SIAM J. NUMER. ANAL.
Vol. 58, No. 6, pp. 3355–3381
© 2020 Society for Industrial and Applied Mathematics

OPTIMAL REDUCED MODEL ALGORITHMS FOR DATA-BASED STATE ESTIMATION*

ALBERT COHEN†, WOLFGANG DAHMEN‡, RONALD DEVORE§, JALAL FADILI¶, OLGA MULA‖, AND JAMES NICHOLS#

Abstract. Reduced model spaces, such as reduced bases and polynomial chaos, are linear spaces V_n of finite dimension n which are designed for the efficient approximation of certain families of parametrized PDEs in a Hilbert space V. The manifold $\mathcal{M}$ that gathers the solutions of the PDE for all admissible parameter values is globally approximated by the space V_n with some controlled accuracy ε_n, which is typically much smaller than when using standard approximation spaces of the same dimension such as finite elements. Reduced model spaces have also been proposed in [Y. Maday et al., *Internat. J. Numer. Methods Ergrg.*, 102 (2015), pp. 933–965] as a vehicle to design a simple linear recovery algorithm of the state $u \in \mathcal{M}$ corresponding to a particular solution instance when the values of parameters are unknown but a set of data is given by m linear measurements of the state. The measurements are of the form $\ell_j(u)$, $j = 1, \dots, m$, where the ℓ_j are linear functionals on V. The analysis of this approach in [P. Binev et al., *SIAM/ASA J. Uncertain. Quantif.*, 5 (2017), pp. 1–29] shows that the recovery error is bounded by $\mu_n \varepsilon_n$, where $\mu_n = \mu(V_n, W)$ is the inverse of an inf-sup constant that describe the angle between V_n and the space W spanned by the Riesz representers of $(\ell_1, \dots, \ell_m)$. A reduced model space which is efficient for approximation might thus be ineffective for recovery if μ_n is large or infinite. In this paper, we discuss the existence and effective construction of an optimal reduced model space for this recovery method. We extend our search to affine spaces which are better adapted than linear spaces for various purposes. Our basic observation is that this problem is equivalent to the search of an optimal affine algorithm for the recovery of $\mathcal{M}$ in the worst case error sense. This allows us to peform our search by a convex optimization procedure. Numerical tests illustrate that the reduced model spaces constructed from our approach perform better than the classical reduced basis spaces.

Key words. reduced models, optimal recovery, sensing, parametrized PDEs, convex optimization

AMS subject classifications. 65M32, 65M15, 68U20, 65D99

DOI. 10.1137/19M1255185

1. Introduction.

1.1. Background and context.

State estimation refers to the general problem of approximately recovering the true state of a physical system of interest from incomplete data. This task is ubiquitous in applied sciences and engineering. One can draw a distinction between two different application scenarios:

*Received by the editors April 8, 2019; accepted for publication (in revised form) July 13, 2020; published electronically November 24, 2020.

https://doi.org/10.1137/19M1255185

Funding: The work of the third author was supported by the Institut Universitaire de France, the ERC Adv grant BREAD, the EMERGENCES grant project "Models and Measures" from the Paris City Council, the National Science Foundation grants DMS 15-21067, DMS 18-17603, DMS 172097 and the Office of Naval Research grants N00014-17-1-2908, N00014-16-1-2706. The work of the first, second, third, fifth and sixth authors were also supported as visitors by the Isaac Newton Institute at Cambridge University.

†Sorbonne Université, Paris, 75005, France (albert.cohen@sorbonne-universite.fr).

‡Department of Mathematics, University of South Carolina, Columbia, SC 29208 USA (wolfgang.anton.dahmen@googlemail.com).

§Texas A&M University, College Station, TX 77843 USA (rdevore@math.tamu.edu).

¶Normandie Université, ENSICAEN, CNRS, Caen, 14050, France (Jalal.Fadili@ensicaen.fr).

‖CEREMADE, CNRS, UMR 7534, Université Paris-Dauphine, PLS University, 75016, Paris, France; Inria, Sorbonne Université and CNRS, Paris, France (mula@ceremade.dauphine.fr).

#Australian National University, ACT 0200, Australia (james.nichols@anu.edu.au).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

(i) The physical properties of the states, sometimes referred to as background information, are approximately modeled by a nonlinear dynamical system which by itself is neither sufficiently accurate nor stable to warrant reliable predictions. This is typically the case in weather prediction, climatology, or generally in atmospheric research. One therefore utilizes observational data to *correct* the model-based predictions, ideally in real time. Such correction mechanisms are often based on statistical hypotheses such as Gaussianity of error distributions. One important approach is *ensemble Kalman filtering* which can be viewed as a recursive Bayesian estimation based on Monte Carlo approximations to the first and second moments of the error distributions [16, 18]. A second class of methods are so-called variational data assimilation schemes like 3 dimensional (3D)-VAR or 4 dimensional (4D)-VAR [19]. The state predicted by the model is then corrected by minimizing a quadratic cost functional involving inverse covariance matrices for the background model error and observation error. In this first scenario, the error bounds between the exact and estimated state are typically expressed in an average sense, based on the accepted simplified statistical model assumptions.

(ii) The physical states of interest are reliably described in terms of a *parameter dependent* family of PDEs which for each parameter instance can be computed within a desired target accuracy. The states are therefore elements of the associated *solution manifold* that consists of all solutions to the PDE as parameters vary. The task is then to estimate a state in (or near to) the solution manifold from only a finite number of measurements generated through a *fixed* number of sensors. A classical example is to estimate a pressure field of a porous media flow from a finite number of pressure head measurements. The parametric model then could arise from a Karhunen–Loève expansion of a random field of permeability coefficients in Darcy's law, and may thus involve a large or even infinite number of parameters. Problems of this type have been investigated over the past decade in the context of uncertainty quantification. Again, a prominent approach is Bayesian inversion where prior information is given in terms of a probability distribution for the parameter, inducing a probability distribution for the state [25]. The objective is then to approximate the posterior probability distribution of states given the data. High dimensionality renders such methods computationally expensive. Alternatively, state estimation can be formulated as a constrained optimization problem. For instance, one could minimize the deviation of state measurements over the solution manifold asking for probabilistic or deterministic error bounds. In practice, one typically chooses first a sufficiently fine discretization of the *high fidelity* continuum model which then gives rise to a large scale (discrete) constrained nonconvex optimization problem that needs to be solved for each instance of data. Ill-posedness of the inversion task necessitates adding regularization terms which introduce a further ambiguous bias. Reduced models are used to alleviate the possibly prohibitive cost of the numerous forward simulations that are needed in the descent method. A central issue is then to judiciously switch between the high fidelty model, given in terms of the fine scale discretization, and the low fidelity reduced model; see [27].

In this article, we consider scenario (ii) but pursue a different approach taking up on recent work in [3, 20]. Although it can be formulated without any reference to a statistical model, it has conceptual similarities with the 3D- and 4D-Var variational approach invoked for scenario (i); see [17] and [26] for such connections. In contrast

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

to Bayesian inversion, this approach yields deterministic error bounds expressed in a worst case sense over the solution manifold, which is the primary interest in this paper. Specifically, we follow [3] and formulate state estimation as an *optimal recovery problem*; see, e.g., [21]. This allows us to formulate optimality benchmarks that steer our development of recovery algorithms.

1.2. Mathematical formulation of the state estimation problem. The *sensing* or *recovery* problem studied in this paper is formulated in a Hilbert space V equiped with some norm $\|\cdot\|$ and inner product $\langle\cdot,\cdot\rangle$: we want to recover an approximation to an unknown function $u \in V$ from data given by m linear measurements

$$\ell_i(u), \quad i = 1, \dots, m, \tag{1.1}$$

where the ℓ_i are m linearly independent bounded linear functionals over V. This problem appears in many different settings. The particular one that motivates our work is the case where $u = u(y)$ represents the *state* of a physical system described as a solution to a parametric PDE

$$\mathcal{P}(u, y) = 0 \tag{1.2}$$

for some unknown finite or infinite dimensional parameter vector $y = (y_j)_{j\geq 1}$ picked from some admissible set Y. The ℓ_i are a mathematical model for sensors that capture some partial information on the unknown solution $u(y) \in V$.

Denoting by $\omega_i \in V$ the Riesz representers of the ℓ_i, such that $\ell_i(v) = \langle \omega_i, v\rangle$ for all $v \in V$, and defining

$$W := \operatorname{span}\{\omega_1, \dots, \omega_m\}, \tag{1.3}$$

the measurements are equivalently represented by

$$w = P_W u, \tag{1.4}$$

where P_W is the orthogonal projection from V onto W. A *recovery algorithm* is a computable map

$$A : W \to V \tag{1.5}$$

and the approximation to u obtained by this algorithm is

$$u^* = A(w) = A(P_W u). \tag{1.6}$$

The construction of A should be based on the available prior information that describes the properties of the unknown u, and the evaluation of its performance needs to be defined in some precise sense. Two distinct approaches are usually followed:

- In the *deterministic setting*, the sole prior information is that u belongs to the set

  $$\mathcal{M} := \{u(y) \ : \ y \in Y\} \tag{1.7}$$

  of all possible solutions. The set $\mathcal{M}$ is sometimes called the *solution manifold*. The performance of an algorithm A over the class $\mathcal{M}$ is usually measured by the "worst case" reconstruction error

  $$E_{\mathrm{wc}}(A, \mathcal{M}) = \sup\{\|u - A(P_W u)\| \ : \ u \in \mathcal{M}\}. \tag{1.8}$$

  The problem of finding an algorithm that minimizes $E_{\mathrm{wc}}(A)$ is called *optimal recovery*. It has been extensively studied for convex sets $\mathcal{M}$ that are balls of smoothness classes [5, 21, 22], which is not the case for (1.7).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

- In the *stochastic setting*, the prior information on u is described by a probability distribution p on V, which is supported on $\mathcal{M}$, typically induced by a probability distribution on Y that is assumed to be known. It is then natural to measure the performance of an algorithm in an averaged sense, for example, through the mean square error

$$E_{\rm ms}(A,p) = \mathbb{E}(\|u - A(P_W u)\|^2) = \int_V \|u - A(P_W u)\|^2 dp(u). \tag{1.9}$$

  This stochastic setting is the starting point for *Bayesian estimation* methods [13]. Let us observe that for any algorithm A one has $E_{\rm ms}(A,p) \le E_{\rm wc}(A,\mathcal{M})^2$.

1.3. Optimal algorithms. The present paper concentrates on the deterministic setting according to the above distinction, although some remarks will be given on the analogies with the stochastic setting. In this setting, the benchmark for the performance of recovery algorithms is given by

$$E^*_{\rm wc}(\mathcal{M}) = \inf_A E_{\rm wc}(A,\mathcal{M}),$$

where the infimum is taken over all possible maps A.

