Y_\mu := \{ y \in Y : \exists u \in U \text{ s.t. } (u, y) \in H^0(\mu) \} \)

\[ = \{ y \in Y : \exists u \in U \text{ s.t. } a_\mu(y, \psi) = b_\mu(u, \psi) \quad \forall \psi \in Y \} \quad (14) \]

as the space of all states in \( H^0(\mu) \). Note that \( Y_\mu \neq \{0\} \) since \( \gamma_\mu > 0 \) by (4).

With \( A(0, \psi), (0, \psi) = 0 \) for all \( \psi \in T^\perp \), \( A \) is in general not \( H \)-coercive. For the application of Brezzi’s theorem \[ 8\], it is, however, already sufficient for \( A \) to be \( H^0(\mu) \)-coercive. This property follows from the coercivity of \( a_\mu \) as shown in the following theorem.

Theorem 2 For \( \mu \in C \), let \( \alpha^0_A(\mu, \lambda) \) be the coercivity constant of \( A \) on \( H^0(\mu) \), i.e.,

\[ \alpha^0_A(\mu, \lambda) := \inf_{h \in H^0(\mu)} \frac{A(h, h)}{||h||^2_H}. \quad (15) \]

Define the ratios

\[ \eta(\mu) := \inf_{(u, y) \in H^0(\mu)} \frac{||y||_Y}{||u||_U} \geq 0 \quad \text{and} \quad \overline{\eta}(\mu) := \sup_{(u, y) \in H^0(\mu)} \frac{||y||_Y}{||u||_U} > 0, \quad (16) \]

and the inf-sup constant

\[ \beta_T(\mu) := \inf_{y \in Y_\mu} \sup_{\tau \in T} \frac{(y, \tau)_Y}{||y||_Y ||\tau||_Y} = \inf_{y \in Y_\mu} \frac{||\Pi_T y||_Y}{||y||_Y} \geq 0. \quad (17) \]

Then

\[ \alpha_A^{LB}(\lambda, \eta(\mu), \overline{\eta}(\mu), \beta_T(\mu)) := \begin{cases} \frac{1 + \lambda \beta_T(\mu)^2 \overline{\eta}(\mu)^2}{1 + \eta(\mu)^2}, & \text{if } \lambda \beta_T(\mu)^2 \leq 1 \\ \frac{1 + \lambda \beta_T(\mu)^2 \eta(\mu)^2}{1 + \eta(\mu)^2}, & \text{otherwise} \end{cases} \quad (18) \]

is a positive lower bound to the \( H^0(\mu) \)-coercivity constant of \( A \), i.e.,

\[ \alpha^0_A(\mu, \lambda) \geq \alpha_A^{LB}(\lambda, \eta(\mu), \overline{\eta}(\mu), \beta_T(\mu)) \geq \frac{1}{1 + \eta(\mu)^2} > 0. \quad (19) \]
Proof Since $a_\mu$ is coercive and $\gamma_\mu(\mu) > 0$ by assumption (4), there exists at least one element $(u, y) \in H^0(\mu)$ with $y \neq 0$; hence, $\bar{\eta}(\mu) > 0$. Let $0 \neq h = (u, y) \in H^0(\mu)$ be arbitrary. We note that $y \in Y_\mu$. With the definitions (16) and (17), we obtain $\bar{\eta}(\mu)||u||_U \leq ||y||_Y \leq \bar{\eta}(\mu)||u||_U$ and $||\eta T y||_Y \geq \beta T(\mu)||y||_Y$, respectively.

If $\lambda \beta T(\mu)^2 \leq 1$, then

$$A(h, h) = ||u||^2_U + \lambda||\eta T y||^2_Y$$

$$\geq ||u||^2_U + \lambda \beta T(\mu)^2 ||y||^2_Y$$

$$\geq (1 - x \bar{\eta}(\mu)^2) ||u||^2_U + (\lambda \beta T(\mu)^2 + x) ||y||^2_Y$$

$$\geq \min\{1 - x \bar{\eta}(\mu)^2, \lambda \beta T(\mu)^2 + x\} ||h||^2_\mathcal{H}$$

for arbitrary $x \in [0, \bar{\eta}(\mu)^{-2}]$. The minimum is maximal when both terms are equal, which is the case for $x = \frac{1 - \lambda \beta T(\mu)^2}{1 + \bar{\eta}(\mu)^2} < \bar{\eta}(\mu)^{-2}$. The minimum then takes the value $\alpha^\text{LB}_A(\lambda, \eta(\mu), \bar{\eta}(\mu), \beta T(\mu))$.

Now suppose $\lambda \beta T(\mu)^2 > 1$. If $y = 0$, then $\bar{\eta}(\mu) = 0$ and $A(h, h) = ||u||^2_U = ||h||^2_\mathcal{H}$. For $y \neq 0$, we obtain for any $x \in [0, \lambda \beta T(\mu)^2]$: $A(h, h) = ||u||^2_U + \lambda||\eta T y||^2_Y$$

$$\geq ||u||^2_U + \lambda \beta T(\mu)^2 ||y||^2_Y$$

$$\geq (1 + x \bar{\eta}(\mu)^2) ||u||^2_U + (\lambda \beta T(\mu)^2 - x) ||y||^2_Y$$

$$\geq \min\{1 + x \bar{\eta}(\mu)^2, \lambda \beta T(\mu)^2 - x\} ||h||^2_\mathcal{H}$$

With $x = \frac{\lambda \beta T(\mu)^2 - 1}{1 + \bar{\eta}(\mu)^2}$, we obtain the second part of $\alpha^\text{LB}_A$.

We briefly comment on the meaning of some of the constants. While the inf-sup constant $\beta T(\mu)$ relates to the experimental design, the ratios $\eta(\mu)$ and $\bar{\eta}(\mu)$ relate to the modeling process. The former depends on the choice of measurements, i.e., the measurement space $\mathcal{T}$, and reflects how well model modifications can be distinguished based on changes in $\mathcal{T}$. The latter ratios $\eta(\mu)$ and $\bar{\eta}(\mu)$ reflect how much change in the state can minimally and maximally be evoked by a model modification in $\mathcal{U}$, respectively. As already mentioned in the proof of Theorem 2, $\bar{\eta}(\mu) > 0$ due to $\gamma_\mu(\mu) > 0$. Note that if $\bar{\eta}(\mu) = 0$ would hold, $\mathcal{U}$ would have no effect on the best-knowledge model (2). The ratio $\eta(\mu)$ equals zero if there exists a superfluous search direction $u \in \mathcal{U}$ whose physical influence $b_\mu(u, \cdot)$ is not sufficiently captured by $\mathcal{Y}$. To express this correlation, we define, for nontrivial spaces $\mathcal{V} \subset \mathcal{U}, \mathcal{W} \subset \mathcal{Y}$, the inf-sup constant

$$\beta_b(\mathcal{V}, \mathcal{W}) := \inf_{u \in \mathcal{V}} \sup_{w \in \mathcal{W}} \frac{b_\mu(u, w)}{||u||_\mathcal{U}||w||_\mathcal{Y}} \geq 0.$$  

We can then bound $\eta(\mu)$ by $\frac{\beta_b(\mathcal{U}, \mathcal{Y}_\mu)}{\gamma_\mu(\mu)} \leq \eta(\mu) \leq \frac{\beta_b(\mathcal{U}, \mathcal{Y}_\mu)}{\alpha_\mu(\mu)}$. The inf-sup constant $\beta_b(\mathcal{U}, \mathcal{Y}_\mu)$ thus provides information on the quality of the model modifications for any parameter $\mu \in \mathcal{C}$ and may prove useful in the design of possible model modifications.  

© Springer
In order to quantify the influence of the noise on the 3D-VAR solution, we first present a result concerning the behavior of $\alpha_A^0(\mu, \lambda)$. To describe asymptotic behavior in $\lambda$, we make use of the Landau symbol $\Theta$ (see, e.g., [9]): A function $f$ lies in the class $\Theta(g)$ if $g$ is an asymptotically tight bound, i.e., there exist $c_1, c_2 > 0$ and a $\lambda_0 \geq 0$ such that $0 \leq c_1 g(\lambda) \leq f(\lambda) \leq c_2 g(\lambda)$ for all $\lambda \geq \lambda_0$.

**Theorem 3** For $\mu \in \mathcal{C}$, either $\alpha_A^0(\mu, \lambda) \leq 1$ for all $\lambda > 0$ or $\alpha_A^0(\mu, \lambda) \in \Theta(\lambda)$ with $\lambda_0 > 0$. The latter is equivalent to both $\eta(\mu) > 0$ and $\beta_T(\mu) > 0$ being satisfied.

**Proof** By Theorem 2 and (11),

$$0 < \frac{1}{1 + \eta(\mu)^2} \leq \alpha_A^{LB}(\lambda, \eta(\mu), \overline{\eta}(\mu), \beta_T(\mu)) \leq \alpha_A^0(\mu, \lambda) \leq \gamma_A = \max\{1, \lambda\},$$

(21)

where $\alpha_A^{LB}(\lambda, \eta(\mu), \overline{\eta}(\mu), \beta_T(\mu)) \in \Theta(\lambda)$ if and only if $\eta(\mu) > 0$ and $\beta_T(\mu) > 0$. It is hence sufficient to show that $\eta_A(\mu) \leq 1$ for all $\lambda > 0$ when $\eta(\mu) = 0$ or $\beta_T(\mu) = 0$. We note that $\lambda_0 > 0$ stems from the upper bound $\alpha_A^0(\mu, \lambda) \leq \max\{1, \lambda\}$ in (21).

We suppose first that $\eta(\mu) = 0$. Let $\lambda > 0$ be fixed. With (16), there exists a sequence $h_n = (u_n, y_n) \in \mathcal{H}(\mu), n \in \mathbb{N}$, such that $||u_n||_{\mathcal{U}} = 1$ and $||y_n||_{\mathcal{Y}} \to 0$ for $n \to \infty$. Then, $\alpha_A^0(\mu, \lambda) \leq \frac{A(h_n, h_n)}{||h_n||_{\mathcal{U}}^2} \leq \frac{1 + ||y_n||_{\mathcal{Y}}^2}{1 + ||y_n||_{\mathcal{Y}}^2} \to 1$ for $n \to \infty$.

If $\beta_T(\mu) \neq 0$, there exists $h_n = (u_n, y_n) \in \mathcal{H}(\mu), n \in \mathbb{N}$, with $||y_n||_{\mathcal{Y}} = 1$, and for $n \to \infty$, $||\Pi_T y_n||_{\mathcal{Y}} \to 0$. In particular, $||u_n||_{\mathcal{U}} > 0$ for all $n \in \mathbb{N}$.

$$\alpha_A^0(\mu, \lambda) \leq \frac{A(h_n, h_n)}{||h_n||_{\mathcal{U}}^2} = \frac{||u_n||_{\mathcal{U}}^2 + ||y_n||_{\mathcal{Y}}^2}{||u_n||_{\mathcal{U}}^2 + ||y_n||_{\mathcal{Y}}^2} + \lambda \frac{||\Pi_T y_n||_{\mathcal{Y}}^2}{||u_n||_{\mathcal{U}}^2 + ||y_n||_{\mathcal{Y}}^2},$$

where the first term is bounded by $(1 + \eta(\mu)^2)^{-1} \leq 1$ and the second converges to 0 for $n \to \infty$ and any fixed $\lambda$.

$\square$

Theorem 3 shows that the lower bound from $\alpha_A^{LB}(\lambda, \eta(\mu), \overline{\eta}(\mu), \beta_T(\mu))$ underpredicts the true $\mathcal{H}(\mu)$-coercivity constant $\alpha_A^0(\mu, \lambda)$ only by a bounded multiplicative factor, i.e., $\alpha_A^0(\mu, \lambda) \in \Theta(\alpha_A^{LB}(\lambda, \eta(\mu), \overline{\eta}(\mu), \beta_T(\mu)))$. We thus gain insight on how the main behavior of $\alpha_A^0(\mu, \lambda)$ is influenced by the quantities $\eta(\mu), \overline{\eta}(\mu),$ and $\beta_T(\mu)$. By optimizing these quantities in the modeling process and with the selection of measurements, we have a direct way of influencing $\alpha_A^0(\mu, \lambda)$ and thereby the stability of the whole 3D-VAR system (12).

