Deutsche Forschungsgemeinschaft

Priority Programme 1962

POD-Based Multiobjective Optimal Control of PDEs with Non-smooth Objectives

Dennis Beermann, Michael Dellnitz, Sebastian Peitz, Stefan Volkwein

SPP 1962 | Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization

Preprint Number SPP1962-024

received on June 19, 2017

Edited by
SPP1962 at Weierstrass Institute for Applied Analysis and Stochastics (WIAS)
Leibniz Institute in the Forschungsverbund Berlin e.V.
Mohrenstraße 39, 10117 Berlin, Germany
E-Mail: spp1962@wias-berlin.de

World Wide Web: http://spp1962.wias-berlin.de/

POD-based multiobjective optimal control of PDEs with non-smooth objectives

Dennis Beermann[1], **Michael Dellnitz**[2], **Sebastian Peitz**[2, *], and **Stefan Volkwein**[1]

[1] Department of Mathematics and Statistics, University of Konstanz
[2] Department of Mathematics, Paderborn University

A framework for set-oriented multiobjective optimal control of partial differential equations using reduced order modeling has recently been developed [1]. Following concepts from localized reduced bases methods, error estimators for the reduced cost functionals are utilized to construct a library of locally valid reduced order models. This way, a superset of the Pareto set can efficiently be computed while maintaining a prescribed error bound. In this article, this algorithm is applied to a problem with non-smooth objective functionals. It is shown that the extension to non-smooth problems can be realized in a straightforward manner and the implications on the numerical results are discussed.

1 Introduction

During the past decade, considering multiple objectives in the development of a system has steadily been gaining importance in a wide range of applications. There are nowadays only few problems where only one objective is of interest. When designing an application where the underlying dynamical system is given by a partial differential equation (PDE), this results in a multiobjective PDE-constrained optimization or optimal control problem. For example, in buildings we want to provide a comfortable room temperature while at the same time minimizing the energy consumption. This example illustrates that many objectives are often equally important and also contradictory such that we are forced to accept a trade-off between them. This results in a *multiobjective optimization problem* (MOP), where multiple objectives have to be minimized at the same time. Similar to scalar optimization problems, we want to find an optimal solution to this problem. However, in the multiobjective situation, we have to identify the set of *optimal compromises*, the so-called *Pareto set* (see [2] for a detailed introduction of theory and algorithms).

When addressing PDE-constrained MOPs, many function evaluations are required and hence, the computational effort quickly becomes prohibitively large. To overcome the problem of expensive function evaluations, model-order reduction is a widely used concept. Here, the underlying PDE is replaced by a surrogate model which can be solved much faster [3]. In this context, reduced-order models (ROMs) based on Galerkin projection and *Proper Orthogonal Decomposition (POD)* [4] have proven to be a powerful tool, in particular in a multi-query context such as parameter estimation, uncertainty quantification or optimization. During the last years, the first publications have appeared where PDE-constrained MOPs are solved using POD reduced-order modeling [5–8]. A set-oriented approach has recently been developed in [1], where – similar to localized reduced basis approaches [9] – a library of locally valid ROMs is used to efficiently solve PDE-constrained MOPs. In this article, these results are extended to non-smooth objective functions, where the non-smoothness is introduced by considering L^1 norms. To measure the error between the objective computed by using the POD Galerkin approximation and the objective based on the Galerkin finite element (FE) approximation, we make use of an a-posteriori error analysis, where the obtained error bound is based on an a-posteriori error bound for the state.