There is a simple mathematical description of an optimal map that meets this benchmark. For any bounded set $S \subset V$ we define its *Chebyshev ball* as the smallest closed ball that contains S. The *Chebyshev ball radius and center* denoted by ${\rm rad}(S)$ and ${\rm cen}(S)$ are the radius and center of this ball. Since the information that we have on u is that it belongs to the set

$$\mathcal{M}_w := \mathcal{M} \cap V_w, \quad V_w := \{v \in V \,:\, P_W v = w\} = w + W^\perp, \tag{1.10}$$

where $W^\perp$ is the orthogonal complement of W in V, it follows that an optimal reconstruction map $A^*_{\rm wc}$ for the worst case error is given by

$$A^*_{\rm wc}(w) = {\rm cen}(\mathcal{M}_w), \tag{1.11}$$

because the Chebyshev ball center of $\mathcal{M}_w$ minimizes the quantity $\sup\{\|u - v\| \,:\, u \in \mathcal{M}_w\}$ among all $v \in V$. The worst case error is therefore given by

$$E^*_{\rm wc}(\mathcal{M}) = E_{\rm wc}(A^*_{\rm wc},\mathcal{M}) = \sup\{{\rm rad}(\mathcal{M}_w) \,:\, w \in P_W(\mathcal{M})\}. \tag{1.12}$$

Note that the map $A^*_{\rm wc}$ is also optimal among all algorithms for each $\mathcal{M}_w$, $w \in P_W(\mathcal{M})$, since

$$E_{\rm wc}(A^*_{\rm wc},\mathcal{M}_w) = \min_A E_{\rm wc}(A,\mathcal{M}_w) = {\rm rad}(\mathcal{M}_w), \quad w \in P_W(\mathcal{M}). \tag{1.13}$$

However, there may exist other maps A such that $E_{\rm wc}(A,\mathcal{M}) = E^*_{\rm wc}(\mathcal{M})$, since we also supremize over $w \in P_W(\mathcal{M})$.

1.4. Linear and affine algorithms based on reduced models. In practice the above map $A^*_{\rm wc}$ cannot be easily constructed due to the fact that the solution manifold $\mathcal{M}$ is a high dimensional and geometrically complex object. One is therefore interested in designing "suboptimal yet good" recovery algorithms and analyzing their performance.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

One vehicle for constructing linear recovery mappings A is to use *reduced modeling*. Generally speaking, reduced models consist of linear spaces $(V_n)_{n\geq 0}$ with increasing dimension $\dim(V_n) = n$ which uniformly approximate the solution manifold in the sense that

$$\operatorname{dist}(\mathcal{M}, V_n) := \max_{u\in\mathcal{M}} \min_{v\in V_n} \|u - v\| \leq \varepsilon_n, \tag{1.14}$$

where

$$\varepsilon_0 \geq \varepsilon_1 \geq \cdots \geq \varepsilon_n \geq \cdots \geq 0 \tag{1.15}$$

are known tolerances. Instances of reduced models for parametrized families of PDEs with provable accuracy are provided by polynomial approximations in the y variable [10, 11] or reduced bases [7, 24, 23]. The construction of a reduced model is typically done offline, possibly using a large training set of instances of $u \in \mathcal{M}$ called *snapshots*. The offline stage potentially has a high computational cost. Once this is done, the online cost of recovering $u^* = A(w)$ from any data w using this reduced model should, in contrast, be moderate.

In [20], a simple reduced-model-based recovery algorithm was proposed in terms of the map

$$A_n(w) := \operatorname{argmin}\{\operatorname{dist}(v, V_n) \ : \ v \in V_w\}, \tag{1.16}$$

which is well-defined provided that $V_n \cap W^\perp = \{0\}$. It turns out that A_n is a linear mapping and so these algorithms are linear. This approach is called the parametrized-background data-weak method, however, we follow the terminology introduced in [3], refering to an algorithm of the form A_n as a *one-space algorithm*. In the latter, it was shown that A_n has a simple interpretation in terms of the cylinder

$$\mathcal{K}_n := \{v \in V \ : \ \operatorname{dist}(v, V_n) \leq \varepsilon_n\} \tag{1.17}$$

that contains the solution manifold $\mathcal{M}$. Namely, the algorithm A_n is also given by

$$A_n(w) = \operatorname{cen}(\mathcal{K}_{n,w}), \quad \mathcal{K}_{n,w} := \mathcal{K}_n \cap V_w, \tag{1.18}$$

and the map is shown to be optimal when $\mathcal{M}$ is replaced by the simpler containment set $\mathcal{K}_n$, that is,

$$A_n = \operatorname*{argmin}_{A:W\to V} E_{\mathrm{wc}}(A, \mathcal{K}_n).$$

The substantial advantage of this approach is that, in contrast to A^*_{wc}, the map A_n can be easily computed by solving simple least-squares minimization problems which amount to finite linear systems. In turn A_n is a linear map from W to V. This map depends on V_n and W, but not on ε_n in view of (1.16). We refer to A_n as the one-space algorithm based on the space V_n.

This algorithm satisfies the performance bound

$$\|u - A_n(P_W u)\| \leq \mu_n \operatorname{dist}(u, V_n \oplus (V_n^\perp \cap W)) \leq \mu_n \operatorname{dist}(u, V_n) \leq \mu_n \varepsilon_n, \tag{1.19}$$

where the last inequality holds when $u \in \mathcal{M}$. Here

$$\mu_n = \mu(V_n, W) := \max_{v\in V_n} \frac{\|v\|}{\|P_W v\|} \tag{1.20}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

is the inverse of the inf-sup constant $\beta_n := \min_{v\in V_n} \max_{w\in W} \frac{\langle v,w\rangle}{\|v\|\,\|w\|}$ which describes the angle between V_n and W. In particular $\mu_n = \infty$ in the event where $V_n \cap W^\perp$ is nontrivial.

An important observation is that the one-space algorithm (1.16) has a simple extension to the setting where V_n is an affine space rather than a linear space, namely, when

$$V_n = \overline{u} + \widetilde{V}_n \tag{1.21}$$

with $\widetilde{V}_n$ a linear space of dimension n and $\overline{u}$ a given offset that is known to us.

At a first sight, affine spaces do not bring any significant improvement in terms of approximating the solution manifold, due to the following elementary observation: if $\mathcal{M}$ is approximated with accuracy ε by an n-dimensional affine space V_n given by (1.21), it is also approximated with accuracy $\widetilde{\varepsilon} \le \varepsilon$ by the $(n+1)$-dimensional linear space

$$\widetilde{V}_{n+1} := V_n \oplus \mathbb{R}\overline{u}. \tag{1.22}$$

However, the choice of an affine reduced model may significantly improve the performance of the one-space algorithm in the case where the parametric solution $u(y)$ is a "small perturbation" of a nominal solution $\overline{u} = u(\overline{y})$ for some $\overline{y} \in Y$ in the sense that

$$\operatorname{diam}(\mathcal{M}) \ll \|\bar{u}\|. \tag{1.23}$$

Indeed, suppose in addition that $\overline{u}$ is badly aligned with respect to the measurement space W in the sense that

$$\|P_W \overline{u}\| \ll \|\bar{u}\|. \tag{1.24}$$

In such a case, any linear space V_n that is well tailored to approximating the solution manifold (for example, a reduced basis space) will contain a direction close to that of $\overline{u}$ and thus, we will have that $\mu_n \gg 1$, rendering the reconstruction by the linear one-space method much less accurate than the approximation error by V_n. The use of the affine mapping (1.21) has the advantage of elimitating the bad direction $\overline{u}$ since μ_n will now be computed with respect to the linear part $\widetilde{V}_n$.

A further perspective, currently under investigation, is to agglomerate local affine models in order to generate a *nonlinear* reduced model. This can be executed, for example, by decomposing the parameter domain Y into K subdomains Y_k and using different affine reduced models for approximating the resulting subsets $\mathcal{M}_k = u(Y_k)$, see, for instance, [6, 15].

1.5. Objective and outline. The standard constructions of reduced models are targeted at making the spaces V_n as efficient as possible for approximating $\mathcal{M}$, that is, making ε_n as small as possible for each given n. For example, for the reduced basis spaces, it is known [2, 12] that a certain greedy selection of snapshots generates spaces V_n such that $\operatorname{dist}(\mathcal{M}, V_n)$ decays at the same rate (polynomial or exponential) as the Kolmogorov n-width

$$\delta_n(\mathcal{M}) := \inf\{\operatorname{dist}(\mathcal{M}, E) \ : \ \dim(E) = n\}. \tag{1.25}$$

However these constructions do not ensure the control of μ_n and therefore these reduced spaces may be much less efficient when using the one-space algorithm for the recovery problem.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

In view of the above observations, the objective of this paper is to discuss the construction of reduced models (both linear and affine) that are better targeted towards the recovery task. In other words, we want to build the spaces V_n to make the recovery algorithm A_n as efficient as possible, given the measurement space W. Note that a different problem is, given $\mathcal{M}$, to optimize the choice of the measurement functionals ℓ_i picked from some admissible dictionary, which amounts to optimizing the space W, as discussed for example in [4]. Here, we consider our measurement system to be imposed on us, and therefore W to be fixed once and for all.

The rest of our paper is organized as follows. In section 2, we detail the affine map A_n associated with V_n, that can be computed in a similar way as in the linear case. Conversely, we show that any affine recovery map may be interpreted as a one-space algorithm for a certain affine reduced model V_n. For a general set $\mathcal{M}$, the existence and construction of an optimal affine recovery map A^*_{wca} for the worst case error is therefore equivalent to the existence and construction of an optimal reduced space for the recovery problem. We then draw a short comparison with the stochastic setting in which the optimal affine map A^*_{msa} for the mean square error (1.9) is derived explicitely from the second order statistics of u.

In section 3, we compute an approximation of A^*_{wca} by convex optimization, based on a training set of snapshots. Two algorithms are considered: subgradient descent and primal-dual proximal splitting. Our numerical results illustrate the superiority of the latter for this problem. The optimal affine map A^*_{wca} significantly outperforms the one-space algorithm A_{n^*} when standard reduced basis spaces V_n are used and an optimal value n^* is selected using the training set. It also outperforms the affine map A^*_{msa} computed from second order statistics of the training set. All three maps significanly outperform the minimal V-norm recovery given by $A(w) = w = P_W u$.

2. Affine one-space recovery. In this section, we show that any linear recovery algorithm is given by a one-space algorithm and that a similar result holds for any affine algorithm. We then go on to describe the optimal one-space algorithms by exploiting this fact.

2.1. The one-space algorithm. We begin by discussing in more detail the one-space algorithm for a linear space V_n of dimension $n \leq m$. As shown in [3], the map A_n associated with V_n has a simple expression after a proper choice of favorable bases has been made for W and V_n through an SVD applied to the cross-Gramian of an initial pair of orthonormal bases. The resulting *favorable bases* $\{\psi_1, \ldots, \psi_m\}$ for W and $\{\varphi_1, \ldots, \varphi_n\}$ for V_n satisfy the equations

$$\langle \psi_i, \varphi_j \rangle = s_i \delta_{i,j}, \tag{2.1}$$

where

$$1 \geq s_1 \geq s_2 \geq \cdots \geq s_n > 0 \tag{2.2}$$

are the singular values of the cross-Gramian. Then, if w is in W, we can write $w = \sum_{j=1}^m w_j \psi_j$ in the favorable basis, and find that

$$A_n(w) = \sum_{j=1}^n s_j^{-1} w_j \varphi_j + \sum_{j=n+1}^m w_j \psi_j. \tag{2.3}$$

Let us observe that the functions ψ_j in the second sum span the space $V_n^\perp \cap W$ while the first sum is the solution of the least-squares problem $\min_{v \in V_n} \|w - P_W v\|$ corrected by the second sum so as to fit the data.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Now consider any linear recovery algorithm $A : W \to V$. Since we are given the measurement observation w, any algorithm A which is a candidate for optimality must satisfy $P_W(A(w)) = w$ (otherwise the reconstruction error would not be minimized). Thus A should have the form

$$A(w) = w + B(w), \tag{2.4}$$

where $B : W \to W^\perp$ with $W^\perp$ the orthogonal complement of W in V. Note that in functional analysis the mappings A of the form (2.4) are called liftings.

Therefore, in going further in this paper, we always require that A has the form (2.4) and concentrate on the construction of good linear maps B. Our next observation is that any algorithm A of this form can always be interpreted as a one-space algorithm A_n for a certain space V_n with $n \le m$.