We next recall that $\beta_B(\mu)$ and $\alpha_A^0(\mu, \lambda)$ are uniformly bounded away from zero as shown in Theorems 1 and 2, respectively. It then follows from Brezzi’s theorem that the 3D-VAR formulation is uniformly well-posed over the parameter domain for all source functions in $\mathcal{H}'$ and $\mathcal{Y}'$ [7, 8]. For the particular choice in (10), we get the following stability bounds from Theorem 5.2, p. 38, in [7] after inserting $||F||_{\mathcal{H}'} = \lambda ||y_d||_{\mathcal{Y}}$ and $||G_{\mu}||_{\mathcal{Y}'} = ||f_{dbk, \mu}||_{\mathcal{Y}'}$. 

$\odot$ Springer
Theorem 4  Problem (12) has a unique solution \((h^*_\mu, p^*_\mu) \in \mathcal{H} \times \mathcal{Y},\) and
\[
\|h^*_\mu\|_\mathcal{H} \leq \frac{\lambda}{\alpha^*_A(\mu, \lambda)} \|y\|_\mathcal{Y} + \|f_{bk,\mu}\|_\mathcal{Y} \left(\frac{\gamma_A}{\alpha^*_A(\mu, \lambda)} + 1\right) \quad \text{(22a)}
\]
\[
\|p^*_\mu\|_\mathcal{Y} \leq \frac{\lambda}{\beta_B(\mu)} \|y\|_\mathcal{Y} \left(\frac{\gamma_A}{\alpha^*_A(\mu, \lambda)} + 1\right) + \frac{\gamma_A ||f_{bk,\mu}||_\mathcal{Y}}{\beta_B(\mu)^2} \left(\frac{\gamma_A}{\alpha^*_A(\mu, \lambda)} + 1\right) \quad \text{(22b)}
\]

We note that the effect of experimental noise on the solution is governed by the stability coefficients in front of \(\|y\|_\mathcal{Y}.\) Since \(\gamma_A = \max\{1, \lambda\},\) they scale like \(\Theta(\lambda/\alpha^*_A(\mu, \lambda))\) in (22a) and \(\Theta(\lambda^2/\alpha^*_A(\mu, \lambda))\) in (22b) for \(\lambda \geq 1.\) Furthermore, we know from Theorem 3 that \(\alpha^*_A(\mu, \lambda)\) is of order \(\Theta(\lambda)\) for \(\lambda \to \infty\) if and only if \(\eta(\mu) > 0\) and \(\beta_T(\mu) > 0.\) It thus follows that the stability coefficients for the model modification and state variable \(h^*_\mu\) are bounded from above independently of \(\lambda\) in this case and that the stability coefficients for the adjoint \(p^*_\mu\) scale like \(\Theta(\lambda).\) If \(\eta(\mu) = 0\) or \(\beta_T(\mu) = 0,\) the scaling is significantly worse with \(\Theta(\lambda)\) and \(\Theta(\lambda^2),\) respectively. Provided \(\eta(\mu) > 0,\) we should thus choose \(T\) such that the inf-sup constant \(\beta_T(\mu)\) is maximized. We discuss such a stability-based generation of the measurement space in Section 5.1.

Remark 4  We briefly comment on the link between the inf-sup constant \(\beta_T(\mu),\) the PBDW formulation, and the 3D-VAR solution. Suppose \(\mathcal{Y}_\mu\) is closed, as is the case if \(\dim \mathcal{U} < \infty,\) and suppose further that \(\beta_T(\mu) > 0.\) Let \(y_{bk,\mu} \in \mathcal{Y}\) be the solution of the best-knowledge problem (2). Then, the minimization
\[
\min_{y \in \mathcal{Y}_\mu, d \in \mathcal{T}} ||d||_{\mathcal{Y}} \text{ such that } (y + d, \tau)_{\mathcal{Y}} = (y_d - y_{bk,\mu}, \tau)_{\mathcal{Y}} \forall \tau \in \mathcal{T}
\]
has a unique, \(\mu\)-dependent solution \((\hat{y}_{\infty,\mu}, d_{\infty,\mu}) \in \mathcal{Y}_\mu \times \mathcal{T}.\) This problem is the PBDW formulation [27] over the best-knowledge space \(\mathcal{Y}_\mu,\) where the basis functions of \(\mathcal{U}\) act as parameters. For an extensive analysis of state estimation with the PBDW approach, we refer to the literature [5, 27–29].

With this notation and the discussion in Remark 3, the model correction \(u^*_\mu = u^*_\mu(\lambda)\) hence converges for \(\lambda \to \infty\) to the unique element \(u_{\infty,\mu}\) with minimal norms under all \(u \in \mathcal{U}\) with \((u, \hat{y}_{\infty,\mu}) \in \mathcal{H}^0(\mu).\) If the inf-sup condition \(\beta_T(\mathcal{U}, \mathcal{Y}_\mu) > 0\) holds, or equivalently \(\eta(\mu) > 0,\) then the limit \(u_{\infty,\mu}\) is already uniquely defined by the property \((u_{\infty,\mu}, y_{\infty,\mu}) \in \mathcal{H}^0(\mu).\) In either case, we have \(y^*_\mu = y^*_\mu(\lambda) \to y_{\infty,\mu} := y_{bk,\mu} + \hat{y}_{\infty,\mu}\) and \(d^*_\mu(\lambda) = y_d - \Pi_T y^*_\mu(\lambda) \to d_{\infty,\mu}\) for \(\lambda \to \infty.\)

It has been shown in [5, 27] that if \(y_d = \Pi_T y_{\text{true}},\) then
\[
\|\Pi y_\mu(y_{\text{true}} - y_{\infty,\mu})\|_{\mathcal{Y}} \leq \frac{1}{\beta_T(\mu)} \inf_{d \in \mathcal{T} \cap \mathcal{Y}_\mu^1} \inf_{y \in y_{bk,\mu} + \mathcal{Y}_\mu} ||y_{\text{true}} - (y + d)||_{\mathcal{Y}}
\]
\[
\|y_{\text{true}} - (y_{\infty,\mu} + d_{\infty,\mu})\|_{\mathcal{Y}} \leq \frac{1}{\beta_T(\mu)} \inf_{d \in \mathcal{T} \cap \mathcal{Y}_\mu^1} \inf_{y \in y_{bk,\mu} + \mathcal{Y}_\mu} ||y_{\text{true}} - (y + d)||_{\mathcal{Y}},
\]
and \(\beta_T(\mu)^{-1}\) is the optimal class performance for the second bound (see [5]). The approximation of \(y_{\text{true}}\) through the 3D-VAR solution therefore also profits in the limit \(\lambda \to \infty\) from a preferably large inf-sup constant \(\beta_T(\mu) > 0.\)
Remark 5 We note that the PBDW approach requires $\dim \mathcal{T} \geq \dim \mathcal{U}$ measurements of the true state for the problem to be well-posed. For the 3D-VAR method, on the other hand, the method remains well-posed for any measurement space—although a “good” choice for $\mathcal{T}$ may increase its stability. Additionally, the regularization parameter $\lambda$ provides a direct control over the amplification of noise. By changing $\lambda$, the 3D-VAR method can be adjusted for different noise levels. A comparatively small $\lambda$ value may result in better state estimate than a large $\lambda$ value given noisy data.

3.4 Numerical solution

For the numerical solution of the 3D-VAR saddle-point system (12), the spaces $\mathcal{U}$ and $\mathcal{Y}$ are replaced with finite-dimensional approximation spaces, e.g., conforming finite element spaces. Let $(\phi_m)_{m=1}^{\mathcal{M}} \subset \mathcal{U}$ and $(\psi_n)_{n=1}^{\mathcal{N}} \subset \mathcal{Y}$ be bases of $\mathcal{U}$ and $\mathcal{Y}$, respectively. For $\mathcal{T}$, we consider the basis $(\tau_l)_{l=1}^{L}$ which can be represented in terms of the basis of $\mathcal{Y}$.

The 3D-VAR solution $(h^{*}_\mu, p^{*}_\mu) = (u^{*}_\mu, y^{*}_\mu, p^{*}_\mu) \in \mathcal{U} \times \mathcal{Y} \times \mathcal{Y}$ is then uniquely defined by the basis coefficients of the model modification, state, and adjoint arguments represented by $u^{*}_\mu = \sum_{m=1}^{\mathcal{M}} u^*_m \phi_m$, $y^*_\mu = \sum_{n=1}^{\mathcal{N}} y^*_n \psi_n$, and $p^*_\mu = \sum_{n=1}^{\mathcal{N}} p^*_n \psi_n$, respectively. The coefficient vector $(u^*, y^*, p^*) \in \mathbb{R}^{\mathcal{M}+2\mathcal{N}}$ is the unique solution of the linear $(\mathcal{M}+2\mathcal{N}) \times (\mathcal{M}+2\mathcal{N})$ saddle-point system

$$
\begin{pmatrix}
U & 0 & B^T(\mu) \\
0 & \lambda & A^T(\mu) \\
B(\mu) & A(\mu) & 0
\end{pmatrix}
\begin{pmatrix}
\begin{bmatrix} u^* \\ y^* \\ p^* \end{bmatrix}
\end{pmatrix}
= \begin{pmatrix}
\begin{bmatrix} 0 \\ \lambda y_d \\ f_{bk}(\mu) \end{bmatrix}
\end{pmatrix}. \tag{23}
$$

Here, $U \in \mathbb{R}^{\mathcal{M} \times \mathcal{M}}$ is the mass matrix on $\mathcal{U}$, $P \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}}$ is the matrix representation of the state projection $\Pi_T$, $A(\mu) \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}}$ is the stiffness matrix, and $B(\mu) \in \mathbb{R}^{\mathcal{N} \times \mathcal{M}}$ is the model modification matrix. The entries are given by $U_{i,j} = (\phi_j, \phi_i)_{\mathcal{U}}$, $P_{i,j} = (\Pi_T \psi_j, \Pi_T \psi_i)_{\mathcal{Y}}$, $A_{i,j}(\mu) = a_\mu(\psi_j, \psi_i)$, and $B_{i,j}(\mu) = -b_\mu(\phi_j, \psi_i)$. On the right side of the equation, we have the best-knowledge source term vector $f_{bk}(\mu) \in \mathbb{R}^{\mathcal{N}}$ given by $f_{bk,i}(\mu) = f_{bk,\mu}(\psi_i)$ and the data vector $y_d \in \mathbb{R}^{\mathcal{N}}$ with $y_{d,i} = (y_d, \psi_i)_{\mathcal{Y}}$. The latter can be obtained from the measurement data $m_d$: For $T \in \mathbb{R}^{L \times L}$, $T_{i,j} := g_i(\tau_j)$ and $S \in \mathbb{R}^{\mathcal{N} \times L}$, $S_{i,j} := (\tau_j, \psi_i)_{\mathcal{Y}}$, we have $y_d = (ST^{-1})m_d$.

We note that the affine nature (5) of the bilinear and linear forms $a_\mu$ and $b_\mu$ and $f_{bk,\mu}$ transfers to the corresponding matrices $A(\mu)$ and $B(\mu)$ and $f_{bk}(\mu)$. Hence, for different parameters, they can be assembled from precomputed, $\mu$-independent matrices, e.g., the stiffness matrix $A(\mu)$ has the form $A(\mu) = \sum_{\vartheta=1}^{Q_a} \vartheta^a_\mu A^\vartheta \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}}$ with $(A^\vartheta)_{i,j} = a^\vartheta(\psi_j, \psi_i)$ for $\vartheta = 1, ..., Q_a$.

4 RB approximation

The stability analysis in Section 3.3, in particular Theorem 4, indicates that the 3D-VAR problem (7) is well-posed for every parameter $\mu \in \mathcal{C}$ with uniformly bounded stability constants. Yet, the implementation of the 3D-VAR method involves solving a
large linear system of equations for each parameter \( \mu \), where the \( \mu \)-dependent parts of the block matrix need to be assembled anew whenever \( \mu \) changes. For real-time and many-query applications over the parameter domain, the computational cost hence becomes large.