2 Considering Non-Smooth Objective Functionals

Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$ be a Lipschitz domain, $T > 0$ a final time and $\mathcal{U} := \mathbb{R}^m$ the control space. We consider the following *multiobjective optimal control problem (MOCP)* which contains L^1 objective functions and is thus non-smooth:

$$\min_{u\in\mathcal{U}} J(y,u) = \frac{1}{2}\left(\int_\Omega |y(T,x) - y_{d,1}(x)|\,\mathrm{d}x,\ \int_\Omega |y(T,x) - y_{d,2}(x)|\,\mathrm{d}x,\ |u|_2^2\right)^T \tag{1a}$$

subject to the semilinear PDE constraints

$$\begin{aligned} y_t(t,x) - \Delta y(t,x) + y^3(t,x) &= \sum_{i=1}^m u_i\chi_i(x) && \text{for } (t,x) \in (0,T)\times\Omega,\\ \tfrac{\partial y}{\partial n}(t,s) &= 0 && \text{for } (t,s) \in (0,T)\times\Sigma,\\ y(0,x) &= y_0(x) && \text{for } x \in \Omega, \end{aligned} \tag{1b}$$

* Corresponding author: e-mail speitz@math.upb.de

and the bilateral control constraints

$$u_{\mathsf{a}} \leq u \leq u_{\mathsf{b}} \quad \text{in}\, \mathcal{U}. \tag{1c}$$

The PDE (1b) is understood in a canonical weak sense for which the spaces $V := H^1(\Omega)$, $H := L^2(\Omega)$ and $V^{h\ell} \subset V^h \subset V$ are of importance. Let V^h denote a FE space and $V^{h\ell}$ be a low-dimensional subspace created by POD. It is then well-known that there are well-defined, non-linear solution operators

$$\mathcal{S} : \mathcal{U} \to L^2(0,T;V) \cap H^1(0,T;V') \cap C([0,T] \times \overline{\Omega}), \quad \mathcal{S}^h : \mathcal{U} \to H^1(0,T;V^h), \quad \mathcal{S}^{h\ell} : \mathcal{U} \to H^1(0,T;V^{h\ell})$$

for the weak, FE-discretized and POD-surrogate formulations of (1b). This allows us to define *reduced cost functions*

$$\hat{J}, \hat{J}^h, \hat{J}^{h\ell} : \mathcal{U} \to \mathbb{R}^3, \quad \hat{J}(u) = J(\mathcal{S}(u), u), \quad \hat{J}^h(u) = J(\mathcal{S}^h(u), u), \quad \hat{J}^{h\ell}(u) = J(\mathcal{S}^{h\ell}(u), u)$$

For a more detailed description, the reader is referred to [1]. The state equation considered therein is the same. However, the objectives in (1a) are non-differentiable in the first two components since the mapping $z \mapsto \|z\|_{L^1(\Omega)}$ is non-differentiable from $L^2(\Omega)$ to $\mathbb{R}$. Therefore, the problem is treated with the global, derivative-free subdivision algorithm presented in [10].

It has been shown in [1] that an a-posteriori error estimator $\Delta^{\mathsf{pr}}(y^{h\ell}) \in \mathbb{R}_+$ exists which satisfies

$$\|y^h(T) - y^{h\ell}(T)\|^2_{L^2(\Omega)} \;\leq\; \Delta^{\mathsf{pr}}(T, y^{h\ell})$$

Using this estimator, we present an error estimator for the cost function J:

Theorem 2.1 *Let $y^h \in H^1(0,T;V^h)$ and $y^{h\ell} \in H^1(0,T;V^\ell)$ be high-fidelity and reduced solutions to* (1b) *for a control $u \in \mathcal{U}$. The cost function error can then be estimated by*

$$\left|J_i(y^h,u) - J_i(y^{h\ell},u)\right| \;\leq\; \Delta^J(y^\ell) := \sqrt{|\Omega|\, \Delta^{\mathsf{pr}}(T, y^\ell)}, \quad i = 1,2 \tag{2}$$