PROPOSITION 2.1. *Let A be any linear map of the form* (2.4)*. Then, there exists a space V_n of dimension $n \le m$ such that A coincides with the one-space algorithm* (2.3) *for V_n.*

Proof. By considering the SVD of the linear transform B, there exists an orthonormal basis $\{\psi_1, \dots, \psi_m\}$ of W and an orthonormal system $\{\omega_1, \dots, \omega_m\}$ in $W^\perp$ such that, with $w = \sum_{j=1}^m w_j \psi_j$,

$$Bw = \sum_{j=1}^m \alpha_j w_j \omega_j, \quad w \in W, \tag{2.5}$$

for some numbers $\alpha_1 \ge \alpha_2 \ge \cdots \ge \alpha_m \ge 0$. Defining the functions

$$\varphi_j = s_j(\psi_j + \alpha_j \omega_j), \quad s_j = (1 + \alpha_j^2)^{-1/2}, \tag{2.6}$$

and defining V_n as the span of those φ_j for which $\alpha_j \neq 0$, we recover the exact form (2.3) of the one-space algorithm expressed in favorable bases. ∎

These results can be readily extended to the case where V_n is an affine space given by (1.21) for some given n-dimensional linear space $\widetilde{V}_n$ and offset $\overline{u}$. In what follows, we systematically use the notation

$$\widetilde{u} = u - \overline{u} \tag{2.7}$$

for the recentered state, and likewise $\widetilde{w} = w - \overline{w}$ with $\overline{w} = P_W \overline{u}$ for the recentered observation. The one-space algorithm associated with V_n has the form

$$A_n(w) = \overline{u} + \widetilde{A}_n(\widetilde{w}), \tag{2.8}$$

where $\widetilde{A}_n$ is the one-space linear algorithm associated with $\widetilde{V}_n$.

Performances bounds similar to those of the linear case are derived in the same way as in [3]: the reconstruction satisfies

$$\|u - A_n(P_W u)\| \le \mu_n \mathrm{dist}(u, \overline{u} + \widetilde{V}_n \oplus (\widetilde{V}_n^\perp \cap W)) \le \mu_n \mathrm{dist}(u, V_n), \tag{2.9}$$

where

$$\mu_n = \mu(\widetilde{V}_n, W) = \max_{v \in \widetilde{V}_n} \frac{\|v\|}{\|P_W v\|} = s_n^{-1} < \infty. \tag{2.10}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

The map A_n is optimal for the cylinders of the form

$$\mathcal{K}_n = \{u \in V \ : \ \mathrm{dist}(u, V_n) \leq \varepsilon_n\} \tag{2.11}$$

since it coincides with the Chebyshev ball center of $\mathcal{K}_{n,w} = \mathcal{K}_n \cap V_w$. In particular, one has

$$E^*_{\mathrm{wc}}(\mathcal{K}_n) = E_{\mathrm{wc}}(A_n, \mathcal{K}_n) = \mu_n \varepsilon_n. \tag{2.12}$$

For a solution manifold $\mathcal{M}$ contained in $\mathcal{K}_n$, one has

$$E^*_{\mathrm{wc}}(\mathcal{M}) \leq E_{\mathrm{wc}}(A_n, \mathcal{M}) \leq \mu_n \mathrm{dist}(\mathcal{M}, \widetilde{V}_n \oplus (\widetilde{V}_n^{\perp} \cap W)) \leq \mu_n \mathrm{dist}(\mathcal{M}, V_n) \leq \mu_n \varepsilon_n, \tag{2.13}$$

and these inequalities are generally strict.

In view of (2.8) the map A_n is affine. A general affine recovery map takes the form

$$A(w) = w + Bw + c, \tag{2.14}$$

where $B : W \to W^{\perp}$ is linear and $c = A(0) \in W^{\perp}$. The following result is a direct consequence of Proposition 2.1.

COROLLARY 2.2. *Let A be an affine map of the form* (2.14). *Then, there exists an affine space $V_n = \overline{u} + \widetilde{V}_n$ such that A coincides with the one-space algorithm* (2.8).

2.2. The best affine map. In view of this result, the search for an affine reduced model V_n that is best tailored to the recovery problem is equivalent to the search of an optimal affine map. Our next result is that such a map always exist when $\mathcal{M}$ is a bounded set.

THEOREM 2.3. *Let $\mathcal{M}$ be a bounded set. Then there exists a map A^*_{wca} that minimizes $E_{\mathrm{wc}}(A, \mathcal{M})$ among all affine maps A.*

Proof. We consider any affine map A of the form (2.14), so that the error is given by

$$E_{\mathrm{wc}}(A, \mathcal{M}) = \sup\{u \in \mathcal{M} \ : \ \|P_{W^{\perp}} u - c - B P_W u\|\} =: F(c, B). \tag{2.15}$$

We begin by remarking that for each $(c, B) \in W^{\perp} \times \mathcal{L}(W, W^{\perp})$, the map $u \mapsto \|P_{W^{\perp}} u - c - B P_W u\|$ is uniformly bounded on the bounded set $\mathcal{M}$. Its supremum $F(c, B)$ is thus a finite positive number, which we may write as

$$F(c, B) = \sup_{u \in \mathcal{M}} F_u(c, B), \tag{2.16}$$

where $F_u(c, B) = \|P_{W^{\perp}} u - c - B P_W u\|$. Each F_u is convex and satisfies the Lipschitz bound

$$|F_u(c, B) - F_u(c', B')| \leq \|c - c'\| + M \|B - B'\|_S \tag{2.17}$$

where $\|B\|_S$

$$\|B\|_S = \max\{\|Bv\| \ : \ v \in W, \ \|v\| = 1\} \tag{2.18}$$

denotes the spectral norm and $M := \sup\{\|P_W u\| \ : \ u \in \mathcal{M}\} < \infty$. This implies that the function F is convex and satisfies the same Lipschitz bound.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

We note that the linear maps of $\mathcal{L}(W, W^\perp)$ are of rank at most m and therefore, given any orthonormal basis $(e_1, \dots, e_m)$ of W, we can equip $\mathcal{L}(W, W^\perp)$ with the Hilbert–Schmidt norm

$$\|B\|_{HS} := \Big(\sum_{i=1}^m \|Be_i\|^2\Big)^{1/2}, \tag{2.19}$$

which is equivalent to the spectral norm since

$$\|B\|_S \le \|B\|_{HS} \le \sqrt{m}\|B\|_S, \quad B \in \mathcal{L}(W, W^\perp). \tag{2.20}$$

In particular F is continuous with respect to the Hilbertian norm

$$\|(c, B)\|_H := \Big(\|c\|^2 + \sum_{i=1}^m \|Be_i\|^2\Big)^{1/2}. \tag{2.21}$$

The function F may not be infinite at infinity: this happens if there exists a nontrivial pair (c, B) such that

$$c + BP_W u = 0, \quad u \in \mathcal{M}.$$

In order to fix this problem, we define the subspace

$$S_0 := \Big\{(c, B) \in W^\perp \times \mathcal{L}(W, W^\perp) \ : \ c + BP_W u = 0, \ u \in \mathcal{M}\Big\}. \tag{2.22}$$

and we denote by S_1 its orthogonal complement in $W^\perp \times \mathcal{L}(W, W^\perp)$ for the inner product associated with the above Hilbertian norm $\|\cdot\|_H$. The function F is constant in the direction of S_0 and therefore we are left to prove the existence of the minimum of F on S_1. For any $(c, B) \in S_1$, there exists $u \in \mathcal{M}$ such that $c + BP_W u \neq 0$. This implies that

$$\lim_{|t|\to+\infty} \|P_{W^\perp} u - tc - tBP_W u\| = +\infty \tag{2.23}$$

and, therefore, that $\lim_{|t|\to+\infty} F_u(t(c, B)) = +\infty$. This shows that F is infinite at infinity when restricted to S_1. Any convex and continuous function in a Hilbert space is weakly lower semicontinuous, and admits a minimum when it is infinite at infinity. We thus conclude the existence of a minimizer (c^*, B^*) of F and therefore

$$A^*_{\mathrm{wca}}(w) = w + c^* + B^* w \tag{2.24}$$

is an optimal affine recovery map. ∎

2.3. The best affine map in the stochastic setting. In the stochastic setting, assuming that u has finite second order moments, the optimal map that minimizes the mean square error (1.9) is given by the conditional expectation

$$A^*_{\mathrm{ms}}(w) = \mathbb{E}(u \mid P_W u = w), \tag{2.25}$$

that is, the expectation of the posterior distribution p_w of u conditioned to the observation of w. Various sampling strategies have been developed in order to approximate the posterior and its expectation; see [13] for a survey. These approaches come at a significant computational cost since they require a specific sampling for each instance w of observed data. In the parametric PDE setting, each sample requires one solve of the forward problem.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

On the other hand, it is well known that an optimal affine map $A^*_{\rm msa}$ for the mean square error can be explicitely derived from the first and second order statistics of u. We briefly recall this derivation by using an arbitrary orthonormal basis $(e_1,\dots,e_m)$ of W that we complement into an orthonormal basis $(e_j)_{j\geq 1}$ of V. We write

$$u=\sum_{j\geq 1} w_j e_j \quad \text{and} \quad \overline{u}=\mathbb{E}(u)=\sum_{j\geq 1}\overline{w}_j e_j, \quad \overline{w}_j:=\mathbb{E}(w_j), \tag{2.26}$$

as well as

$$\widetilde{u}=u-\overline{u}=\sum_{j\geq 1}\widetilde{w}_j e_j, \quad \widetilde{w}_j:=w_j-\overline{w}_j. \tag{2.27}$$

An affine recovery map of the form (2.14) leaves the coordinates $w_1,\dots,w_m$ unchanged and recovers for each $i\geq 1$

$$w^*_{m+i}=c_i+\sum_{j=1}^m b_{i,j}w_j, \tag{2.28}$$

which can be rewritten as

$$w^*_{m+i}=\overline{w}_{m+i}+d_i+\sum_{j=1}^m b_{i,j}\widetilde{w}_j. \tag{2.29}$$

Since $E_{\rm ms}(A)=\sum_{i\geq 1}\mathbb{E}(|w^*_{m+i}-w_{m+i}|^2)$, the numbers d_i and $b_{i,j}$ are found by separately minimizing each term. By the Pythagorean theorem one has

$$\mathbb{E}(|w^*_{m+i}-w_{m+i}|^2)=|d_i|^2+\mathbb{E}\Big(\Big|\sum_{j=1}^m b_{i,j}\widetilde{w}_j-\widetilde{w}_{m+i}\Big|^2\Big), \tag{2.30}$$

which shows that we should take $d_i=0$. Minimizing the second term leads to the orthogonal projection equations

$$\sum_{j=1}^m b_{i,j}t_{j,l}=t_{m+j,l}, \quad l=1,\dots,m, \tag{2.31}$$

which involve the entries of the covariance matrix

$$\mathbf{S}:=(t_{i,j}), \quad t_{i,j}:=\mathbb{E}(\widetilde{w}_i\widetilde{w}_j). \tag{2.32}$$

Therefore, with the block decomposition

$$\mathbf{S}=\begin{pmatrix}\mathbf{S}_{1,1} & \mathbf{S}_{1,2}\\ \mathbf{S}_{2,1} & \mathbf{S}_{2,2}\end{pmatrix}, \tag{2.33}$$

corresponding to the splitting of rows and columns from $\{1,\dots,m\}$ and $\{m+1,m+2,\dots\}$, one obtains that the matrix $\mathbf{B}=(b_{i,j})$ that defines the optimal affine map satisfies $\mathbf{S}_{1,1}\mathbf{B}^{\rm T}=\mathbf{S}_{1,2}$ and, therefore,

$$\mathbf{B}=\mathbf{S}_{2,1}\mathbf{S}^{-1}_{1,1}, \tag{2.34}$$

where we have used the symmetry of $\mathbf{S}$. In other words,

$$A^*_{\rm msa}(w)=w+P_{W^\perp}\overline{u}+B\tilde{w}, \tag{2.35}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

where the linear transform $B \in \mathcal{L}(W, W^\perp)$ is represented by the matrix $\mathbf{B}$ in the basis $(e_j)_{j\geq 1}$.