We address these problems by introducing an RB scheme: We first derive an RB formulation of the 3D-VAR method and employ the extensive stability analysis of Section 3.3 to the RB spaces to emphasize the connection to the measurement space. We comment briefly on the numerical implementation of the RB method, before we provide an error analysis between the solutions of the original (truth) 3D-VAR problem (7) and its RB approximation. We derive computationally efficient a posteriori error bounds for the error in the model modification, state estimate, adjoint solution, and the misfit between observations in the measurement space.

We briefly comment on the notation. We require several reduced basis quantities in this section which are direct analogues of their truth counterparts, e.g., of the inf-sup constant \( \beta_T(\mu) \). In order to avoid various repetitions, quantities with an indexed \((\cdot)_R\) are defined as before with the FE spaces replaced by the corresponding RB spaces.

### 4.1 RB problem statement

The main contributions to the computational cost of the 3D-VAR method (7) come from the evaluation of the PDE over the state space \( \mathcal{Y} \) in the direction of all possible model modifications \( \mathcal{U} \). If the amount of model modifications is restricted to the most essential directions, and the solution of the PDE is approximated efficiently with a reduced order model, then the 3D-VAR saddle-point system reduces in size and may be solved faster. To this end, we assume to be given closed, non-trivial subspaces \( \mathcal{U}_R \subset \mathcal{U} \) and \( \mathcal{Y}_R \subset \mathcal{Y} \), which we refer to as RB spaces. In practice, \( \mathcal{U}_R \) and \( \mathcal{Y}_R \) are typically low-dimensional. In Section 5, we discuss how they can be chosen based on the intrinsic structure of the PDE. Note that we do not consider a reduction of the measurement space \( \mathcal{T} \), i.e., the RB problem is based on the same data and no measurements are ignored.

Our reduced-basis 3D-VAR problem has then the form

\[
\min_{(u_R, y_R, d) \in \mathcal{U}_R \times \mathcal{Y}_R \times \mathcal{T}} \frac{1}{2} ||u_R||^2_{\mathcal{U}} + \frac{\lambda}{2} ||d||^2_{\mathcal{Y}} \text{ s.t.} \\
\begin{align*}
&\quad a_{\mu}(y_R, \psi_R) = f_{bk, \mu}(\psi_R) + b_{\mu}(u_R, \psi_R) \quad \forall \psi_R \in \mathcal{Y}_R \\
&\quad (y_R + d, \tau)_\mathcal{Y} = (y_d, \tau)_\mathcal{Y} \quad \forall \tau \in \mathcal{T}.
\end{align*}
\]

(24)

Similar to before, the reduced basis minimization (24) can be written as an equivalent saddle-point problem over the spaces \( \mathcal{H}_R := \mathcal{U}_R \times \mathcal{Y}_R \) and \( \mathcal{Y}_R \): Given \( \mu \in \mathcal{C} \), find \((h^*_R, \mu, \psi^*_R, \mu) = (u^*_R, \mu, y^*_R, \mu, \mu, p^*_R, \mu) \in \mathcal{H}_R \times \mathcal{Y}_R \) such that

\[
\begin{align*}
A(h^*_R, k_R) + B_{\mu}(k_R, p^*_R, \mu) &= F(k_R) \quad \forall k_R \in \mathcal{H}_R \\
B_{\mu}(h^*_R, \psi_R) &= G_{\mu}(\psi_R) \quad \forall \psi_R \in \mathcal{Y}_R.
\end{align*}
\]

(25)

The forms \( A, B_{\mu}, F, \) and \( G_{\mu} \) are the ones defined in (10) restricted to the respective spaces.

\( \odot \) Springer
Following standard nomenclature in the RB literature, we distinguish between the original 3D-VAR formulation (7) and its RB approximation (24) by referring to the former as the truth problem and the latter as the RB problem. On a structural level, the only formal difference between the truth and the RB problem is that $\mathcal{T}$ is a closed subspace of the state space $\mathcal{Y}$, while in general $\mathcal{T} \not\subset \mathcal{Y}_R$. However, this property remained unused within the stability analysis in Section 3.3, which therefore directly applies to the RB case when $\mathcal{U}$ and $\mathcal{Y}$ are formally replaced with $\mathcal{U}_R$ and $\mathcal{Y}_R$. We therefore omit a detailed stability analysis as well as the a priori error analysis. The latter directly follows from standard a priori theory for Galerkin projections of saddle-point problems [7, 8]. More specifically, in the present case, we can even apply the results for symmetric problems with a coercive bilinear form $A$ from [12, 13].

### 4.2 Computational procedure for the RB system

In the finite-dimensional setting and after a preparatory offline phase, the RB problem (25) can be solved independently of the dimensions $N = \dim \mathcal{Y}$ and $\mathcal{M} = \dim \mathcal{U}$ of the high-fidelity spaces. For given basis functions $(\phi_{R,m})_{m=1}^M$ of $\mathcal{U}_R$ and $(\psi_{R,n})_{n=1}^N$ of $\mathcal{Y}_R$, the basis coefficients $(u^*_R, y^*_R, p^*_R) \in \mathbb{R}^{M+2N}$ of $(h^*_R, p^*_R) = (u^*_{R,\mu}, y^*_{R,\mu}, p^*_{R,\mu}) \in \mathcal{U} \times \mathcal{Y}$ can be computed by solving the linear system

\[
\begin{pmatrix}
U_R & 0 & B_R(\mu) \\
0 & \lambda P_R & A^T_R(\mu) \\
B^T_R(\mu) & A_R(\mu) & 0
\end{pmatrix}
\begin{pmatrix}
u^*_R \\
y^*_R \\
p^*_R
\end{pmatrix} =
\begin{pmatrix}
0 \\
\lambda y_{d,R} \\
f_{bk,R}(\mu)
\end{pmatrix}.
\]

(26)

Once $\mathcal{U}_R$ and $\mathcal{Y}_R$ have been chosen, e.g., with the approaches presented in Section 5, the matrices and vectors can be obtained from the previous ones in (23) by multiplication with the basis representations of $(\phi_{R,m})_{m=1}^M$ in $(\phi_{m})_{m=1}^M$ and $(\psi_{R,n})_{n=1}^N$ in $(\psi_{n})_{n=1}^N$. This computation needs only to be performed once, since the $\mu$-dependent forms $A_R(\mu)$, $B_R(\mu)$, and $f_{bk,R}(\mu)$ can be assembled from small, stored, $\mu$-independent matrices in $O(Q_aN^2 + Q_bMN + Q_fN)$ within the online phase thanks to the affine decomposition (5).

For any measurement data $\textbf{m}_j \in \mathbb{R}^L$, the vector $y_{d,R}$ can be obtained via $y_{d,R} = (S_R T^{-1}) \textbf{m}$, where $T \in \mathbb{R}^{L \times L}$, $T_{i,j} := g_I(\tau_j)$, and $(S_R)_{i,j} := (\tau_j, \psi_{R,i})|_{\mathcal{T}}$ for the basis $(\tau_l)_{l=1}^L$ of $\mathcal{T}$. Assuming that $S_R T^{-1}$ is precomputed, each new measurement set $\textbf{m}_j$ necessitates computational cost of order $O(NL)$. Once assembled, the saddle-point system (26) can be solved in $O((M + 2N)^3)$. Altogether, the online solves are completely independent of the dimensionality of the truth spaces.

### 4.3 A posteriori error estimation

Given the truth and RB saddle-point problems (12) and (25), we may follow the approach in [11, 13] to derive a posteriori error bounds for $h^*_\mu - h^*_{R,\mu}$ and $p^*_\mu - p^*_{R,\mu}$. However, these bounds provide no individual information on the error between the model corrections $u^*_\mu$ and $u^*_{R,\mu}$ or between the state estimates $y^*_\mu$ and $y^*_{R,\mu}$. We therefore pursue the approach from [19] to develop separate error bounds for the
errors $e_u := u^*_\mu - u^*_R,\mu \in \mathcal{U}$ in the model correction $e_y := y^*_\mu - y^*_R,\mu \in \mathcal{Y}$ in the state estimate and $e_p := p^*_\mu - p^*_R,\mu \in \mathcal{Y}$ in the adjoint solution. In addition, we provide an error bound for the difference $ed := d^*_\mu - d^*_R,\mu \in \mathcal{T}$ between the misfits $d^*_\mu := y_d^* - \Pi_T y^*_\mu \in \mathcal{T}$ and $d^*_R,\mu := y_d^* - \Pi_T y^*_R,\mu \in \mathcal{T}$.

We first define the residual functions

$$ru : \mathcal{U} \to \mathbb{R} : ru(\phi) := b\mu(\phi, p^*_R,\mu) - (u^*_R,\mu,\phi),$$
$$rp : \mathcal{Y} \to \mathbb{R} : rp(\psi) := \lambda(\psi, d^*_R,\mu) - a\mu(\psi, p^*_R,\mu),$$
$$ry : \mathcal{Y} \to \mathbb{R} : ry(\psi) := f_bk,\mu(\psi) + b\mu(u^*_R,\mu,\psi) - a\mu(y^*_R,\mu,\psi),$$

all three of which are linear, continuous, and $\mu$-dependent. By subtracting (25) from (12) and varying over the whole spaces $\mathcal{U}$, $\mathcal{Y}$, and $\mathcal{T}$ individually, the error tuple $(eu, ey, ed, ep) \in \mathcal{U} \times \mathcal{Y} \times \mathcal{Y} \times \mathcal{T}$ solves the variational system

$$(eu, \phi)_{\mathcal{U}} - b\mu(\phi, ep) = ru(\phi) \quad \forall \phi \in \mathcal{U},$$
$$(ep, \psi)_{\mathcal{Y}} - \lambda(ed,\psi) = rp(\psi) \quad \forall \psi \in \mathcal{Y},$$
$$(ey, \psi)_{\mathcal{Y}} - a\mu(ep,\psi) = ry(\psi) \quad \forall \psi \in \mathcal{Y},$$
$$(ed, \tau)_{\mathcal{T}} + (ey, \tau)_{\mathcal{Y}} = 0 \quad \forall \tau \in \mathcal{T}. (28d)$$

The first three equations provide alternative representations of $ru$, $rp$, and $ry$, which can be used to bound their norms in regard to the approximation errors by

$$||ru||_{\mathcal{U}'} \leq ||eu||_{\mathcal{U}} + \gamma_b(\mu) ||ep||_{\mathcal{Y}},$$
$$||rp||_{\mathcal{Y}'} \leq \gamma_a(\mu) ||ep||_{\mathcal{Y}} + \lambda ||ed||_{\mathcal{Y}},$$
$$||ry||_{\mathcal{Y}'} \leq \gamma_a(\mu) ||ey||_{\mathcal{Y}} + \gamma_b(\mu) ||eu||_{\mathcal{U}}.$$

(29)

Hence, small errors are reflected by the residual norms. In case the truth and the RB solution coincide, $ru$, $rp$, and $ry$ equal zero. Following the alternative approach in [19], we obtain the following:

**Theorem 5** Let $\mu \in C$ be given and define

$$p_u := ||ru||_{\mathcal{U}'} + \frac{\gamma_b(\mu)}{\alpha_a(\mu)} ||rp||_{\mathcal{Y}}$$
$$q_u := \frac{2}{\alpha_a(\mu)} ||rp||_{\mathcal{Y}'} ||ry||_{\mathcal{Y}} + \frac{\lambda}{4\alpha_a(\mu)^2} ||ry||_{\mathcal{Y}'}^2$$
$$p_d := \frac{1}{\alpha_a(\mu)} ||ry||_{\mathcal{Y}'}$$
$$q_d := \frac{2}{\lambda \alpha_a(\mu)} ||rp||_{\mathcal{Y}'} ||ry||_{\mathcal{Y}} + \frac{1}{4\lambda} p_u^2.$$  