Proof. Using the Hölder's inequality we have

$$\begin{aligned}\left|J_i(y^h,u) - J_i(y^{h\ell},u)\right| &= \left|\|y^h(T) - y_{d,i}\|_{L^1(\Omega)} - \|y^{h\ell}(T) - y_{d,i}\|_{L^1(\Omega)}\right| \\ &\leq \|y^h(T) - y^{h\ell}(T)\|_{L^1(\Omega)} \leq \|1\|_{L^2(\Omega)} \cdot \|y^h(T) - y^{h\ell}(T)\|_{L^2(\Omega)} \\ &\leq \sqrt{|\Omega|\, \Delta^{\mathsf{pr}}(T, y^{h\ell})}\end{aligned}$$

which gives the proof. □

3 An Algorithm for Set-Oriented Multiobjective Optimal Control of PDEs using POD

The aim in multiobjective optimization is to compute the *set of optimal compromises* – or the set of *non-dominated solutions* – for the concurrent objectives in (1). Non-dominated solutions are characterized by the fact that there exists no admissible solution which is superior in all objectives. For example, consider $u_1, u_2 \in \mathcal{U}_{\mathsf{ad}}$. If u_1 dominates u_2, then

$$\begin{aligned} & \hat{J}_i(u_1) \leq \hat{J}_i(u_2) \text{ for } i = 1, \dots, k \\ \text{and} \quad & \hat{J}_i(u_1) < \hat{J}_i(u_2) \text{ for at least one } i \in \{1, \dots, k\}, \end{aligned} \tag{3}$$

where k is the number of objectives. The set of non-dominated solutions is also known as the *Pareto set*, it's image is the *Pareto front*. Here the Pareto set is computed using a global, derivative-free subdivision algorithm for function values with uncertainties [10]. If these uncertainties can be bounded, i.e. $\Delta_i^J \leq \Delta^J_{\mathsf{max},i}$ for $i = 1, \dots, k$, then the non-dominance property (3) can be extended to *confident dominance*: u_1 confidently dominates u_2, if

$$\begin{aligned} & \hat{J}_i(u_1) + \Delta^J_{\mathsf{max},i} \leq \hat{J}_i(u_2) - \Delta^J_{\mathsf{max},i} \text{ for } i = 1, \dots, k \\ \text{and} \quad & \hat{J}_i(u_1) + \Delta^J_{\mathsf{max},i} < \hat{J}_i(u_2) - \Delta^J_{\mathsf{max},i} \text{ for at least one } i \in \{1, \dots, k\}. \end{aligned} \tag{4}$$

In the algorithm, the parameter domain is divided into a finite set of disjoint boxes (represented by a finite set of *sample points*) which are alternatingly *subdivided* and *selected* using a *non-dominance test*. In the latter, all boxes containing only confidently dominated points according to (4) are discarded. This way an arbitrarily close covering of the *set of almost Pareto optimal points* (i.e. all points satisfying (4)) is obtained.

Using the error estimate (2) from the previous section, we can develop an algorithm which uses a library $\mathcal{R}$ of locally valid ROMs in order to approximate the objectives for a large number of sample points (the indices of which are contained in the set $\mathcal{N}$) that need to be evaluated for the non-dominance test. Since the evlaution of the sample points requires an evaluation of (1b), the computation quickly becomes numerically infeasible. To this end, we want to approximate the FE solution using ROMs. The detailed procedure of evaluating all sample points using locally valid ROMs from $\mathcal{R}$ is described in Algorithm 1. Based on the 2-norm, the closest ROM is selected for each sample point and the function value as well as the error estimator are computed. If the accuracy is sufficient for all points, i.e. the prescribed error bound $\Delta^J_{\mathsf{sc,max}}$ is satisfied, then we proceed with the non-dominance test. Otherwise, an additional ROM is added to the library at the point u_{ref} with the largest error. This procedure is repeated until all points are approximated sufficiently accurately. In order to increase the numerical efficiency, the error estimator Δ^J is tightened by a heuristic factor C_{sc}, where $\Delta^J_{\mathsf{sc}} = \Delta^J / C_{\mathsf{sc}}$ (for details, see [1]).