The optimal affine recovery map $A^*_{\rm msa}$ agrees with the optimal map $A^*_{\rm ms}$ in the particular case where u has Gaussian distribution, therefore entirely characterized by its average $\overline{u}$ and covariance matrix $\mathbf{S}$. To see this, assume for simplicity that V is finite dimensional. The distribution of $\mathbf{u} = (w_j)_{j\geq 1}$ has density proportional to $\exp(-\frac{1}{2}\langle \mathbf{T}\widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}\rangle)$, where $\mathbf{T} = \mathbf{S}^{-1}$. We expand the quadratic form into

$$\frac{1}{2}\langle \mathbf{T}\widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}\rangle = \frac{1}{2}\langle \mathbf{T}_{1,1}\widetilde{\mathbf{w}}, \widetilde{\mathbf{w}}\rangle + \langle \mathbf{T}_{2,1}\widetilde{\mathbf{w}}, \widetilde{\mathbf{w}}_\perp\rangle + \frac{1}{2}\langle \mathbf{T}_{2,2}\widetilde{\mathbf{w}}_\perp, \widetilde{\mathbf{w}}_\perp\rangle, \tag{2.36}$$

where $\widetilde{\mathbf{w}}_\perp = (\widetilde{w}_{m+j})_{j\geq 1}$ and $\widetilde{\mathbf{w}} = (\widetilde{w}_j)_{j=1,\dots,m}$, and where

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_{1,1} & \mathbf{T}_{1,2} \\ \mathbf{T}_{2,1} & \mathbf{T}_{2,2} \end{pmatrix} \tag{2.37}$$

is a block decomposition similar to that of $\mathbf{S}$. The distribution of the vector $\widetilde{\mathbf{w}}_\perp$ conditional to the observation of $\widetilde{\mathbf{w}}$ is also Gaussian and its expectation coincides with the minimum of the quadratic form

$$Q_{\mathbf{w}}(\widetilde{\mathbf{w}}_\perp) = \frac{1}{2}\langle \mathbf{T}_{2,2}\widetilde{\mathbf{w}}_\perp, \widetilde{\mathbf{w}}_\perp\rangle + \langle \mathbf{T}_{2,1}\widetilde{\mathbf{w}}, \widetilde{\mathbf{w}}_\perp\rangle. \tag{2.38}$$

Therefore

$$\mathbb{E}(\widetilde{\mathbf{w}}_\perp \,|\, \widetilde{\mathbf{w}}) = -\mathbf{T}_{2,2}^{-1}\mathbf{T}_{2,1}\widetilde{\mathbf{w}} = \mathbf{S}_{2,1}\mathbf{S}_{1,1}^{-1}\widetilde{\mathbf{w}} = \mathbf{B}\widetilde{\mathbf{w}}, \tag{2.39}$$

which shows that

$$A^*_{\rm ms}(w) = \mathbb{E}(u \,|\, P_W u = w) = A^*_{\rm msa}(w). \tag{2.40}$$

One main interest of the above discussed stochastic setting is that the best affine map is now explicitely given by the second order statistics, in view of (2.34). This contrasts with the deterministic setting in which the optimal affine map is obtained by minimization of the convex functional F from (2.16) and does not generally have a simple explicit expression. Algorithms for solving this minimization problem are the object of the next section.

Only for particular cases where $\mathcal{M}$ has a simple geometry does the best affine map $A^*_{\rm wca}$ in the deterministic setting have a simple expression. One typical example is when $\mathcal{M}$ is an ellipsoid described by an equation of the form

$$\langle \mathbf{T}\widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}\rangle \leq 1 \tag{2.41}$$

for a symmetric positive matrix $\mathbf{T}$. Then, the set $\mathcal{M}_w = \mathcal{M} \cap V_w$ is also an ellipsoid associated with the above quadratic form $Q_{\mathbf{w}}$. The coordinates of its center are therefore given by the equation $\widetilde{\mathbf{w}}_\perp = -\mathbf{T}_{2,2}^{-1}\mathbf{T}_{2,1}\widetilde{\mathbf{w}}$, which is the same as that defining the conditional expectation in (2.39). This shows that, in the particular case of an ellipsoid, (i) the optimal map $A^*_{\rm wc}$ agrees with the optimal affine recovery map $A^*_{\rm wca}$ for the worst case error, and (ii) it has an explicit expression which agrees with the optimal map $A^*_{\rm msa}$ for the mean square error when the prior is a Gaussian with $\mathbf{T}$ as inverse covariance matrix.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

3. Algorithms for optimal affine recovery.

3.1. Discretization and truncation.

We have seen that the optimal affine recovery map is obtained by minimizing the convex function

$$F(c,B) = \sup_{u\in\mathcal{M}} \|P_{W^\perp}u - c - BP_W u\|, \tag{3.1}$$

over $W^\perp \times \mathcal{L}(W, W^\perp)$. This optimization problem cannot be solved exactly for two reasons:

(i) The sets $W^\perp$ as well as $\mathcal{L}(W, W^\perp)$ are infinite dimensional when V is infinite dimensional.

(ii) One single evaluation of $F(c, B)$ requires, in principle, to explore the entire manifold $\mathcal{M}$.

The first difficulty is solved by replacing V by a subspace Z_N of finite dimension $\dim(Z_N) = N$ that approximates the solution manifold $\mathcal{M}$ with an accuracy of smaller order than that expected for the recovery error. One possibility is to use a finite element space $Z_N = V_h$ of sufficiently small mesh size h. However its resulting dimension $N = N(h)$ needed to reach the accuracy could still be quite large. An alternative is to use reduced model spaces Z_N which are more efficient for the approximation of $\mathcal{M}$, as we discuss further.

We therefore minimize $F(c, B)$ over $\widetilde{W}^\perp \times \mathcal{L}(W, \widetilde{W}^\perp)$, where $\widetilde{W}^\perp$ is the orthogonal complement of W in the space $W + Z_N$, and obtain an affine map $\widetilde{A}_{\rm wca}$ defined by

$$\tilde{A}_{\rm wca}(w) = w + \overline{c} + \overline{B}w, \quad (\overline{c}, \overline{B}) := \operatorname{argmin}\{F(c,B) :\ c \in \widetilde{W}^\perp, B \in \mathcal{L}(W, \widetilde{W}^\perp)\}. \tag{3.2}$$

In order to compare its performance with that of $A^*_{\rm wca}$, we first observe that

$$\|P_{W^\perp}u - P_{\widetilde{W}^\perp}u\| \le \varepsilon_N := \sup_{u\in\mathcal{M}} \operatorname{dist}(u, Z_N). \tag{3.3}$$

For any $(c, B) \in W^\perp \times \mathcal{L}(W, W^\perp)$, we define $(\widetilde{c}, \widetilde{B}) \in \widetilde{W}^\perp \times \mathcal{L}(W, \widetilde{W}^\perp)$ by $\widetilde{c} = P_{\widetilde{W}^\perp} c$ and $\widetilde{B} = P_{\widetilde{W}^\perp} \circ B$. Then, for any $u \in \mathcal{M}$,

$$\begin{aligned}\|P_{W^\perp}u - \widetilde{c} - \widetilde{B}u\| &\le \|P_{\widetilde{W}^\perp}(P_{W^\perp}u - c - BP_W u)\| + \|P_{W^\perp}u - P_{\widetilde{W}^\perp}u\| \\ &\le \|P_{W^\perp}u - c - BP_W u\| + \varepsilon_N.\end{aligned}$$

It follows that we have the framing

$$E(A^*_{\rm wca}, \mathcal{M}) \le E(\widetilde{A}_{\rm wca}, \mathcal{M}) \le E(A^*_{\rm wca}, \mathcal{M}) + \varepsilon_N, \tag{3.4}$$

which shows that the loss in the recovery error is at most of the order ε_N.

To understand how large N should be, let us observe that a recovery map A of the form (2.14) takes it value in the linear space

$$F_{m+1} = \mathbb{R}c + \operatorname{ran}(B), \tag{3.5}$$

which has dimension $m + 1$. It follows that the recovery error is always larger than the approximation error by such a space. Therefore

$$E_{wc}(A^*_{\rm wca}, \mathcal{M}) \ge \delta_{m+1}(\mathcal{M}), \tag{3.6}$$

where $\delta_{m+1}(\mathcal{M})$ is the Kolmogorov n-width defined by (1.25) for $n = m+1$. Therefore, if we could use the space $Z_n := E_n$ that exactly achieve the infimum in (1.25), we

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

would be ensured that, with $N = m+1$, the additional error $\varepsilon_N = \delta_{m+1}(\mathcal{M})$ in (3.4) is of smaller order than $E_{wc}(A^*_{\rm wca}, \mathcal{M})$. As a result we would obtain the framing

$$E(A^*_{\rm wca}, \mathcal{M}) \le E(\widetilde{A}_{\rm wca}, \mathcal{M}) \le 2E(A^*_{\rm wca}, \mathcal{M}). \tag{3.7}$$

In practice, since we do not have access to the n-width spaces, we use instead the reduced basis spaces $Z_n := V_n$ which are expected to have comparable approximation performances in view of the results from [2, 12]. We take N larger than m but of comparable order.

The second difficulty is solved by replacing the set $\mathcal{M}$ in the supremum that defines $F(c, B)$ by a discrete training set $\widetilde{\mathcal{M}}$, which corresponds to a discretization $\widetilde{Y}$ of the parameter domain Y, that is,

$$\widetilde{\mathcal{M}} := \{u(y) \,:\, y \in \widetilde{Y}\} \tag{3.8}$$

with finite cardinality.

We therefore minimize over $\widetilde{W}^\perp \times \mathcal{L}(W, \widetilde{W}^\perp)$ the function

$$\widetilde{F}(c, B) = \sup_{u \in \widetilde{\mathcal{M}}} \|P_{W^\perp} u - c - BP_W u\|, \tag{3.9}$$

which is computable. The additional error resulting from this discretization can be controlled from the resolution of the discretization. Namely, let $\varepsilon > 0$ be the smallest value such that $\widetilde{\mathcal{M}}$ is an ε-approximation net of $\mathcal{M}$, that is, $\mathcal{M}$ is covered by the V-balls $B(u, \varepsilon)$ for $u \in \widetilde{\mathcal{M}}$. Then, we find that

$$\widetilde{F}(c, B) \le F(c, B) \le \widetilde{F}(c, B) + \varepsilon \|B\|_{\mathcal{L}(W, \widetilde{W}^\perp)}, \tag{3.10}$$

which shows that the additional recovery error will be of the order of ε amplified by the norm of the linear part of the optimal recovery map.

One difficulty is that the cardinality of ε-approximation nets become potentially intractable for small ε as the parameter dimension becomes large, due to the curse of dimensionality. This difficulty also occurs when constructing a reduced basis by a greedy selection process which also needs to be performed in sufficiently dense discretized sets. Recent results obtained in [9] show that in certain relevant instances ε approximation nets can be replaced by random training sets of smaller cardinality. One interesting direction for further research is to apply similar ideas in the context of the present paper.