Then for the unique solution $(e_u, e_y, e_d, e_p) \in \mathcal{U} \times \mathcal{Y} \times \mathcal{Y} \times \mathcal{T}$ of (4.3), we have

$$||e_u||_{\mathcal{U}} \leq \frac{1}{2} p_u + \sqrt{\frac{1}{4} p_u^2 + q_u}$$
$$||e_y||_{\mathcal{Y}} \leq \frac{1}{2} \alpha_a(\mu) ||rp||_{\mathcal{Y}'} + \frac{\gamma_b(\mu)}{\alpha_a(\mu)} ||eu||_{\mathcal{U}}$$
$$||e_d||_{\mathcal{Y}} \leq \frac{1}{2} p_d + \sqrt{\frac{1}{4} p_d^2 + q_d}$$
$$||e_p||_{\mathcal{Y}} \leq \frac{1}{2} \alpha_a(\mu) ||rp||_{\mathcal{Y}'} + \frac{\lambda}{\alpha_a(\mu)} ||ed||_{\mathcal{Y}}.$$  

(30)
Proof. To bound $||e_p||_Y$, we use the coercivity (3) of $a_\mu$ and set $\psi := e_p$ in (28b) to obtain
\[
\alpha_a(\mu) ||e_p||_Y \leq \frac{a_\mu(e_p, e_p)}{||e_p||_Y} = \frac{r_p(e_p) + \lambda(e_d, e_p)_Y}{||e_p||_Y} \leq ||r_p||_Y + \lambda ||e_d||_Y.
\] (31)

The bound for $||e_p||_Y$ is derived similarly with $\psi := e_y \in Y$ in (28c); we thus get
\[
\alpha_a(\mu) ||e_y||_Y \leq \frac{\alpha_a(e_y, e_y)}{||e_y||_Y} = \frac{r_y(e_y) + b_\mu(e_u, e_y)}{||e_y||_Y} \leq ||r_y||_Y + \gamma_b(\mu) ||e_u||_U.
\] (32)

For $||e_u||_U$ and $||e_d||_Y$, we start with the choice $\phi = e_u \in U$ in (28a),
\[
||e_u||_U^2 \overset{(28a)}{=} r_u(e_u) + b_\mu(e_u, e_p)
\]
\[
\overset{(28c)}{=} r_u(e_u) - r_y(e_p) + a_\mu(e_y, e_p)
\]
\[
\overset{(28b)}{=} r_u(e_u) - r_y(e_p) + r_p(e_y) + \lambda(e_d, e_y)_Y
\]
\[
\overset{(28d)}{=} r_u(e_u) - r_y(e_p) + r_p(e_y) - \lambda ||e_d||_Y^2
\]
\[
\leq ||r_u||_U^2 ||e_u||_U + ||r_y||_Y ||e_y||_Y + ||r_p||_Y ||e_y||_Y - \lambda ||e_d||_Y^2.
\] (33)

The use of the previous bounds (31) and (32) in the second and third term yields
\[
||e_u||_U^2 + \lambda ||e_d||_Y^2 \leq (||r_u||_U^2 + \frac{\gamma_b(\mu)}{\alpha_a(\mu)} ||r_p||_Y ||e_d||_Y) ||e_u||_U + 2 ||r_p||_Y ||r_y||_Y + \frac{\lambda ||r_y||_Y^2}{\alpha_a(\mu)} + \frac{\lambda ||r_y||_Y^2}{\alpha_a(\mu)} ||e_d||_Y.
\] (34)

For the last term, we have $\frac{\lambda}{\alpha_a(\mu)} ||r_y||_Y ||e_d||_Y \leq \frac{\lambda}{4\alpha_a(\mu)} ||r_y||_Y^2 + \lambda ||e_d||_Y^2$ by Young’s inequality. After subtracting $\lambda ||e_d||_Y^2$ from either side of (34), we have
\[
||e_u||_U^2 \leq (||r_u||_U^2 + \frac{\gamma_b(\mu)}{\alpha_a(\mu)} ||r_p||_Y ||e_d||_Y) ||e_u||_U + 2 \frac{2}{\alpha_a(\mu)} ||r_p||_Y ||r_y||_Y + \frac{\lambda ||r_y||_Y^2}{4\alpha_a(\mu)^2},
\]
which yields the bound for $||e_u||_U$ with the quadratic formula. For the bound of $||e_d||_Y$, we use Young’s inequality on the first term in the inequality (34) to get
\[
(||r_u||_U^2 + \frac{\gamma_b(\mu)}{\alpha_a(\mu)} ||r_p||_Y ||e_u||_U) ||e_u||_U \leq \frac{1}{4} (||r_u||_U^2 + \frac{\gamma_b(\mu)}{\alpha_a(\mu)} ||r_p||_Y^2)^2 + ||e_u||_U^2.
\]

Then, $||e_u||_U^2$ can be subtracted from either side of (34), which leaves
\[
\lambda ||e_d||_Y^2 \leq \frac{\lambda ||r_y||_Y^2}{\alpha_a(\mu)} ||e_d||_Y + \frac{2}{\alpha_a(\mu)} ||r_p||_Y ||r_y||_Y + \frac{1}{4} (||r_u||_U^2 + \frac{\gamma_b(\mu)}{\alpha_a(\mu)} ||r_p||_Y^2)^2.
\]

The bound for $||e_d||_Y$ can now also be obtained with the quadratic formula.

The error bounds in Theorem 5 can be used in the finite-dimensional setting to bound the approximation error without computing the truth solution. First, note that the bounds decrease with the coercivity constant $\alpha_a(\mu)$, which can therefore be substituted with its efficiently computable lower bound $\alpha_a^{LB}(\mu) \leq \alpha_a(\mu)$. Analogously,
\( \gamma_b(\mu) \) can be replaced with its upper bound \( \gamma_b^{UB}(\mu) \). Due to the affine parameter dependence assumed in (5), the residual norms can be computed from (27) in an offline-online procedure independent of the space dimensions \( N = \dim Y \) and \( M = \dim U \). Their computation is a standard procedure in the RB literature and therefore omitted (see, e.g., [34]). To summarize, we then obtain a posteriori bounds \( \Delta_1u(\mu) \), \( \Delta_1y(\mu) \), \( \Delta_1d(\mu) \), and \( \Delta_1p(\mu) \) for \( ||e_u||_U, ||e_y||_Y, ||e_d||_Y, \) and \( ||e_p||_Y \) that may be computed fast and efficiently. We note that the boundedness (29) of the residual norms with respect to the errors ensures a correspondence between the a posteriori error bounds and the behavior of the true error.

5 Space construction

In the previous sections, we have assumed the measurement space \( T \) and the RB spaces \( U_R \) and \( Y_R \) to be given. In this section, we present ideas on how the previously discussed properties concerning stability and approximation quality may be used for constructing the spaces. More specifically, we first introduce a greedy-OMP algorithm for the stability-based generation of \( T \) and subsequently present an approach for the construction of the reduced basis spaces.

5.1 Greedy-OMP algorithm

In some applications, prior information can allow for a low-dimensional model correction space \( U \). One reason would be that, due to the known presence of noise, the focus of interest lies in the large-scale behavior of the model correction rather than noise-sensitive details and oscillations. Such a low-dimensional choice can be of great benefit, since it can counteract noise amplification already for low-dimensional measurement spaces, as indicated by our stability analysis in Section 3.3. In the following, we propose a greedy-OMP algorithm that chooses a measurement space \( T \) such that (1) the influence of noise upon the 3D-VAR model correction and state estimates is bounded independently of \( \lambda \) and \( \mu \) and (2) solutions of the modified model (6) are distinguishable for different parameters.

We presuppose that \( M := \dim U \) is small and that \( Y \) can sufficiently capture the modifications induced by \( U \) so that \( \eta(\mu) > 0 \) holds uniformly on \( \mathcal{C} \). Then, by the stability analysis in Section 3.3, the amplification of noise is bounded for the 3D-VAR model correction and state solution \( h^*_\mu \) \( \lambda \)-independently if \( \beta_T(\mu) > 0 \). For any fixed parameter \( \mu \), we can compute the space \( Y_\mu \) of state modifications by considering the forward problem

\[
\text{For } u_0 \in U, \text{ find } y_\mu(u_0) \in Y \text{ s.t. } a_\mu(y_\mu(u_0), \psi) = b_\mu(u_0, \psi) \quad \forall \psi \in Y. \quad (35)
\]

Then, \( Y_\mu = \text{span}(y_\mu(\phi_m))_{m=1}^M \) for any basis \( (\phi_m)_{m=1}^M \) of \( U \).

We assume to be given a library \( \mathcal{L} \subset Y' \), from which we may select different measurement functionals \( g_l \) whose Riesz representation then spans \( T \) (see (1)). For any fixed parameter \( \mu \), a generalized OMP algorithm can be applied to the space \( Y_\mu \) to iteratively expand \( T \) and—under some restrictions upon the library—enforce \( \beta_T(\mu) > 0 \) for \( L = \dim T \geq M = \dim Y_\mu \) (see [6]). Theoretically, \( T \) can be
expanded by repeating this procedure for sufficiently many parameters. However, the necessity to solve the forward problem (35) for \( u_0 = \phi_m, m = 1, \ldots, M \) and each parameter accumulates a high computational cost.

We can improve this procedure with our RB approximation: We set \( \mathcal{U}_R := \mathcal{U} \) and assume to have an RB space \( \mathcal{Y}_R \subset \mathcal{Y} \) with \( \eta_R(\mu) > 0 \) that sufficiently approximates \( \mathcal{Y}_\mu \) in the following sense: For any \( \mu \in \mathcal{C} \), there exists an \( \varepsilon_\mu, 0 \leq \varepsilon_\mu \ll 1 \), with:

\[
\text{If } (u, y) \in \mathcal{H}^0(\mu) \text{ and } (u, y_R) \in \mathcal{H}^0_R(\mu) \text{ then } ||y - y_R||_\mathcal{Y} \leq \varepsilon_\mu ||y||_\mathcal{Y}.
\] (36)

We emphasize that the RB spaces here may be considered completely separate from the RB 3D-VAR method. We use the same notation as in Section 4 only for simplicity, as it avoids the needless repetition of definitions and results.

For \( u_0 \in \mathcal{U} = \mathcal{U}_R \), let \( y(u_0) \in \mathcal{Y}_\mu \) and \( y_R(u_0) \in \mathcal{Y}_R, \mu \) denote the unique elements with \( (u_0, y(u_0)) \in \mathcal{H}^0(\mu) \) and \( (u_0, y_R(u_0)) \in \mathcal{H}^0_R(\mu) \), respectively. Then

\[
\frac{||y_R(u_0)||_\mathcal{Y}}{||y(u_0)||_\mathcal{Y}} \geq \frac{||y(u_0)||_\mathcal{Y} - ||y(u_0) - y_R(u_0)||_\mathcal{Y}}{||y(u_0)||_\mathcal{Y}} \geq 1 - \varepsilon_\mu.
\] (37)

The inf-sup constant \( \beta(\mu) \) on the high-fidelity space \( \mathcal{Y} \) can then be bounded with respect to \( \beta_{\mathcal{T},R}(\mu) \) on the RB space \( \mathcal{Y}_R \) via

\[
\beta(\mu) = \inf_{u_0 \in \mathcal{U}} \sup_{\tau \in \mathcal{T}} \frac{(y(u_0), \tau)}{||y(u_0)||_\mathcal{Y} ||\tau||_\mathcal{Y}} = \inf_{u_0 \in \mathcal{U}} \sup_{\tau \in \mathcal{T}} (y(u_0), \tau)_{\mathcal{Y}} + \frac{||y(u_0) - y_R(u_0)||_\mathcal{Y} ||\tau||_\mathcal{Y}}{||y(u_0)||_\mathcal{Y} ||\tau||_\mathcal{Y}} \geq \inf_{u_0 \in \mathcal{U}} \sup_{\tau \in \mathcal{T}} (y_R(u_0), \tau)_{\mathcal{Y}} - \varepsilon_\mu \geq (1 - \varepsilon_\mu) \inf_{u_0 \in \mathcal{U}} \sup_{\tau \in \mathcal{T}} (y_R(u_0), \tau)_{\mathcal{Y}} - \varepsilon_\mu = (1 - \varepsilon_\mu) \beta_{\mathcal{T},R}(\mu) - \varepsilon_\mu.
\]

Hence, \( \mathcal{T} \) is a stabilizing choice for the truth 3D-VAR method, if the inf-sup condition \( \beta_{\mathcal{T},R}(\mu) > 0 \) holds and \( \varepsilon_\mu \) is sufficiently small.

The greedy-OMP algorithm 1 generates a measurement space \( \mathcal{T} \) from the library \( \mathcal{L} \subset \mathcal{Y} \) to increase \( \beta_{\mathcal{T},R}(\mu) \) over a training set \( \mathcal{X}_{\text{train}} \subset \mathcal{C} \). The initial computation of \( \mathcal{Y}_R, \mu \) for \( \mu \in \mathcal{X}_{\text{train}} \) (line 1), and the searches for \( y_L \) (line 5) and \( \mu_{L+1} \) (line 8) in each iteration, requires only online operations, i.e., independent of the FE dimensions. The selection of the measurement functional \( B_L \) (line 6) depends on the FE dimension \( \mathcal{N} = \dim \mathcal{Y} \) and the size of the library, and the expansion of \( \mathcal{T} \) (line 7) needs \( \mathcal{N} \)-dependent computations for the Riesz representation and orthonormalization. However, both (lines 6 and 7) need to be performed only once per iteration. In the algorithm, the measurement functional \( B_L \) is chosen in lines 5 and 6 so that the expansion of \( \mathcal{T} \) in line 7 increases \( \beta_{\mathcal{T},R}(\mu_L) \). This is an application to the space \( \mathcal{Y}_R, \mu_L \) of the worst-case OMP algorithm in [6], which is in turn a greedy generalization of classical OMP algorithms. Other selection strategies for increasing the inf-sup constant with regard to measurement functionals exist and can be substituted in lines 5 and 6; we refer to [6] and [27] for further options and both analytical and numerical comparisons.
Algorithm 1 Greedy orthogonal matching pursuit.

Require: \( \Xi_{\text{train}} \subset \mathcal{C} \) training set, \( \mathcal{L} \subset \mathcal{Y}^\prime \) library, \( \beta_0 > 0 \) target value, \( \mu_1 \in \Xi_{\text{train}} \), \( L_{\text{max}} \)

1: Compute \( \mathcal{Y}_{\mathcal{R},\mu} \) for each \( \mu \in \Xi_{\text{train}} \)
2: \( \mathcal{T} \leftarrow \{0\}, L \leftarrow 0, \beta \leftarrow 0 \)
3: while \( \beta \leq \beta_0 \) and \( L \leq L_{\text{max}} \) do
4: \( L \leftarrow L + 1 \)
5: \( y_L \leftarrow \arg\max \{||y - \Pi_T y||y : y \in \mathcal{Y}_{\mathcal{R},\mu_L}, ||y||y = 1\} \)
6: \( g_L \leftarrow \arg\max_{g \in \mathcal{L}} \|g(y_L - \Pi_T y_L)||g||y' \)
7: Expand \( \mathcal{T} \) with the Riesz representation of \( g_L \)
8: Find \( \mu_{L+1} \in \arg\min_{\mu \in \Xi_{\text{train}}} \beta_{\mathcal{T},\mathcal{R}}(\mu) \)
9: \( \beta \leftarrow \beta_{\mathcal{T},\mathcal{R}}(\mu_{L+1}) \)
10: end while

If the 3D-V AR method is to be employed for parameter estimation, then \( \mathcal{T} \) should be able to distinguish between solutions of the modified model (6) for different parameters. To this end, we define, for \( \mu, \nu \in \mathcal{C} \),

\[
\beta_{\mathcal{T}}(\mu, \nu) := \inf_{y \in \mathcal{Y}_{\mathcal{R},(\mu,\nu)}} \sup_{\tau \in \mathcal{T}} \frac{(y, \tau)_Y}{||y||_Y||\tau||_Y}
\]

with \( \mathcal{Y}_{\mathcal{R},(\mu,\nu)} := \text{span} \{ y_{\mathcal{R},\mu}, y_{\mathcal{R},\nu}, \mathcal{Y}_{\mathcal{R},\mu}, \mathcal{Y}_{\mathcal{R},\nu} \} \).

Here, \( y_{\mathcal{R},\mu} \) and \( y_{\mathcal{R},\nu} \) are the best-knowledge solutions of (2) for the respective parameters. If \( y_{\mu} = y_{\mu}(u_1) \) and \( y_{\nu} = y_{\nu}(u_2) \) are solutions of the modified model (6) for parameters \( \mu, \nu \in \mathcal{C} \) and modifications \( u_1, u_2 \in \mathcal{U} \), then \( y_{\mu}, y_{\nu} \in \mathcal{Y}_{\mathcal{R},(\mu,\nu)} \). The experimentally observable difference \( ||\Pi_T (y_{\mu} - y_{\nu})||y \geq \beta_{\mathcal{T}}(\mu, \nu)||y_{\mu} - y_{\nu}||y \) is hence bounded from below relative to their actual distance. For parameter estimation, \( \mathcal{T} \) should thus be chosen such that \( \beta_{\mathcal{T}}(\mu, \nu) > 0 \) on \( \mathcal{C} \times \mathcal{C} \). We can generate \( \mathcal{T} \) similarly to before, with a slight modification of the greedy-OMP algorithm 1: In addition to \( \beta_{\mathcal{T},\mathcal{R}}(\mu) \), we consider

\[
\beta_{\mathcal{T},\mathcal{R}}(\mu, \nu) := \inf_{y \in \mathcal{Y}_{\mathcal{R},(\mu,\nu)}} \sup_{\tau \in \mathcal{T}} \frac{(y, \tau)_Y}{||y||_Y||\tau||_Y}
\]  

over a training set in \( \mathcal{C} \times \mathcal{C} \) with \( \mathcal{Y}_{\mathcal{R},(\mu,\nu)} := \text{span} \{ y_{\mathcal{R},\mu}, y_{\mathcal{R},\nu}, \mathcal{Y}_{\mathcal{R},\mu}, \mathcal{Y}_{\mathcal{R},\nu} \} \) with RB approximations \( y_{\mathcal{R},\mu} \) and \( y_{\mathcal{R},\nu} \) of the best-knowledge states \( y_{\mathcal{R},\mu} \) and \( y_{\mathcal{R},\nu} \). Note that this additionally requires \( \mathcal{Y}_{\mathcal{R}} \) to approximate the best-knowledge model (2). The spaces \( \mathcal{Y}_{\mathcal{R},(\mu,\nu)} \) can be computed efficiently online from the RB space \( \mathcal{Y}_{\mathcal{R}} \). It is therefore possible to consider different sets of training parameters in the course of the algorithm to reduce the required amount of memory.

5.2 Construction of reduced basis spaces

We next consider the construction of the reduced basis spaces \( \mathcal{Y}_{\mathcal{R}} \) and \( \mathcal{U}_{\mathcal{R}} \). If the measurement space \( \mathcal{T} \) and the measurements—and thus the data state \( y_d \in \mathcal{T} \)—are known a priori, we can directly use a greedy approach similar to the one for optimal
control problems in [19] (see [31]). However, here we propose a different approach which does not require $\mathcal{T}$ or $y_d$ to be known a priori.

By varying in (12) separately over the spaces $\mathcal{U}$ and $\mathcal{Y}$ and inserting $d^*_\mu = y_d - \Pi_{\mathcal{T}} y^*_\mu$, the following system is obtained for the model correction $u^*_\mu$, state $y^*_\mu$, and adjoint solution $p^*_\mu$:

\begin{align}
(u^*_\mu, \phi)_{\mathcal{U}} - b_\mu(\phi, p^*_\mu) &= 0 & \quad \forall \phi \in \mathcal{U} \\
 a_\mu(\psi, p^*_\mu) - \lambda(\psi, d^*_\mu)_{\mathcal{Y}} &= 0 & \quad \forall \psi \in \mathcal{Y} \\
 a_\mu(y^*_\mu, \psi) - b_\mu(u^*_\mu, \psi) &= f_{\text{bk}, \mu}(\psi) & \quad \forall \psi \in \mathcal{Y} \\
 (y^*_\mu + d^*_\mu, \tau)_{\mathcal{Y}} &= (y_d, \tau)_{\mathcal{Y}} & \quad \forall \tau \in \mathcal{T}.
\end{align}

We assume that $\mathcal{U}_R \subset \mathcal{U}$ is fixed and low-dimensional so that $u^*_\mu$ can be approximated well enough in $\mathcal{U}_R$ for the expected range of measurement data and the desired level of detail. As $u^*_\mu$ is not known, the RB space $\mathcal{Y}_R$ needs to provide good approximations to the solution of (39c) for each model modification in $\mathcal{U}_R$ relative to its norm. We realize this by constructing a state space $\mathcal{Y}_y \subset \mathcal{Y}$ as an RB approximation of the forward problem

For $\mu \in \mathcal{C}$ and $f \in \{ f_{\text{bk}, \mu} \} \cup \{ b_\mu(\phi_{\mathcal{R}, m}, \cdot) : 1 \leq m \leq M \} \subset \mathcal{Y}'$:

Find $y \in \mathcal{Y}$ s.t. $a_\mu(y, \psi) = f(\psi) \quad \forall \psi \in \mathcal{Y}$. (40)

If $\mathcal{T}$ is not fixed by the experimental design, it can now be generated from $\mathcal{Y}_y$ with the greedy-OMP algorithm 1. Given $\mathcal{T}$ and following an analogous argument as before for the generation of $\mathcal{Y}_y$, we can obtain an adjoint space $\mathcal{Y}_p \subset \mathcal{Y}$ if we iteratively replace $d^*_\mu$ in (39b) with orthonormal basis functions $(\tau_l)_{l=1}^L$ of $\mathcal{T}$ and perform an RB approximation of each equation over the parameter domain $\mathcal{C}$. By considering a relative target accuracy, we can use $\lambda = 1$ for the RB space generation, but note that since the approximation quality of the adjoint solution $p^*_\mu$ scales with $\lambda$, the target accuracy for the RB adjoint approximation should be chosen with respect to the largest regularization parameter $\lambda$ that the RB 3D-VAR method is expected to be used for. Finally, we set $\mathcal{Y}_R := \mathcal{Y}_y + \mathcal{Y}_p$.

The following sketch summarizes this consecutive construction of the spaces:

$\mathcal{U}_R \xrightarrow{(39c)} \mathcal{Y}_y \xrightarrow{\text{Alg. 1}} \mathcal{T} \xrightarrow{(39b)} \mathcal{Y}_p \quad \mathcal{Y}_R = \mathcal{Y}_y + \mathcal{Y}_p$.

This stepwise selection of the spaces avoids two possible drawbacks: First, the data state $y_d$ needs not to be known at the start of the offline phase; the RB spaces are hence not influenced by measurement noise. Second, once the spaces are constructed, they can be used repeatedly for different measurement data $y_d$ without necessitating additional offline costs.

As $\mathcal{Y}_R = \mathcal{Y}_y + \mathcal{Y}_p$ comprises $M + L + 1$ RB approximation spaces, $\mathcal{Y}_R$ may become large. The target accuracy and the training set should thus be chosen carefully to reduce the offline computational time. Once the data state $y_d$ becomes known, $\mathcal{Y}_R$ can be condensed by following the two-step RB approach in [10], i.e., by subsequently applying a greedy algorithm to derive spaces of smaller dimensions.
6 Numerical results

In this section, we present numerical experiments to verify our theoretical results. All computations were performed with MATLAB® on a computer with 2.5 GHz Intel Core i5 processor and 4 GB of RAM.