3.2. Optimization algorithms. As already brought up in the previous section, the practical computation of $\widetilde{A}_{\rm wc}$ consists in solving

$$\min_{(c,B) \in \widetilde{W}^\perp \times \mathcal{L}(W, \widetilde{W}^\perp)} \sup_{u \in \widetilde{\mathcal{M}}} \|P_{W^\perp} u - c - BP_W u\|^2. \tag{3.11}$$

The numerical solution of this problem is challenging due to its lack of smoothness (the objective function is convex but nondifferentiable) and its high dimensionality (for a given target accuracy ε_N, the cardinality of $\widetilde{\mathcal{M}}$ might be large). One could use classical subgradient methods, which are simple to implement. However these schemes only guarantee a very slow $O(k^{-1/2})$ convergence rate of the objective function, where k is the number of iterations. This approach did not give satisfactory results in our case; due to the slow convergence, the solution update of one iteration falls below machine precision before approaching the minimum close enough; see Figure 3.1. This

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

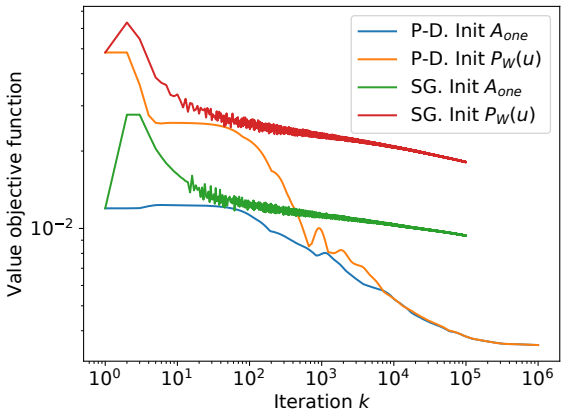


FIG. 3.1. *Convergence of the objective function for two different optimization algorithms and starting guesses. P.D. = Primal-Dual splitting. S.G. = Subgradient. Here,* $m = 40$.

has motivated the use of a primal-dual splitting method which is known to ensure a $O(1/k)$ convergence rate on the partial duality gap. We next describe this method, but only briefly, as a detailed analysis would make us deviate too far from the main topic of this paper. A complete analysis with further examples of application will be presented in a forthcoming work [14].

We assume without loss of generality that $\dim(W + V_N) = m + N$ and that $\dim \widetilde{W}^{\perp} = N$. Let $\{\psi_i\}_{i=1}^{m+N}$ be an orthonormal basis of $W + V_N$ such that $W = \operatorname{span}\{\psi_1, \ldots, \psi_m\}$. Since for any $u \in V$,

$$P_{W+V_N} u = \sum_{i=1}^{m+N} u_i \psi_i,$$

the components of u in W can be given in terms of the vector $\mathbf{w} = (u_i)_{i=1}^{m}$ and the ones in $\widetilde{W}^{\perp}$ with $\mathbf{u} = (u_{i+m})_{i=1}^{N}$.

We now consider the finite training set

$$\widetilde{\mathcal{M}} := \{u^1, \ldots, u^J\}, \quad J := \#(\widetilde{\mathcal{M}}) < \infty, \tag{3.12}$$

and denote by $\mathbf{w}^j$ and $\mathbf{u}^j$ the vectors associated with the snapshot functions u^j for $j = 1, \ldots, J$. One may express the problem (3.11) as the search for

$$\min_{\substack{(\mathbf{R},\mathbf{b}) \in \\ \mathbb{R}^{N\times m}\times\mathbb{R}^N}} \max_{j=1,\ldots,J} \|\mathbf{u}^j - \mathbf{R}\mathbf{w}^j - \mathbf{b}\|_2^2. \tag{3.13}$$

Concatenating the matrix and vector variables $(\mathbf{R}, \mathbf{b})$ into a single $\mathbf{x} \in \mathbb{R}^{m(N+1)}$, we rewrite the above problem as

$$\min_{\mathbf{x}\in\mathbb{R}^{m(N+1)}} \max_{j=1,\ldots,J} f_j(\mathbf{Q}_j \mathbf{x}), \tag{3.14}$$

where $\mathbf{Q}_j \in \mathbb{R}^{N\times m(N+1)}$ is a sparse matrix built using the coefficients of $\mathbf{w}^j$ and $f_j(\mathbf{y}) := \|\mathbf{u}^j - \mathbf{y}\|_2^2$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

The key observation to build our algorithm is that problem (3.14) can be equivalently written as a minimization problem on the epigraphs, i.e.,

$$\begin{aligned} &\min_{(\mathbf{x},t)\in\mathbb{R}^{m(N+1)}\times\mathbb{R}^+} t \quad \text{subject to} \quad f_j(\mathbf{Q}_j\mathbf{x}) \le t, \quad j=1,\dots,J \\ \iff &\min_{(\mathbf{x},t)\in\mathbb{R}^{m(N+1)}\times\mathbb{R}^+} t \quad \text{subject to} \quad (\mathbf{Q}_j\mathbf{x},t) \in \mathrm{epi}_{f_j}, \quad j=1,\dots,J, \end{aligned} \tag{3.15}$$

or, in a more compact (and implicit) form,

$$\min_{(\mathbf{x},t)\in\mathbb{R}^{m(N+1)}\times\mathbb{R}^+} t + \sum_{j=1}^{J} \iota_{\mathrm{epi}_{f_j}}(\mathbf{Q}_j\mathbf{x},t), \tag{$\mathrm{P}_{\mathrm{epi}}$}$$

where, for any nonempty set S the indicator function ι_S has value 0 on S and $+\infty$ on S^c.

This problem takes the following canonical expression, which is amenable to a primal-dual proximal splitting algorithm:

$$\min_{(\mathbf{x},t)\in\mathbb{R}^{m(N+1)}\times\mathbb{R}} G(\mathbf{x},t) + F\circ L(\mathbf{x},t). \tag{3.16}$$

Here, G is the projection map for the second variable

$$G(\mathbf{x},t) = t, \tag{3.17}$$

the linear operator L is defined by

$$L(\mathbf{x},t) := ((\mathbf{Q}_1\mathbf{x},t),(\mathbf{Q}_2\mathbf{x},t),\dots,(\mathbf{Q}_J\mathbf{x},t)) \tag{3.18}$$

and acts from $\mathbb{R}^{m(N+1)}\times\mathbb{R}$ to $\times_{j=1}^{J}(\mathbb{R}^N\times\mathbb{R})$ and the function F acting from $\times_{j=1}^{J}(\mathbb{R}^N\times\mathbb{R})$ to $\mathbb{R}$ is defined by

$$F\Big((\mathbf{v}_1,t_1),\dots,(\mathbf{v}_J,t_J)\Big) := \sum_{j=1}^{J} \iota_{\mathrm{epi}_{f_j}}(\mathbf{v}_j,t_j). \tag{3.19}$$

Note that F is the indicator function of the Cartesian product of epigraphs.

Before introducing the primal-dual algorithm, some remarks are in order:

(i) We recall that if ϕ is a proper closed convex function on $\mathbb{R}^d$, its proximal mapping prox_ϕ is defined by

$$\mathrm{prox}_\phi(y) = \mathrm{argmin}_{\mathbb{R}^d}\Big(\phi(x) + \frac{1}{2}\|x-y\|_2^2\Big). \tag{3.20}$$

(ii) The adjoint operator L^* is given by

$$L^*\Big((\mathbf{v}_1,t_1),\dots,(\mathbf{v}_J,t_J)\Big) := \left(\sum_{j=1}^{J}\mathbf{Q}_j^T\mathbf{v}_j, \sum_{j=1}^{J} t_j\right). \tag{3.21}$$

It can be easily shown that the operator norm of L satisfies $\|L\|^2 \le J + \sum_{j=1}^{J}\|\mathbf{Q}_j\|^2$.

(iii) Both G and F are simple functions in the sense that their proximal mappings, prox_G and prox_F, can be computed in closed form. See [14] for details.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

The iterations of our primal-dual splitting method read for $k \geq 0$,

(3.22)
$$
\begin{aligned}
(\mathbf{x}, t)^{k+1} &= \mathrm{prox}_{\gamma_G G}\Big((\mathbf{x}, t)^k - \gamma_G L^*\Big(\Big((\mathbf{v}_1, \xi_1), \ldots, (\mathbf{v}_J, \xi_J)\Big)^k\Big)\Big),\\
(\bar{\mathbf{x}}, \bar{t})^{k+1} &= (\mathbf{x}, t)^{k+1} + \theta\Big((\mathbf{x}, t)^{k+1} - (\mathbf{x}, t)^k\Big),\\
\Big((\mathbf{v}_1, \xi_1), \ldots, (\mathbf{v}_J, \xi_J)\Big)^{k+1} &= \mathrm{prox}_{\gamma_F \hat{F}}\Big(\Big((\mathbf{v}_1, \xi_1), \ldots, (\mathbf{v}_J, \xi_J)\Big)^k + \gamma_F L(\bar{\mathbf{x}}, \bar{t})^{k+1}\Big),
\end{aligned}
$$

where $\hat{F}$ is the Fenchel–Legendre transform of F, $\gamma_G > 0$ and $\gamma_F > 0$ are such that $\gamma_G \gamma_F < 1/\|L\|^2$, and $\theta \in [-1, +\infty[$ (it is generally set to $\theta = 1$ as in [8]).

Algorithm 1 gives some guidelines and summarizes in an informal pseudocode style the main steps of the primal-dual approach (the implementation of the routine "BuildQ" is left to the reader).

Algorithm 1 Primal-dual algorithm: $\mathbf{R}, \mathbf{b} = \mathrm{PD}(\widetilde{\mathcal{M}}, \mathcal{M}_{\mathrm{greedy}}, W, K_{\max})$.

Input:
- training manifold $\widetilde{\mathcal{M}}$ for primal-dual iterations
- training manifold $\mathcal{M}_{\mathrm{greedy}}$ for greedy algorithm
- basis $\{\omega_i\}_{i=1}^m$ of measurement space W
- maximum number of iteration $K_{\max}$

Generate basis $\{v_i\}_{i=1}^N$ of V_N // e.g., with a greedy algorithm over $\mathcal{M}_{\mathrm{greedy}}$; see (3.27)
Build orthonormal basis $\{\psi_i\}_{i=1}^{m+N}$ of $W + V_N$ with a Gram–Schmidt procedure over $\{w_1, \ldots, w_m, v_1, \ldots, v_N\}$. In this way, $\widetilde{W}^\perp = \mathrm{span}\{\psi_{m+1}, \ldots, \psi_{m+N}\}$
Qlist, wlist, ulist = []
for all $u \in \widetilde{\mathcal{M}}$ **do** // Build matrices $\mathbf{Q}^j$ of (3.14)
  $\mathbf{w} = \{\langle u, \psi_i\rangle\}_{i=1}^m$, $\mathbf{u} = \{\langle u, \psi_i\rangle\}_{i=m+1}^{N+m}$, $\mathbf{Q} = \mathrm{BuildQ}(\mathbf{w})$
  Qlist.append($\mathbf{Q}$), wlist.append($\mathbf{w}$), ulist.append($\mathbf{u}$)
end for
Estimate $\|L\|$ // e.g., with power method
Set γ_G and γ_F such that $\gamma_G \gamma_F < 1/\|L\|^2$
$\bar{\mathbf{x}} = \mathbf{x} = \mathrm{zeros}(m(N+1))$ // starting guess for $A : W \to V$ set to $A(w) = w$
$t = 1$ // starting guess for $t > 0$
$\Big((\mathbf{v}_1, \xi_1), \ldots, (\mathbf{v}_J, \xi_J)\Big) = L(\mathbf{x}, t)$ // starting guess dual variables
for k in $[0, K_{\max}]$ **do** // primal-dual iterations
  $(\mathbf{x}_{\mathrm{old}}, t_{\mathrm{old}}) = (\mathbf{x}, t)$
  $\Big((\mathbf{v}_1, \xi_1), \ldots, (\mathbf{v}_J, \xi_J)\Big) = \mathrm{prox}_{\gamma_F \hat{F}}\Big(\Big((\mathbf{v}_1, \xi_1), \ldots, (\mathbf{v}_J, \xi_J)\Big) + \gamma_F L(\bar{\mathbf{x}}, \bar{t})\Big)$
  $(\mathbf{x}, t) = \mathrm{prox}_{\gamma_G G}\Big((\mathbf{x}, t) - \gamma_G L^*\Big(\Big((\mathbf{v}_1, \xi_1), \ldots, (\mathbf{v}_J, \xi_J)\Big)\Big)\Big)$
  $(\bar{\mathbf{x}}, \bar{t}) = (\mathbf{x}, t) + \theta\Big((\mathbf{x}, t) - (\mathbf{x}_{\mathrm{old}}, t_{\mathrm{old}})\Big)$
end for
Retrieve $\mathbf{R}, \mathbf{c}$ by appropriately reshaping $\mathbf{x}$
Output: $\mathbf{R}, \mathbf{c}$