6.1 Model description

We consider the steady-state temperature distribution \( y \) of a thermal block \( \Omega = (0, 1) \times (0, 1) \) that comprises up to three different material properties within \( \Omega_1 \), \( \Omega_2 \), and \( \Omega_0 \). We consider parameters \( \mu = (\mu_1, \mu_2) \in C := [1, 10]^2 \), where \( \mu_i, i \in \{1, 2\} \) is the ratio of the thermal conductivity of \( \Omega_i \) to \( \Omega_0 \). We divide the outer boundary \( \partial \Omega = \Gamma_{in} \cup \Gamma_D \cup \Gamma_N \), with \( \Gamma_N = \{0, 1\} \times [0, 1] \), \( \Gamma_D = [0, 1] \times \{1\} \), and \( \Gamma_{in} = [0, 1] \times \{0\} \), and confer different boundary conditions on each by \( \nabla y \cdot n = 0 \) a.e. on \( \Gamma_N \) (zero Neumann flux), \( y|_{\Gamma_D} = 0 \) a.e. on \( \Gamma_D \) (zero Dirichlet), and \( \nabla y \cdot n = u \) a.e. on \( \Gamma_{in} \) (Neumann flux). Here, \( n \) is the outer unit normal to \( \Omega \) and \( u \) is a yet unspecified function in \( U_c := L^2(\Gamma_{in}) \). Across the subdomains, we require that the temperature \( y \) as well as the heat flux is continuous. An outline of this setup is provided in Fig. 1a.

For the state space, we define \( Y_c := \{ y \in H^1(\Omega) : y|_{\Gamma_D} = 0 \} \) with inner product \( (y, w)_{Y_c} := \int_{\Omega} \nabla y(x) \cdot \nabla w(x) dx \). For the space discretization in the state variable, we use a linear finite element space \( Y \) of dimension 4210. The bilinear forms \( a_{\mu} : Y_c \times Y_c \to \mathbb{R} \) and \( b : U_c \times Y_c \to \mathbb{R} \) associated with the thermal block problem are then given by

\[
\begin{align*}
a_{\mu}(y, w) & := \sum_{i=0}^{2} \mu_i \int_{\Omega_i} \nabla y(x) \cdot \nabla w(x) dx, \\
b(u, w) & := \int_{\Gamma_{in}} u(x) w(x) dS(x),
\end{align*}
\]

Fig. 1 a Illustration of the domain and the imposed outer boundary conditions. b Greedy placement of the measurement sensors
where $\mu_0 := 1$. Both forms are continuous, and $a_\mu$ is coercive with $\alpha_a(\mu) = \min\{1, \mu_1, \mu_2\} \geq 0.1$.

For the state and parameter estimation, we assume that the unknown true state $y_{\text{true}} \in \mathcal{Y}$ is given as the unique solution of the weak PDE

$$a_{\mu_{\text{true}}}(y_{\text{true}}, \psi) = b(u_{\text{true}}, \psi) \quad \forall \psi \in \mathcal{Y},$$

(41)

with $u_{\text{true}}(x_1) := 1.5 + 0.3 \sin(2\pi x_1)$ for $x_1 \in \Gamma_{\text{in}}$ and $\mu_{\text{true}} = (7, 0.3)$. For the numerical computation of this state, we approximate $u_{\text{true}}$ with 69 linear finite elements on $\Gamma_{\text{in}}$. For the best-knowledge model, we presume a steady, homogeneous heat inflow in the form of a uniform Neumann flux $u_{\text{start}} \equiv 1$ on $\Gamma_{\text{in}}$. We hence obtain $f_{\beta_{\text{bk}}, \mu} = b(u_{\text{start}}, \cdot) \in \mathcal{Y}'$. In this numerical experiment, we aim to estimate $\mu_{\text{true}}$, $u_{\text{true}}$, and $y_{\text{true}}$ from a few measurements of $y_{\text{true}}$.

### 6.2 Space generation with prior knowledge

In this section, we first specify the framework for the application of the 3D-VAR method. We then use the approach described in the previous section to generate the RB spaces and the measurement space $\mathcal{T}$.

Due to the diffusion of heat, state changes brought forth by local details in the Neumann flux smooth out as the distance to the boundary $\Gamma_{\text{in}}$ increases. The observation of such details would hence necessitate the placement of observation functionals very close to the boundary, and the local reconstruction of the Neumann flux would then be sensitive to noise. For the model modifications, we hence make the educated guess that $u_{\text{true}}$ can be approximated sufficiently in the space $P_3(\Gamma_{\text{in}})$ of polynomials on $\Gamma_{\text{in}}$ with a degree smaller or equal to 3; accordingly, we fix $\mathcal{U} := P_3(\Gamma_{\text{in}})$ equipped with the $L^2(\Gamma_{\text{in}})$ inner product. As a consequence of this decision, we accept that we can approximate $u_{\text{true}}$ only up to the accuracy $||\Pi_{\mathcal{U}^\perp} u_{\text{true}}||_{L^2(\Gamma_{\text{in}})} \approx 1.9877e-02$. Similarly, for the true parameter $\mu_{\text{true}}$, $y_{\text{true}}$ can only be approximated up to $||\Pi_{\mathcal{Y}^\perp} y_{\text{true}}||_{\mathcal{Y}} \approx 4.7576e-03$.

We next generate the measurement space $\mathcal{T}$ and the RB spaces $\mathcal{U}_R$ and $\mathcal{Y}_R$. Since we already restricted $\mathcal{U}$ to a small dimension, we chose $\mathcal{U}_R := \mathcal{U}$ without discarding any further model modifications. The state space $\mathcal{Y}_y \subset \mathcal{Y}$ is generated from the first four Legendre polynomials by a weak greedy algorithm over a $41 \times 41$ regular training grid on the logarithmic parameter domain with target accuracy $10^{-5}$ relative to the norm of the state solution. The algorithm terminated with $\text{dim} \mathcal{Y}_y = 64$. We note that if we were to use this state space for a $\mu$-independent PBDW state estimate of $y_{\text{true}}$, we would need at least 64 measurement functionals for well-posedness, whereas any number of measurements is sufficient for well-posedness of the 3D-VAR method.

The measurement space $\mathcal{T}$ was obtained from $\mathcal{Y}_y$ with the greedy-OMP algorithm 1, which selected 16 measurement functionals $g_l \in \mathcal{Y}'$, $l \in \{1, \ldots, 16\}$. The library consists of Gaussian functionals in $\mathcal{Y}'$ with standard deviation 0.01 and centers in a $97 \times 97$ regular grid on $(0.02, 0.98)^2 \subset \Omega$. We use the target value $\beta_0 := 0.5$ for the inf-sup constants $\beta_{\mathcal{T}, \mathcal{R}}(\mu)$ and $\beta_{\mathcal{T}, \mathcal{R}}(\mu, \nu)$ with parameters $\mu, \nu$ in a $21 \times 21$ grid on $\mathcal{C}$. Figure 2a shows the development of the minimal inf-sup constants $\beta_{\mathcal{T}, \mathcal{R}}(\mu)$ and
Development of the target variables of the greedy-OMP algorithm versus the dimension of \( T \).

(a) Greedy-OMP Algorithm

(b) \( H^0(\mu) \)-coercivity

\[ \min_{\mu, \lambda} \lambda(\mu) \quad \text{versus} \quad \dim(T) \]

\[ \min_{\mu, \lambda} \lambda(\mu) \quad \text{versus} \quad \dim(T) \]

\[ \alpha_0^A(\mu, \lambda) \quad \text{versus} \quad \lambda \]

\[ \alpha_{LB}^A(\lambda, \eta(\mu), \eta(\mu), T(\mu)) \] for \( \mu = \mu_{\text{true}} \)

\( \beta_{T, R}(\mu, v) \) over the respective training set as the greedy-OMP algorithm 1 expands \( T \). Due to the margin at the domain boundary for the placement of the sensors, we cannot expect the inf-sup constant to reach 1 asymptotically. Since \( \min_{\mu} \beta_{T, R}(\mu) \) exceeds the target value already for five measurement functionals, most measurements have been chosen to increase \( \beta_{T, R}(\mu, v) \) over \( C^2 \). The centers of the chosen measurement functionals are indicated in Fig. 1b. The four measurements near \( \Gamma_{\text{in}} \) are most important for determining the optimal model correction, whereas the other measurements are most important for estimating unknown parameters. In Fig. 2b, the \( H^0(\mu) \)-coercivity constant \( \alpha_0^A(\mu, \lambda) \) of \( A \) and the corresponding lower bound \( \alpha_{LB}^A(\lambda, \eta(\mu), \eta(\mu), T(\mu)) \) for \( \mu = \mu_{\text{true}} \) is larger than \( \alpha_{LB}^A \) by a factor of 1.0325.

To finish the generation of \( \mathcal{Y}_R \), we expand the state space \( \mathcal{Y}_y \) with an adjoint space \( \mathcal{Y}_p \) with \( \dim \mathcal{Y}_p = 95 \). This space is computed from an orthonormal basis of \( T \) with an RB approximation of the adjoint (39b). Since prior tests indicate that due to the local influence of the measurement functionals, the weak greedy algorithm only chooses training parameters very close to the boundary \( \partial C \), the computation of \( \mathcal{Y}_p \) was done using a relative target accuracy of \( 10^{-5} \) on 40 regularly spaced training parameters on \( \partial C \) in the logarithmic plane. After the computation of \( \mathcal{Y}_p \), the target accuracy was confirmed on a fine test grid over the whole parameter domain. Altogether, the offline phase for the RB space generation, including the selection of appropriate measurement functionals, finished after 463 seconds with \( \dim \mathcal{U}_R = 4, \dim T = 16 \) and \( \dim \mathcal{Y}_R = 159 \).

### 6.3 Reduced basis approximation

In this section, we evaluate (1) the computational efficiency of the RB solution, (2) its accuracy and the effectivity of the a posteriori error bounds, and (3) the
effects of noise on the approximation of $u_{\text{true}}$. To this end, we take measurements $\mathbf{m}_d = (g_l(y_{\text{true}}))_{l=1}^{16} \in \mathbb{R}^{16}$ of the true state to compute $y_d = \Pi_T y_{\text{true}}$. Noisy measurements are constructed by adding a random variable drawn from the probability distribution $\mathcal{N}(0, 0.01^2)$ upon each measurement before the computation of $y_d$. This corresponds to a noise level of approximately 1.5% in each measurement compared with the difference to the measurements of the best-knowledge state $y_{bk, \mu_{\text{true}}}$.

First, we evaluate the computational efficiency of the RB method in comparison to the truth 3D-VAR method: For 200 random parameters in $C$ and $\lambda = 100$, the truth 3D-VAR solution was computed from different noisy measurements. The mean computation time was 7.08 s. The RB 3D-VAR solution was then computed over the same parameters with the same noisy data for comparison. With an average online computation time of 4.2 ms for the computation of the RB 3D-VAR solution and

![Fig. 3](https://example.com/fig3.png)

**Fig. 3** Mean relative error (blue, dashed, o-marks) and relative error bound (blue, dashed, x-mark) in the RB approximation for $\lambda = 100$ from noisy data for the model correction, state, observable misfit, and adjoint for the stepwise approach with different $U_R = \mathbb{P}_j(G_m)$, $j = 1, \ldots, 4$. In each plot, the red, continuous, o-marked curve marks the maximum of the relative truth error (labeled shorthand as evaluation at $\mu_{\text{max}}$) for each $j$, whereas the red, continuous, x-marked curve shows the corresponding relative error bound evaluated at the same parameter and the same $j$. The mean and maximum are computed over 200 random parameters. **a** Model correction. **b** State. **c** Observable misfit. **d** Adjoint.
1.3 ms for the computation of the a posteriori error bounds, the RB method showed a mean speedup of 1,276.

To assess the effectivity of the a posteriori error bounds, we generate additional pairs $(U_j^R, Y_j^R) \subset U \times Y$, $j \in \{1, 2, 3\}$ of RB spaces with $U_j^R = \mathbb{P}_{j-1}(\Gamma_{in})$ the span of polynomials on $\Gamma_{in}$ of a degree smaller or equal to $j - 1$. We kept the measurement space $T$ fixed but otherwise followed the approach as outlined in Section 6.2 to obtain $Y_j^R$. For the same parameters and the same noisy data as before, we evaluate the error between the truth 3D-VAR solution and the RB solution on these additional spaces. Figure 3 shows the mean error and mean a posteriori error bound relative to the norm of the solution variable versus the dimension $j = \dim U_j^R$ of the RB model correction space. Additionally, the maximum relative error and the corresponding relative error bound are indicated. As the polynomial $x^3$ cannot be approximated in $U_j^R$ for $j \leq 3$, the large portion of $x^3$ in the truth 3D-VAR model modification $u_\mu^*$ leads to the strong decrease in the error for $\dim U_R = 4$. We observe that the effectivity of the error bounds is generally small and constant, except for the case $U_R = U$ where $e_u$ and $e_d$ lie below the natural limits of the error bounds $\Delta_u(\mu)$ and $\Delta_d(\mu)$ dictated by the

![Figure 3](image-url)

**Fig. 3** Qualitative behavior of the 3D-VAR solution as an approximation of $u_{true}$ for the parameter $\mu_{true}$.

(a) $\lambda = 1$  
(b) $\lambda = 10$  
(c) $\lambda = 100$  
(d) $\lambda = 1000$

![Figure 4](image-url)

**Fig. 4** Qualitative behavior of the 3D-VAR solution as an approximation of $u_{true}$ for the parameter $\mu_{true}$.

(a) $\lambda = 1$.  
(b) $\lambda = 10$.  
(c) $\lambda = 100$.  
(d) $\lambda = 1000$.  

© Springer
square root of the machine accuracy. For different regularization parameters \( \lambda \leq 10^3 \), we observe a posteriori error bounds of a similar magnitude.

Given that the RB model correction \( u^*_{R,\mu} \) approximates \( u^*_\mu \) so accurately, we use the RB method to investigate how well the qualitative, global behavior of the true Neumann flux \( u^*_{true} \) can be reflected by the 3D-V AR method, if the true parameter \( \mu_{true} \) is provided. We hence compute \( u^*_{R,\mu_{true}} \) for 100 different noisy data sets and for regularization parameters \( \lambda = 10^i, i = 0, ..., 3 \). The results are shown in Fig. 4. We observe that for \( \lambda = 1 \), we obtain a good compromise between \( u^*_{true} \) and our initial guess \( u_{start} \equiv 1 \) in the best-knowledge model, and the measurement error only leads to small deviations. As we increase \( \lambda \) in favor of the biased measurements, the estimates start to differ, but, as can be seen when comparing \( \lambda = 10^2 \) to \( \lambda = 10^3 \), the difference to the noise-free solution is bounded.

6.4 Parameter estimation

We finally test three different approaches for parameter estimation with the 3D-V AR method to obtain approximations for \( \mu_{true} = (7, 0.3) \), \( u^*_{true} \), and \( y^*_{true} \); to assess the estimation quality in general, we first use noise-free data. As the involved minimization process takes a long time to converge, we employ the RB 3D-V AR method for comparison. Finally, we evaluate the sensitivity of the parameter estimate with respect to noise in the data.

We define three different cost functionals \( J^\lambda_i : \mathcal{C} \to \mathbb{R} \) by

\[
J^\lambda_1(\mu) := \frac{1}{2} ||u^*_\mu||^2_U + \frac{\lambda}{2} ||d^*_\mu||^2_Y \\
J^\lambda_2(\mu) := \frac{1}{2} ||v_d - (g_i(y^*_\mu))^{16}||^2_Y
\]

where \( (u^*_\mu, y^*_\mu, p^*_\mu) \in \mathcal{U} \times \mathcal{Y} \times \mathcal{Y} \) is the 3D-V AR truth solution (12) and \( d^*_\mu = y_d - \Pi_T y^*_\mu \) is the observable misfit. We then approximate \( \mu_{true} \) with

\[
\mu^*_{i,\lambda} \in \arg \min_{\mu \in \mathcal{C}} J^\lambda_i(\mu), \quad i \in \{1, 2, 3\}.
\]

We thus obtain a bilevel optimization problem whose inner optimization requires the solution of the 3D-V AR problem For \( i = 2 \), (43) finds the parameter \( \mu^*_{2,\lambda} \in \mathcal{C} \), for which the measurements \( (g_i(y^*_\mu))^{16} \) are closest to the actual measurement values \( v_d \). In contrast, for \( i = 3 \), the state \( y^*_{i,\lambda} \) is closest to \( y_d + T^\perp \) over all 3D-V AR state solutions. Due to our prior choice for \( \mathcal{U} \) with \( u_{true} \notin \mathcal{U} \), the measurements are not necessarily obtainable and we cannot expect \( \mu^*_{i,\lambda} \) to converge to \( \mu_{true} \) for \( \lambda \to \infty \).

We first consider the truth parameter estimation problem. The problems in (43) for \( i \in \{1, 2, 3\} \) were solved within \( \mathcal{C} \) from the starting point \( (\mu_1, \mu_2) = (1, 1) \) using the MATLAB function \texttt{fminsearch}, which uses the gradient-free simplex search algorithm described in [20]. We chose the target accuracy of \( 1e-12 \) in both the minimum value and the minimizing parameter. Each minimization took between 25 and 28 minutes. Table 1 shows the results obtained for \( \lambda = 10^j, j = 0, ..., 3 \). Generally, the parameter estimate improves as \( \lambda \) increases and the 3D-V AR solution
Table 1 The parameter estimates obtained from solving the minimizations (43) for different $\lambda$, the log-
arithmic distance to the true parameter $\mu_{\text{true}}$, the number of function evaluations, and the approximation 
quality of $u_{\text{true}}$ and $y_{\text{true}}$ for $\mu = \mu_{i}^{*,\lambda}$ through $u_{\mu}:= u_{\text{start}} + u_{\mu}^{*}$ and $y_{\mu}^{*}$.

| $i$ | $\lambda$ | $(\mu_{i}^{*,\lambda})_{1}$ | $(\mu_{i}^{*,\lambda})_{2}$ | Log. dist. | Eval. | $||u_{\text{true}} - u_{\mu,i}^{*}\|_{U}$ | $||y_{\text{true}} - y_{\mu,i}^{*}\|_{Y}$ |
|-----|-----------|----------------------------|----------------------------|-----------|-------|-----------------------------|-----------------------------|
| 1   | $10^{0}$  | 4.7545                     | 0.2173                     | 2.1878e-01| 235   | 4.5524e-01                  | 2.3916e-01                  |
|     | $10^{1}$  | 6.0571                     | 0.2666                     | 8.1061e-02| 224   | 1.9754e-01                  | 9.8988e-02                  |
|     | $10^{2}$  | 6.8618                     | 0.2952                     | 1.1115e-02| 219   | 3.5157e-02                  | 1.5374e-02                  |
|     | $10^{3}$  | 6.9896                     | 0.2997                     | 8.1349e-04| 226   | 2.1062e-02                  | 5.0379e-03                  |
| 2   | $10^{0}$  | 3.2061                     | 0.2370                     | 3.5420e-01| 228   | 4.0448e-01                  | 2.7434e-01                  |
|     | $10^{1}$  | 5.6337                     | 0.2928                     | 9.4896e-02| 225   | 1.1741e-01                  | 7.9138e-02                  |
|     | $10^{2}$  | 6.8335                     | 0.2997                     | 1.0461e-02| 243   | 2.6032e-02                  | 1.1198e-02                  |
|     | $10^{3}$  | 6.9876                     | 0.3001                     | 7.9396e-04| 238   | 2.1109e-02                  | 5.0631e-03                  |
| 3   | $10^{0}$  | 5.6402                     | 0.2513                     | 1.2134e-01| 226   | 3.4210e-01                  | 1.9846e-01                  |
|     | $10^{1}$  | 6.9380                     | 0.2980                     | 4.8032e-03| 232   | 1.0528e-01                  | 5.2142e-02                  |
|     | $10^{2}$  | 7.0038                     | 0.3002                     | 3.3277e-04| 235   | 2.5843e-02                  | 9.1612e-03                  |
|     | $10^{3}$  | 7.0047                     | 0.3002                     | 3.9092e-04| 228   | 2.1138e-02                  | 5.0527e-03                  |

favors closeness to the measurement data, with $i = 3$ providing the best parameter estimate. However, its accuracy appears to stagnate for $\lambda > 10^{2}$, presumably because of $u_{\text{true}} \notin \mathcal{U}$. The other two problems yielded comparative results for $\lambda = 1$ but did not improve as fast as the third for larger $\lambda$.

Concerning $u_{\text{true}}$ and $y_{\text{true}}$, we observe that the approximation through the 3D-
VAR solution at the parameter estimate $\mu_{i}^{*,\lambda}$ improves with increasing $\lambda$ and comes close to the best-fit errors $||\Pi_{\mathcal{U}^{\perp}}u_{\text{true}}||_{L^{2}(\Gamma_{in})} \approx 1.9877e-02$ and $||\Pi_{\mathcal{Y}^{\perp}_{\text{true}}}y||_{\mathcal{Y}} \approx 4.7576e-03$, which resulted from limiting $\mathcal{U}$ to $\mathbb{P}_{3}(\Gamma_{in})$. Note that even though the parameter estimate for $i = 3$ becomes slightly worse when changing from $\lambda = 10^{2}$ to $10^{3}$, the approximation of $u_{\text{true}}$ and $y_{\text{true}}$ still improves.

To speed up the parameter estimation, we replace the truth 3D-VAR problem with its RB approximation and consider the “reduced” cost functionals

$$J_{1,i,R}(\mu) := \frac{1}{2} ||u_{R,\mu}^{*}(\mu)||_{\mathcal{U}}^{2} + \frac{\lambda}{2} ||d_{R,\mu}^{*}(\mu)||_{\mathcal{Y}}^{2}$$

$$J_{2,i,R}(\mu) := \frac{1}{2} ||v_{\text{data}} - (g_{1}(y_{R,\mu}^{*,\mu}(\mu)))||_{\mathcal{Y}}^{2}$$

$$J_{3,i,R}(\mu) := \frac{1}{2} ||\Pi_{\mathcal{T}}y_{R,\mu}^{*}||_{\mathcal{Y}}^{2}$$

where for $\mu \in \mathcal{C}$, we have $(u_{R,\mu}^{*}, y_{R,\mu}^{*}, p_{R,\mu}^{*}) \in \mathcal{U}_{R} \times \mathcal{Y}_{R} \times \mathcal{Y}_{R}$ is the RB 3D-VAR solution and $d_{R,\mu}^{*} = y_{d} - \Pi_{\mathcal{T}}y_{R,\mu}^{*}$ is the misfit. We then take

$$\mu_{i,R}^{*,\lambda} \in \arg \min_{\mu \in \mathcal{C}} J_{i,i,R}(\mu) i \in \{1, 2, 3\},$$

$\square$ Springer
as RB parameter estimate. The distance between \( \mu_{i,R}^{*,\lambda} \) and \( \mu_{i}^{*,\lambda} \) using noise-free data, as well as the computational time for the RB parameter estimate and the corresponding speedup compared with using (42), is listed in Table 2 (panel a). The approximation of \( \mu_{i}^{*,\lambda} \) through \( \mu_{i,R}^{*,\lambda} \) was very precise with a maximal logarithmic distance of 3.3e\(^{-08}\) in the parameter domain. In each case, the minimization finished within 1.1 s, resulting in speedup factors between 1498 and 1680 compared with the truth evaluation (42) requiring 25–28 min per parameter estimate. The speedup thus justifies the initial offline cost of 7.7 min for the RB space generation, especially when considering that the RB spaces need only be generated once and can then be used \( (\cdot) \) for all combinations of \( \lambda \) and \( i \) in (45) (and also other cost functions), thereby allowing for conclusions to be drawn from the comparison of multiple parameter estimates and \( (\cdot) \) to repeat the parameter estimation repeatedly for different measurement data.