To illustrate the relevance of this algorithm for our purposes, we compare its performance with a standard subgradient method. Figure 3.1 plots the convergence history of the objective function across the iterations of both optimization methods

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

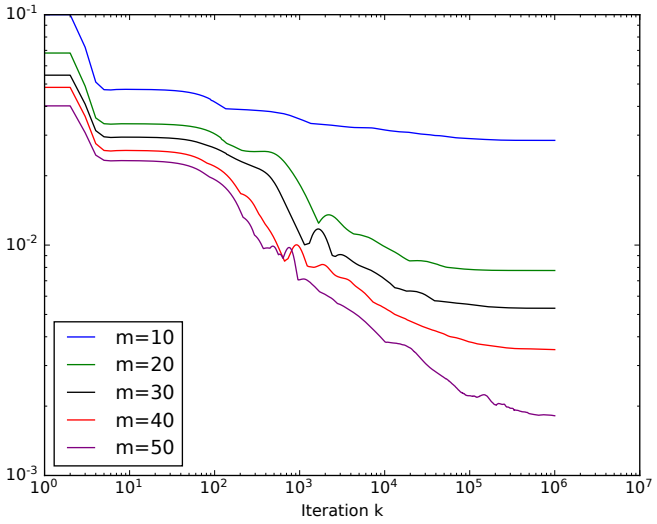


FIG. 3.2. *Convergence of the objective function in the primal-dual iterations for* $m = 10, 20, 30, 40, 50$.

for the example described in the next section ($m = 40$, $N = 110$, and $J = 10^3$). Two different reconstruction maps have been considered as starting guesses: the minimal V-norm recovery map given by $A(w) = w = P_W u$, and the one-space algorithm A_{n^*} based on reduced basis spaces V_n with an optimal choice n^* for n. The convergence plot shows the superiority of the primal-dual method which converges to the same minimal value of the objective function after 10^5 iterations regardless of the intialization, while the subgradient method fails to reach it since its increments fall below machine precision.

For the same numerical example described next, we vary m and consider $m = 10, 20, 30, 40, 50$. Figure 3.2 gives the convergence of the reconstruction error over the training set $\widetilde{\mathcal{M}}$ across the primal-dual iterations (for simplicity, we took P_{W_m} as the starting guess for $A^{(m)}_{\mathrm{wca}}$). To make sure that we reach convergence, we perform 10^6 iterations for each case. As expected, we observe in this figure that the final value of the objective function decreases as we increase the value of m (the reconstruction error decreases as we increase the number of measurements).

3.3. Numerical tests. We present some numerical experiments, aiming primarily at comparing the three above discussed recovery maps in terms of the maximum reconstruction error: the one-space affine map A_n, the best affine map A^*_{msa} for the mean square error, and the best affine map A^*_{wca}. for the worst case error. In addition, we also consider the mimimum V norm reconstruction map $A(w) = w = P_W u$. The results highlight the superiority of the best affine algorithm with respect to the reconstruction error. This comes however at the cost of a computationally intensive training phase as previously described.

We consider the elliptic problem

$$\begin{cases} -\mathrm{div}\Big(a(y)\nabla u\Big) = f, & x \in D, \\ u(x) = 0, & x \in \partial D, \end{cases} \tag{3.23}$$

on the unit square $D =]0,1[^2$ with a certain parameter dependence in the field a.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

More precisely, for a given $p \geq 1$, we consider "checkerboard" random fields where $a(y)$ is piecewise constant on a $p \times p$ subdivision of the unit square:

$$D = \bigcup_{i,j=0}^{p-1} S_{i,j}$$

with

$$S_{i,j} := \left[\frac{i}{p}, \frac{i+1}{p}\right[\times \left[\frac{j}{p}, \frac{j+1}{p}\right[, \qquad i,j \in 0,\dots,p-1.$$

The random field is defined as

$$a(y) = 1 + \frac{1}{2}\sum_{i,j=0}^{p-1} \chi_{S_{i,j}} y_{i,j}, \tag{3.24}$$

where χ_S denotes the characteristic function of a set S, and the $y_{i,j}$ are random coefficients that are independent each with identical uniform distribution on $[-1,1]$. Thus, our vector of parameters is

$$\mathbf{y} = (y_{i,j})_{i,j=0}^{p-1} \in \mathbb{R}^{p\times p}.$$

In our numerical tests, we take $p = 4$, that is, 16 parameters, and work in the ambient space $V = H_0^1(D)$. All the sets of snapshots used for training and validating the reconstruction algorithms have been computed by first generating a certain number J of random parameters $\mathbf{y}^1,\dots,\mathbf{y}^J$ with each $\mathbf{y}^i \in [-1,1]^{p\times p}$, and then solving the variational form of (3.23) in $V = H_0^1(D)$ using $\mathbb{P}_1$ finite elements on a regular grid of mesh size $h = 2^{-7}$. This gives the corresponding solutions $u_h^i = u_h(\mathbf{y}^i)$ that are used in the computations. To ease the reading, in the following we drop the dependence on h in the notation.

The sensor measurements are modeled with linear functionals that are local averages of the form

$$\ell_{\mathbf{x},\tau}(u) = \int_D u(\mathbf{r})\varphi_\tau(\mathbf{r}-\mathbf{x})\,\mathrm{d}\mathbf{r}, \tag{3.25}$$

where

$$\varphi_\tau(\mathbf{r}) \propto \exp(-|\mathbf{r}|/2\tau^2) \tag{3.26}$$

is a radial function such that $\int \varphi_\tau = 1$. The parameter $\tau > 0$ represents the spread around the center $\mathbf{x}$. For the observation space W of our example, we randomly select $m = 50$ centers $\mathbf{x}_i \in [0.1, 0.9]^2$ and spreads $\tau_i \in [0.05, 0.1]$, and compute the Riesz representers $\omega_{\mathbf{x}_i,\tau}$ of $\ell_{\mathbf{x}_i,\tau}$ in $H_0^1(D)$. We then set

$$W \coloneqq \operatorname{span}\{\omega_{\mathbf{x}_i,\tau}\}_{i=1}^M$$

which is a space of dimension $m = 50$. Figure 3.3 shows the m centers $\mathbf{x}_i$. As an example, the figure also plots the function $\omega_{\mathbf{x}_i,\tau}$ for $i = 10$, which has center $\mathbf{x}_i = (0.23, 0.75)$ and spread $\tau_i = 0.06$.

As explained in section 3.1, the first step to compute the best algorithm in practice consists in replacing $V = H_0^1(D)$ by a finite dimensional space that approximates the solution manifold $\mathcal{M}$ at an accuracy smaller than the one expected for the recovery

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

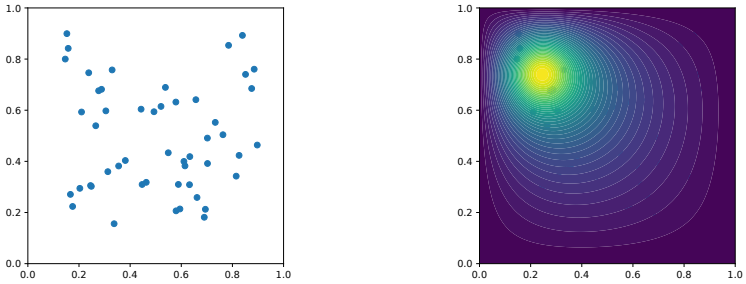


FIG. 3.3. *Sensor locations and the function* $\omega_{\mathbf{x}_i,\tau_i}$ *for* $i = 10$ *(*$\mathbf{x}_i = (0.23, 0.75)$ *and* $\tau_i = 0.06$*).*

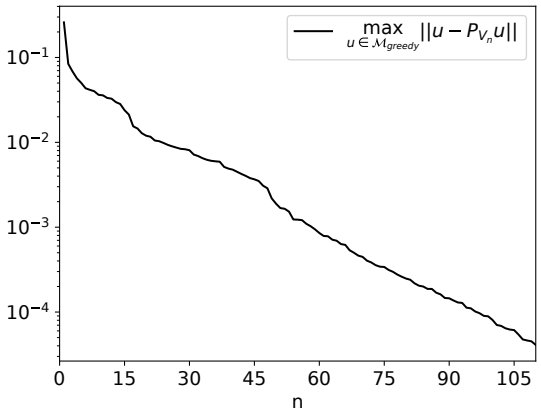


FIG. 3.4. *Greedy algorithm: decay of the error* $e_n^{(greedy)} = \max_{u \in \mathcal{M}_{greedy}} ||u - P_{V_n} u||$.

error. Here, we replace V by $W + V_N$, where V_N is a reduced basis of dimension $N = 110$ that has been generated by running the classical greedy algorithm from [7] over a training set $\mathcal{M}_{\text{greedy}}$ of 10^3 snapshots. We recall that an idealized version is defined for $n \geq 1$ as

$$u_n \in \operatorname*{argmax}_{u \in \mathcal{M}_{\text{greedy}}} \|u - P_{V_{n-1}} u\|, \quad V_n := V_{n-1} \oplus \mathbb{R} u_n = \operatorname{span}\{u_1, \ldots, u_n\}, \tag{3.27}$$

with the convention $V_0 := \{0\}$. Figure 3.4 gives the decay of the error

$$e_n^{(\text{greedy})} = \max_{u \in \mathcal{M}_{\text{greedy}}} \|u - P_{V_n} u\|$$

across the greedy iterations.

We next estimate the truncation accuracy ε_N defined in (3.3). This has been done by computing the maximum of the error $\|u - P_{V_N} u\|$ over the training set $\mathcal{M}_{\text{greedy}}$ supplemented by a test set $\mathcal{M}_{\text{test}}$, also of 10^3 snapshots. We obtain the estimate

$$\max_{u \in \mathcal{M}_{\text{greedy}} \cup \mathcal{M}_{\text{test}}} \|u - P_{V_N} u\| \leq \varepsilon_N = 5.10^{-5}.$$

In the comparison of the three different reconstruction algorithms, we want to illustrate the impact of the number of measurements that are used. To do this, we consider the nested subspaces

$$W_m = \operatorname{span}\{\omega_{\mathbf{x}_i,\tau_i}\}_{i=1}^m \subset W$$

for $m = 10$, 20, 30, 40, 50 so that $W_{50} = W$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

For the computation of the best affine algorithm, we generate a new training set $\widetilde{\mathcal{M}}$ of 10^3 snapshots which we project into $W + V_N$. This projected set, which we denote by $P_{W+V_N}\widetilde{\mathcal{M}}$ with a slight abuse of notation, is used to compute

$$\widetilde{A}^{(m)}_{\text{wca}}(u) = \widetilde{c}^{(m)} + \widetilde{B}^{(m)} P_{W_m} u, \quad m = 10,\, 20, \ldots,\, 50,$$

by running the primal-dual algorithm of section 3.2. We have added the indices m to stress that the algorithm depends on it.