Given the precise approximation quality, we then use the RB method to obtain parameter estimates from noisy data for 100 different noise vectors. Table 2 (panel b) provides the minimum, mean, and maximum distance between the RB parameter estimates \( \mu_{i}^{*,\lambda} \) obtained from noisy and noise-free data. We observe that the noise influenced the parameter estimate independently of \( \lambda \) and \( i \). Further analysis showed that there was more deviation in the \((\mu_{i,R}^{*,\lambda})_1\) argument than in \((\mu_{i,R}^{*,\lambda})_2\) and that the

Table 2 Parameter estimation with the RB 3D-VAR method for noise-free data: distance to the truth parameter estimate \( \mu_{i,R}^{*,\lambda} \) in the logarithmic parameter plane, the computational time, and the speedup compared with (42) (panel a). Minimum, mean, and maximum distance on the logarithmic parameter plane between RB parameter estimates obtained from noisy data and the unbiased RB parameter estimate (panel b)

| \( i \) | \( \lambda \) | Noise-free data (a) | Noisy data (log. dist. to noise-free \( \mu_{i,R}^{*,\lambda} \) (b) |
|---|---|---|---|
|  |  | Dist. to \( \mu_{i,R}^{*,\lambda} \) Time [s] Speedup | Min | Mean | Max |
| 1 | \( 10^0 \) | 3.2603e\(^{-08}\) 1.1017 1498.8 | 9.5939e\(^{-04}\) 2.1314e\(^{-02}\) 8.1847e\(^{-02}\) |
| 2 | \( 10^0 \) | 4.6456e\(^{-09}\) 1.0347 1527.1 | 1.8072e\(^{-03}\) 2.4407e\(^{-02}\) 9.0962e\(^{-02}\) |
| 3 | \( 10^0 \) | 7.4937e\(^{-09}\) 1.0147 1489.7 | 2.0184e\(^{-03}\) 2.6720e\(^{-02}\) 9.5310e\(^{-02}\) |
|  | \( 10^1 \) | 6.7160e\(^{-10}\) 1.0178 1576.8 | 1.7981e\(^{-03}\) 2.6575e\(^{-02}\) 9.5909e\(^{-02}\) |
| 1 | \( 10^0 \) | 4.5371e\(^{-09}\) 0.9358 1679.5 | 7.2490e\(^{-04}\) 1.3379e\(^{-02}\) 4.4470e\(^{-02}\) |
| 2 | \( 10^0 \) | 1.4731e\(^{-08}\) 0.9443 1662.7 | 1.6227e\(^{-03}\) 2.0544e\(^{-02}\) 6.9932e\(^{-02}\) |
| 3 | \( 10^0 \) | 5.5204e\(^{-09}\) 1.0675 1608.7 | 2.7511e\(^{-03}\) 2.4773e\(^{-02}\) 8.4723e\(^{-02}\) |
|  | \( 10^1 \) | 2.9233e\(^{-09}\) 1.0386 1587.1 | 2.9854e\(^{-03}\) 2.5363e\(^{-02}\) 8.6687e\(^{-02}\) |
| 1 | \( 10^0 \) | 2.1966e\(^{-08}\) 0.9785 1591.7 | 1.0941e\(^{-03}\) 2.3011e\(^{-02}\) 8.7283e\(^{-02}\) |
| 2 | \( 10^0 \) | 1.3851e\(^{-08}\) 0.9963 1666.6 | 2.4658e\(^{-03}\) 2.6271e\(^{-02}\) 9.5628e\(^{-02}\) |
| 3 | \( 10^0 \) | 2.9369e\(^{-09}\) 0.9434 1710.8 | 1.7993e\(^{-03}\) 2.6604e\(^{-02}\) 9.5979e\(^{-02}\) |
|  | \( 10^1 \) | 2.9860e\(^{-09}\) 0.9914 1613.8 | 1.7759e\(^{-03}\) 2.6611e\(^{-02}\) 9.5978e\(^{-02}\) |
mean over the biased parameter estimates converged against the noise-free $\mu^{*,\lambda}_{i,R}$ when increasing the number of samples.

7 Conclusion

In this paper, we have proposed and analyzed a data-weak variant of the 3D-VAR method for parameterized PDEs. Through a data-informed perturbation of the model, the method generates an intermediate state between the best-knowledge model solution and the observation in the measurement space. We have reformulated the 3D-VAR method as a saddle-point problem and performed a stability analysis to reveal the relationship between the 3D-VAR method, the model, and the measurement space. In particular, we showed a necessary and sufficient condition on the design of the measurement space that results in improved stability and the mitigation of noise amplification. This condition can be used in the modeling process and in the experimental design for the selection of suitable model modifications and measurement functionals.

For an efficient computation in a parameterized real-time or many-query setting, a certified RB method was introduced. We developed a posteriori error bounds for the computationally efficient approximation of the error in the model modification, state, adjoint, and observable misfit between the truth and the RB solution. We proposed a greedy-OMP algorithm for choosing the measurement space and a construction of the RB spaces which does not require the measurement data to be known a priori. We presented numerical results for parameter and state estimation for a steady heat conduction problem with uncertain parameters and an unknown Neumann boundary condition. The numerical results confirm the validity of our approach as well as the theoretical findings.

Funding information This work was supported by the Excellence Initiative of the German federal and state governments and the German Research Foundation through Grants GSC 111 and 33849990/GRK2379 (IRTG Modern Inverse Problems).

References

1. Barrault, M., Maday, Y., Nguyen, N.C., Patera, A.T.: An ‘empirical interpolation’ method: application to efficient reduced-basis discretization of partial differential equations. Comptes Rendus Mathematique 339(9), 667–672 (2004)
2. Bennett, A.: Array design by inverse methods. Prog. Oceanogr. 15(2), 129–156 (1985)
3. Bennett, A.F.: Inverse Methods in Physical Oceanography. Cambridge University Press (1992)
4. Bennett, A.F.: Inverse Modeling of the Ocean and Atmosphere. Cambridge University Press (2002)
5. Binev, P., Cohen, A., Dahmen, W., DeVore, R., Petrova, G., Wojtaszczyk, P.: Data assimilation in reduced modeling. SIAM/ASA J. Uncert. Quantif. 5(1), 1–29 (2017)
6. Binev, P., Cohen, A., Mula, O., Nichols, J.: Greedy algorithms for optimal measurements selection in state estimation using reduced models. Preprint (2017)
7. Boffi, D., Brezzi, F., Demkowicz, L.F., Durán, R.G., Falk, R.S., Fortin, M.: Mixed Finite Elements, Compatibility Conditions and Applications. Springer (2008)
8. Brezzi, F.: On the existence, uniqueness and approximation of saddle-point problems arising from lagrangian multipliers. Revue française d’automatique, informatique, recherche opérationnelle 8(2), 129–151 (1974)
9. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms. MIT Press (2009)
10. Eftang, J.L., Huynh, D.B.P., Knezevic, D., Patera, A.T.: A two-step certified reduced basis method. J. Sci. Comput. 51, 28–58 (2012)
11. Gerner, A., Veroy, K.: Certified reduced basis methods for parametrized saddle point problems. SIAM J. Sci. Comput. 34(5), A2812–A2836 (2012)
12. Gerner, A.L.: Certified reduced basis methods for parametrized saddle point problems. RWTH Aachen University, Ph.D. thesis (2012)
13. Gerner, A.L., Veroy, K.: Reduced Basis A Posteriori Error Bounds for Symmetric Parametrized Saddle Point Problems ArXiv e-prints (2012)
14. Grepl, M.A., Maday, Y., Nguyen, N.C., Patera, A.T.: Efficient reduced-basis treatment of non-affine and nonlinear partial differential equations. ESAIM: Mathematical Modelling and Numerical Analysis 41(3), 575–605 (2007)
15. Grotsch, C.W.: Generalized Inverses of Linear Operators: Representation and Approximation. Marcel Dekker, inc New York and Basel (1977)
16. Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, S.: Optimization with PDE Constraints Springer Science + Business Media B.V (2009)
17. Huynh, D., Rozza, G., Sen, S., Patera, A.: A successive constraint linear optimization method for lower bounds of parametric coercivity and inf–sup stability constants. Comptes Rendus Mathematique 345(8), 473–478 (2007)
18. Kärcher, M., Boyaval, S., Grepl, M.A., Veroy, K.: Reduced basis approximation and a posteriori error bounds for 4D-Var data assimilation. Optim. Eng. 19, 663–695 (2018)
19. Kärcher, M., Tokoutsis, Z., Grepl, M.A., Veroy, K.: Certified reduced basis methods for parametrized elliptic optimal control problems with distributed controls. J. Sci. Comput. 75(1), 276–307 (2018)
20. Lagarias, J.C., Reeds, J.A., Wright, M.H., Wright, R.E.: Convergence properties of the nelder-mead simplex method in low dimensions. SIAM J. Optim. 9(1), 112–147 (1998). 10.1137/S1052623496303470
21. Law, K., Stuart, A., Zygalakis, K.: Data Assimilation Springer (2015)
22. Le Dimet, F.X., Talagrand, O.: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. Tellus A 38A(2), 97–110 (1986)
23. Lorenz, A.C.: A global three-dimensional multivariate statistical interpolation scheme. Mon. Weather. Rev. 109(4), 701–721 (1981)
24. Lorenz, A.C.: Analysis methods for numerical weather prediction. Q. J. Roy. Meteorol. Soc. 112(474), 1177–1194 (1986)
25. Maday, Y., Mula, O.: A Generalized Empirical Interpolation Method: Application of Reduced Basis Techniques to Data Assimilation, pp. 221–235, Springer, Milan, Milano (2013)
26. Maday, Y., Mula, O., Patera, A.T., Yano, M.: The generalized empirical interpolation method: stability theory on Hilbert spaces with an application to the Stokes equation. Comput. Methods Appl. Mech. Eng. 287, 310–334 (2015)
27. Maday, Y., Patera, A.T., Penn, J.D., Yano, M.: A parameterized-background data-weak approach to variational data assimilation: formulation, analysis, and application to acoustics. Int. J. Numer. Methods Eng. 102(5), 933–965 (2014)
28. Maday, Y., Patera, A.T., Penn, J.D., Yano, M.: PBDW state estimation: noisy observations; configuration-adaptive background spaces; physical interpretations. ESAIM: Proceedings 50, 144–168 (2015)
29. Maday, Y., Taddei, T.: Adaptive PBDW approach to state estimation: noisy observations; user-defined update spaces ArXiv e-prints (2017)
30. Nadal, E., Chinesta, F., Dez, P., Fuenmayor, F., Denia, F.: Real time parameter identification and solution reconstruction from experimental data using the proper generalized decomposition. Comput. Methods Appl. Mech. Eng. 296, 113–128 (2015)
31. Nellesen, N.: A certified reduced basis method for parameterized 3D-VAR data assimilation. Master’s thesis, RWTH Aachen (2018)
32. Reich, S., Cotter, C.: Probabilistic Forecasting and Bayesian Data Assimilation Cambridge University Press (2015)
33. De los Reyes, J.: Numerical PDE-Constrained Optimization Springer Briefs in Optimization (2015)
34. Rozza, G., Huynh, D.B.P., Patera, A.T.: Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. Archives of Computational Methods in Engineering 15(3), 1 (2007)
35. Taddei, T., Patera, A.: A localization strategy for data assimilation; application to state estimation and parameter estimation. SIAM J. Sci. Comput. 40(2), B611–B636 (2018)
36. collab=Taddei, T.: An adaptive parametrized-background data-weak approach to variational data assimilation. ESAIM: M2AN 51(5), 1827–1858 (2017)
37. Tikhonov, A.N., Goncharsky, A., Stepanov, V.V., Yagola, A.G.: Numerical Methods for the Solution of Ill-Posed Problems Springer Science+Business Media Dordrecht (1995)
38. Trémolet, Y.: Accounting for an imperfect model in 4D-Var. Q. J. Roy. Meteorol. Soc. 132(621), 2483–2504 (2006)
39. Vermeulen, P.T.M., Heemink, A.W.: Model-reduced variational data assimilation. Mon. Weather. Rev. 134(10), 2888–2899 (2006)
40. Willcox, K.: Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition. Computers & Fluids 35(2), 208–226 (2006)
41. Yano, M., Penn, J.D., Patera, A.T.: A model-data weak formulation for simultaneous estimation of state and model bias. Comptes Rendus Mathematique 351(23), 937–941 (2013)

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Nicole Aretz-Nellesen1 · Martin A. Grepl2 · Karen Veroy1,3

Nicole Aretz-Nellesen
nellesen@aices.rwth-aachen.de

Karen Veroy
veroy@aices.rwth-aachen.de

1 Aachen Institute for Advanced Study in Computational Engineering Science (AICES), RWTH Aachen University, Schinkelstraße 2, 52062 Aachen, Germany
2 Numerical Mathematics (IGPM), RWTH Aachen University, Templergraben 55, 52056 Aachen, Germany
3 Faculty of Civil Engineering, RWTH Aachen University, Schinkelstraße 2, 52062 Aachen, Germany