For the comparison with the three other reconstruction algorithms, we evaluate

$$e^{(m)}_{\text{wca}} = \max_{u \in \mathcal{M}_{\text{test}}} ||u - \widetilde{A}^{(m)}_{\text{wca}}(P_{W_m} u)||, \quad m = 10,\, 20, \ldots,\, 50.$$

We stress the fact that the three sets $\mathcal{M}_{\text{greedy}}$, $\widetilde{\mathcal{M}}$, and $\mathcal{M}_{\text{test}}$ are *different*. We compare this value with the performance of a straightforward reconstruction with the minimal V-norm recovery map,

$$e^{(m)}_{\text{mvn}} = \max_{u \in \mathcal{M}_{\text{test}}} ||u - P_{W_m} u||, \quad m = 10,\, 20, \ldots,\, 50,$$

with the mean square approach,

$$e^{(m)}_{\text{msa}} = \max_{u \in \mathcal{M}_{\text{test}}} ||u - \widetilde{A}^{(m)}_{\text{msa}}(P_{W_m} u)||, \quad m = 10,\, 20, \ldots,\, 50,$$

and with the best one-space affine algorithm,

$$e^{(m)}_{\text{one}} = \min_{1 \le n \le m} e^{(m,n)}_{\text{one}},$$

where

$$e^{(m,n)}_{\text{one}} = \max_{u \in \mathcal{M}_{\text{test}}} ||u - A^{(m)}_n(P_{W_m} u)||, \quad m = 10,\, 20, \ldots,\, 50. \tag{3.28}$$

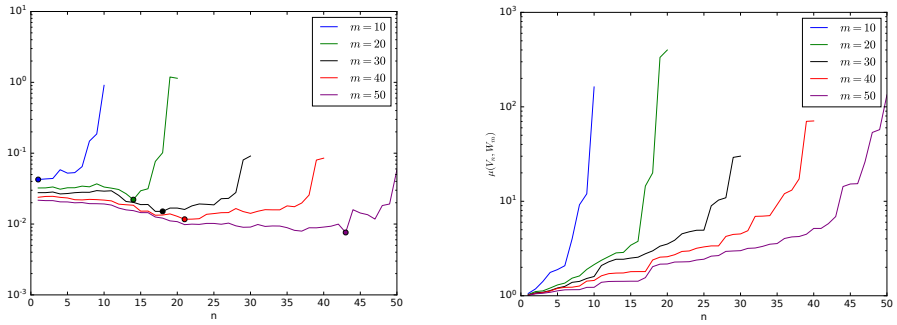
Some remarks on the computation of the one-space algorithm are in order. First of all, we have used the average

$$\bar{u} := \frac{1}{\#\mathcal{M}_{\text{greedy}}} \sum_{u \in \mathcal{M}_{\text{greedy}}} u$$

as our offset. For $m \le M$ and $n \le m$ given, the one-space affine algorithm $A^{(m)}_n$ is the one involving the spaces W_m and $\widetilde{V}_n = \bar{u} + V_n$, where $V_n = \text{span}\{u_1, \ldots, u_n\}$. Its performance is given by $e^{(m,n)}_{\text{one}}$ in formula (3.28). Figure 3.5(a) shows $e^{(m,n)}_{\text{one}}$ as a function of n and m. Note that, for a fixed m, the error $e^{(m,n)}_{\text{one}}$ reaches a minimal value $e^{m}_{\text{one}} = e^{(m,n^*)}_{\text{one}}$ for a certain dimension $n^* = n^*(m)$ of the reduced model, given by a thick dot in the figure. This behavior is due to the trade-off between the increase of the approximation properties of $\widetilde{V}_n$ as n grows and the degradation of the stability of the algorithm, given by the increase of $\mu(\widetilde{V}_n, W_m)$ with n. For our comparison purposes, we use $A^{(m)}_{\text{one}} = A^{(m)}_{n^*(m)}$, that is, the best possible one-space algorithm based on the reduced basis spaces.

Figure 3.6 shows the reconstruction errors $e^{(m)}_{\text{wca}}$, $e^{(m)}_{\text{mvn}}$, $e^{(m)}_{\text{msa}}$, and $e^{(m)}_{\text{one}}$ of the four different approaches for $m = 10,\, 20, \ldots,\, 50$. We also append a table with the values. We observe that a straightforward reconstruction with the minimal V-norm algorithm

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



(a) $e_{\text{one}}^{(m,n)}$: reconstruction error with $A_n^{(m)}$. (b) Stability factor $\mu(\widetilde{V}_n, W_m)$.

FIG. 3.5. *One-space algorithm.*

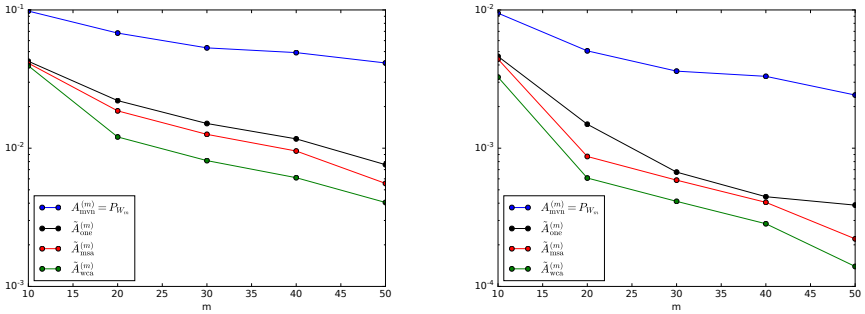


FIG. 3.6. *Comparison of the reconstruction errors (left: $H_0^1(D)$ norm; right: $L^2(D)$ norm).*

performs poorly in terms of approximation error and its quality improves only very mildly as we increase the number m of measurements. This justifies considering our three other, more sophisticated, reconstruction algorithms. In this respect, the results confirm first of all that $\widetilde{A}_{\text{wca}}^{(m)}$ is the best reconstruction algorithm. The mean square approach appears to be slightly superior to the one-space algorithm but still worse than the best affine algorithm. Note that the accuracy improvement between the best affine algorithm and the one-space and mean square algorithms is of about a half-order of magnitude for each m.

Last but not least, we give some illustrations on the reconstruction algorithms applied to a particular snapshot function u from the test set $\mathcal{M}_{\text{test}}$. The target function is given in Figure 3.7 and Figures 3.8 and 3.9 show the resulting reconstructions of u from $P_{W_m} u$ with our four different algorithms and for $m = 20$ and 40. Visually, the reconstructed functions look very similar. However, the difference in quality can be better appreciated in the plots of the spatial errors $|u(\mathbf{x}) - A^{(m)}(u)(\mathbf{x})|$ as well as in the derivatives and their corresponding spatial errors.

Let us briefly discuss the complexity of the primal-dual algorithm. At each iteration of the algorithm, the main bottleneck is the computation of L^* (see (3.21)). It requires one to do J matrix-vector products with the matrices $Q_j \in \mathbb{R}^{N \times m(N+1)}$ and then do a summation of the resulting vectors. The cost of these operations thus increases linearly with J in terms of computational time and memory resources. In fact, the limitation in memory was the main reason to fix $J = 10^3$ and not work with a larger number of training snapshots. Let us make a quick count on the cost

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

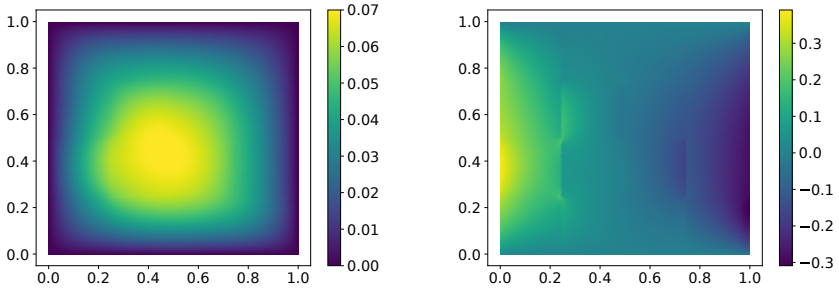
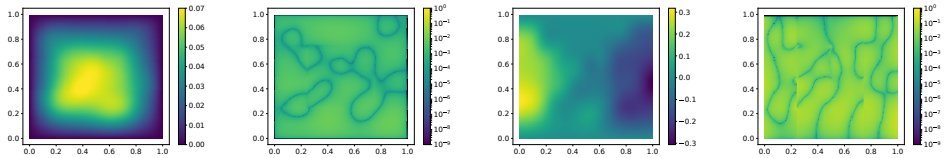
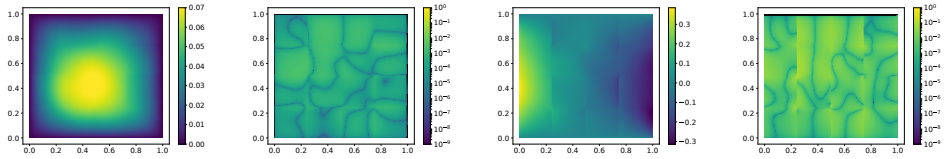


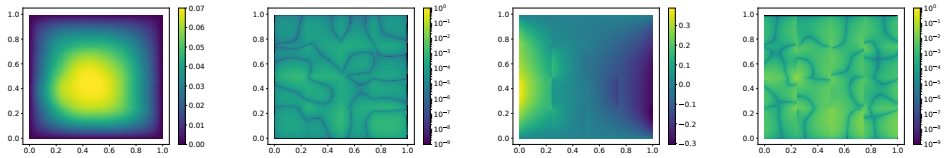
FIG. 3.7. *Function u (left) and $\partial u/\partial x$ (right). The reconstruction of this function is given in Figures* 3.8 *and* 3.9.



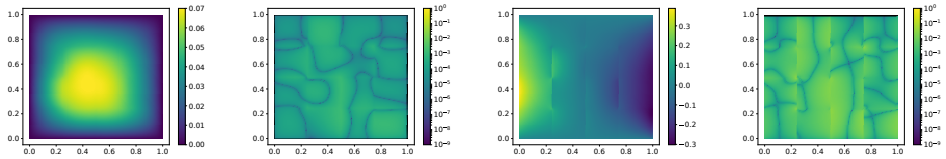
(a) Minimal V-norm: $P_{W_{20}}(u)$.



(b) One-space affine: $A^{(20)}_{\text{one}}\left(P_{W_{20}}(u)\right)$



(c) Mean square algorithm: $\widetilde{A}^{(20)}_{\text{msa}}\left(P_{W_{20}}(u)\right)$

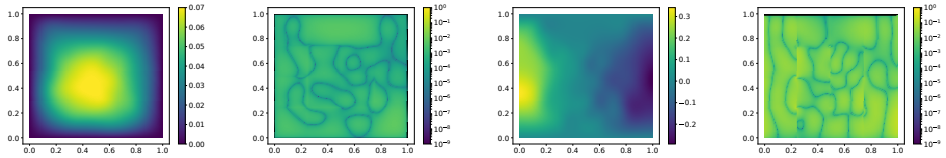


(d) Best affine: $\widetilde{A}^{(20)}_{\text{wca}}\left(P_{W_{20}}(u)\right)$

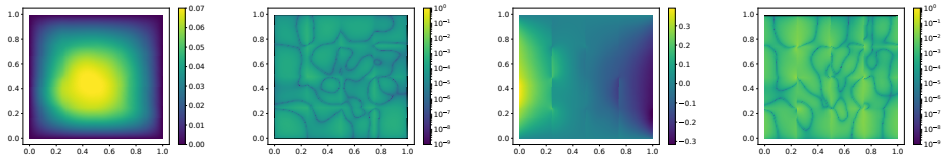
FIG. 3.8. *Reconstruction of the given function u ($m = 20$). For each reconstruction strategy* (i) *the two first figures are $A^{(m)}(u)(\mathbf{x})$ and the spatial errors $|u(\mathbf{x}) - A^{(m)}(u)(\mathbf{x})|$;* (ii) *the two last figures are $\frac{\partial A^{(m)}(u)}{\partial x}(\mathbf{x})$ and the spatial errors $|\frac{\partial u}{\partial x}(\mathbf{x}) - \frac{\partial A^{(m)}(u)}{\partial x}(\mathbf{x})|$.*

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

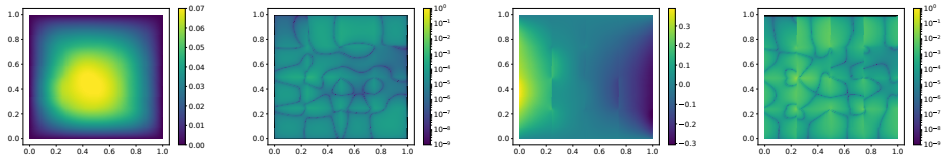
Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



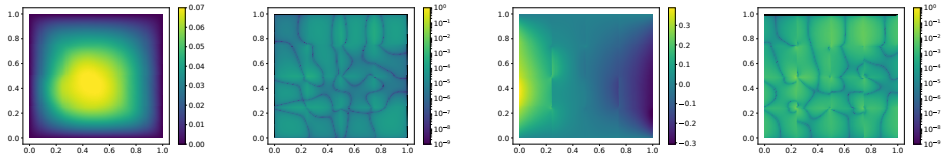
(a) Minimal V-norm: $P_{W_{40}}(u)$.



(b) One-space affine: $A^{(40)}_{\text{one}}\left(P_{W_{40}}(u)\right)$



(c) Mean square algorithm: $\widetilde{A}^{(40)}_{\text{msa}}\left(P_{W_{40}}(u)\right)$



(d) Best affine: $\widetilde{A}^{(40)}_{\text{wca}}\left(P_{W_{40}}(u)\right)$

FIG. 3.9. *Reconstruction of the given function u ($m = 40$). For each reconstruction strategy* (i) *the two first figures are $A^{(m)}(u)(\mathbf{x})$ and the spatial errors $|u(\mathbf{x}) - A^{(m)}(u)(\mathbf{x})|$;* (ii) *the two last figures are $\frac{\partial A^{(m)}(u)}{\partial x}(\mathbf{x})$ and the spatial errors $|\frac{\partial u}{\partial x}(\mathbf{x}) - \frac{\partial A^{(m)}(u)}{\partial x}(\mathbf{x})|$.*

in terms of the number of elements to store at each iteration. The matrices Q_j are sparse. For each row, there are $m+1$ nonnegative coefficients. Therefore we need to store $N(m+1)$ coefficients for each matrix; therefore, a total number of $JN(m+1)$ coefficients. In our case, $N = 110$ was carefully fixed to guarantee that

$$\max_{u \in \mathcal{M}_{\text{greedy}} \cup \mathcal{M}_{\text{test}}} \|u - P_{V_N} u\| \leq \varepsilon_N = 5.10^{-5}.$$

We have m ranging between 10 and 50. Thus the number of nonnegative elements that we have to store for each Q_j ranges between 1210 and 5610. Therefore, taking $J = 10^3$ as in our computation, we need to handle a total number of coefficients ranging between $1.21.10^6$ and $5.61.10^6$.

4. Conclusions. In this paper, we have studied the notion of a best affine recovery map for a general state estimation problem, that is, the map A^*_{wca} that minimizes

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

the worst case error $E_{\mathrm{wc}}(A, \mathcal{M})$ among all affine maps. This map is the solution to a convex optimization problem. Up to the additional perturbation induced from replacing $\mathcal{M}$ by a discrete training set $\widetilde{\mathcal{M}}$, it can be efficiently computed by a primal-dual optimization algorithm. Since any affine recovery map is associated with a reduced basis V_n plus an offset $\bar{u}$, the optimal affine map amounts to applying the one-space method from [20] using an affine reduced model space $\bar{u} + V_n$ which is optimal for the reconstruction task. Our numerical tests confirm that this choice outperforms standard reduced basis spaces, which are not specifically constructed for the recovery problem, but rather for the approximation of $\mathcal{M}$.

Our approach is readily applicable to any type of parametric PDEs, ranging from linear PDEs with affine parameter dependence to nonlinear PDEs with nonaffine parameter dependence. We outline its main limitations:

- The first essential limitation lies in its confinement to linear or affine recovery algorithms. Let us stress that, while state estimation is a linear inverse problem in the sense that the observed data w are generated from u by a linear projection, the optimal recovery map is typically nonlinear, due to the complex nonlinear geometry of the solution manifold $\mathcal{M}$ that constitutes the prior. Therefore, going beyond the results provided by our method requires the development of nonlinear recovery strategies. One possible approach, currently under investigation, is to (i) consider a collection of affine reduced model spaces $\{\bar{u}_k + V^k \,:\, k = 1, \dots K\}$, each of them of dimension $n_k \leq m$; (ii) use the observed data w to properly select a particular space from this collection; and (iii) apply the affine recovery algorithm using this particular data-dependent space. One standard way to obtain such local reduced model spaces is by splitting the parameter domain and searching for local reduced bases or proper orthogonal decomposition as, for example, proposed in [1] for forward modeling or in [15] for state estimation, as well as in [6]. However, the optimal affine recovery approach discussed in the present paper could also be used in order to improve on such constructions.
- The developed approach implicitly assumes that the parametric PDE model is perfect although, in full generality, the true physical state may not belong to $\mathcal{M}$. It is also assumed that measurements are noiseless. One way to readily extend this approach to the search of optimal affine maps that take into account model bias and measurement noise is as follows: suppose that the model bias is of size $\delta > 0$ in the sense that the real physical state u belongs to the offset

$$\mathcal{M}_\delta := \{v \in V \,:\, \exists y \in Y \text{ s.t. } \|v - u(y)\| \leq \delta\}.$$

  Suppose further that measurements are given with some deterministic noise, that is, we are given $P_W u + \eta$ such that $\|\eta\| \leq \sigma$ for some noise level σ. Then, the optimal affine map is given by

$$\min_{A \text{ affine}} \max_{\substack{u \in \mathcal{M}_\delta, \\ \|\eta\| \leq \sigma}} \|u - A(P_W u + \eta)\|.$$

  Once again we may emulate this optimization by introducing a discrete training set. Let us stress that such an optimization problem requires the knowledge of the size of the model bias δ and the noise level σ. While σ may be known for some applications, δ is very hard to estimate in practice.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

An assessment of the obtained estimation accuracy relies, however, on the availability of computable bounds for the distance of the reduced spaces from the solution manifold which may depend on the type of the PDE model.

REFERENCES

[1] D. Amsallem, C. Farhat, and M. J. Zahr, *Nonlinear model order reduction based on local reduced order bases*, Internat. J. Numer. Methods Engrg., 92 (2012), pp. 891–916.

[2] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk, *Convergence rates for greedy algorithms in reduced basis methods*, SIAM J. Math. Anal., 43 (2011), pp. 1457–1472.

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

[4] P. Binev, A. Cohen, O. Mula, and J. Nichols, *Greedy algorithms for optimal measurements selection in state estimation using reduced models*, SIAM J. Uncertain. Quantif., 6 (2018), pp. 1101–1126.

[5] B. Bojanov, *Optimal recovery of functions and integrals*, in First European Congress of Mathematics, Vol. I, Paris, 1992, Progr. Math. 119, Birkhauser, Basel, 1994, pp. 371–390.

[6] A. Bonito, A. Cohen, R. DeVore, D. Guignard, P. Jantsch, and G. Petrova, *Nonlinear methods for model reduction*, ESAIM Math. Model. Numer. Anal., to appear, 2020, https://doi.org/10.1051/m2an/2020057.

[7] A. Buffa, Y. Maday, A. T. Patera, C. Prud'homme, and G. Turinici, *A priori convergence of the greedy algorithm for the parameterized reduced basis*, ESAIM Math. Model. Numer. Anal., 46 (2012), pp. 595–603.

[8] A. Chambolle and T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, J. Math. Imaging Vision, 40 (2011), pp. 120–145.

[9] A. Cohen, W. Dahmen, R. DeVore, and J. Nichols, *Reduced Basis Greedy Selection Using Random Training Sets*, ESAIM Math. Model. Numer. Anal., 54 (2020), pp. 1509–1524.

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

[11] A. Cohen, R. DeVore, and C. Schwab, *Analytic regularity and polynomial approximation of parametric stochastic elliptic PDEs*, Anal. Appl., 9 (2011), pp. 11–47.

[12] R. DeVore, G. Petrova, and P. Wojtaszczyk, *Greedy algorithms for reduced bases in Banach spaces*, Constr. Approx., 37 (2013), pp. 455–466.

[13] M. Dashti and A. M. Stuart, *The Bayesian approach to inverse problems*, Handbook of Uncertainty Quantification, R. Ghanem, D. Higdon, and H. Owhadi, eds., Springer, Cham, Switzerland, 2015.

[14] J. Fadili and O. Mula, *Primal-Dual Splitting for the Max of Convex Functions*, in progress.

[15] F. Galarce, J.-F. Gerbeau, D. Lombardi, and O. Mula, *State Estimation with Nonlinear Reduced Models Application to the Reconstruction of Blood Flows with Doppler Ultrasound Images*, preprint, arXiv:1904.13367, 2019.

[16] P. L. Houtekamer and H. L. Mitchell, *Ensemble Kalman filtering*, Quart. J. R. Meteorol. Soc., 131 (2005), pp. 3269–3289.

[17] M. Kärcher, S. Boyaval, M. A. Grepl, and K. Veroy, *Reduced basis approximation and a posteriori error bounds for 4D-Var data assimilation*, J. Optim. Engrg., 19 (2018), pp. 663–695.

[18] K. Law, A. Stuart, and K. Zygalakis, *Data Assimilation, A Mathematical Introduction*, Texts in Appl. Math. 62, Springer, Cham, Switzerland, 2015.

[19] A. C. Lorenc, *A global three-dimensional multivariate statistical interpolation scheme*, Mon. Weather Rev., 109 (1981), pp. 701–721.

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

[21] C. A. Micchelli and T. J. Rivlin, *Lectures on optimal recovery*, in Numerical analysis, Lancaster, England, 1984, Lecture Notes in Math. 1129, Springer, Berlin, 1985, pp. 21–93.

[22] E. Novak and H. Wozniakowski, *Tractability of Multivariate Problems, Volume* I: *Linear Information*, EMS Tracts Math. 6, European Mathematical Society, Zurich, 2008.

[23] G. Rozza, D. B. P. Huynh, and A. T. Patera, *Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations $\tilde{N}$ application to transport and continuum mechanics*, Arch. Comput. Methods Engrg., 15 (2008), pp. 229–275.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[24] S. Sen, *Reduced-basis approximation and a posteriori error estimation for many-parameter heat conduction problems*, Numer. Heat Transf. B, 54 (2008), pp. 369–389.
[25] A. M. Stuart, *Inverse problems: A Bayesian perspective*, Acta Numer., 19 (2010), pp. 451–559.
[26] T. Taddei, *An adaptive parametrized-background data-weak approach to variational data assimilation*, ESAIM Math. Model. Numer. Anal., 51 (2017), pp. 1827–1858.
[27] B. Peherstorfer, K. Willcox, and M. Gunzburger, *Survey of multifidelity methods in uncertainty propagation, inference, and optimization*, SIAM Rev., 60 (2018), pp. 550–591.

Downloaded 05/18/21 to 129.252.33.113. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.