Algorithm 1 (Greedy localized reduced basis approach)

Require: $\Delta^J_{\mathsf{sc,max}} \in \mathbb{R}^k$, C_{sc}, set of sample points $\mathcal{N} \subset \mathcal{U}_{\mathsf{ad}}$;
Consider all sample points as *insufficiently approximated*, i.e. $\mathcal{I} = \mathcal{N}$;
while $\mathcal{I} \neq \emptyset$ **do**
  for $i = 1, \dots, |\mathcal{I}|$ **do**
    Identify the *closest* ROM with respect to the 2-norm:

$$i_{\mathsf{close}} = \underset{j \in \{1,\dots,|\mathcal{R}|\}}{\arg\min} \left\| u^i - u^j_{\mathsf{ref}} \right\|_2.$$

    Compute $\hat{J}^{h\ell}(u^i)$ using ROM i_{close};
    Evaluate the error $\Delta^J_{\mathsf{sc}}(u^i)$ for ROM i_{close} using (2);
    if $\Delta^J_{\mathsf{sc}}(u^i) \leq \Delta^J_{\mathsf{sc,max}}$ **then**
      Accept $\hat{J}^{h\ell}(u^i)$ as sufficiently accurate;
      Remove i from the set $\mathcal{I}$;
  Identify the sample point with the largest error:

$$i_{\mathsf{max}} = \underset{s \in \mathcal{I}}{\arg\max}\, \Delta^J_{\mathsf{sc}}(u^s)$$

  Add ROM to library $\mathcal{R}$ with $u_{\mathsf{ref}} = u^{i_{\mathsf{max}}}$;
Remove all ROMs from $\mathcal{R}$ that have not been used;

4 Results

The Pareto sets and fronts of (1) obtained using both a FEM discretization and the localized reduced basis approach are shown in Figure 1. The control space is $\mathcal{U} = [-1, 1]^4$, which is also the size of the initial box in the subdivision algorithm. For further details of the numerical setup, the reader is referred to [1], from which only the norm in the objective functions has been changed here. We observe a good agreement between both the Pareto sets and the Pareto fronts. The error bound $\Delta^J_{\mathsf{sc,max}}$ for the objectives $\hat{J}^{h\ell}_1$ and $\hat{J}^{h\ell}_2$ is satisfied as desired. In order to also bound the error in the control space, further assumptions on the objectives have to be made. When comparing the two Pareto sets, we see that the inexact solution is *thicker*, which is a consequence of the inexactness in the objective functions and the condition (4) for confident dominance.

From a computational point of view, the ROM-based approach is significantly more efficient. Over all subdivision steps, the number of expensive FEM computations can be reduce by a factor of approximately 130. This ratio even increases during later iterations of the subdivision scheme – in the 20$^{\text{th}}$ subdivision step, it is approximately 220.

It is worth mentioning that when comparing Fig. 1 with [1, Fig. 5], the shape of the Pareto front has changed drastically from the smooth setup to this one. While the Pareto sets are qualitatively comparable, it is interesting to note that the distribution of ROMs turns out to be much more uniform in the smooth setting [1] than in Figure 1, where we observe an intense clustering in some regions of the admissible set. This can most likely be attributed to the non-smoothness of the cost functions in the current example, by which certain regions in the control space can only be poorly approximated by reduced models. In order to further increase the numerical efficiency, it would be interesting to investigate the reason for this and to adapt the algorithm, e.g. by directly treating these regions with FEM computations.

5 Conclusion

The present work demonstrates that the subdivision algorithm combined with localized reduced-order models for PDE-constrained MOCPs from [1] can easily be extended to non-smooth objectives. To this end, the cost function error estimators have been adapted to L^1 norms. The performance of the algorithm is very similar to the smooth setup, which demonstrates the strength of this derivative-free strategy. Additional work needs to be done both algorithmically and theoretically in order

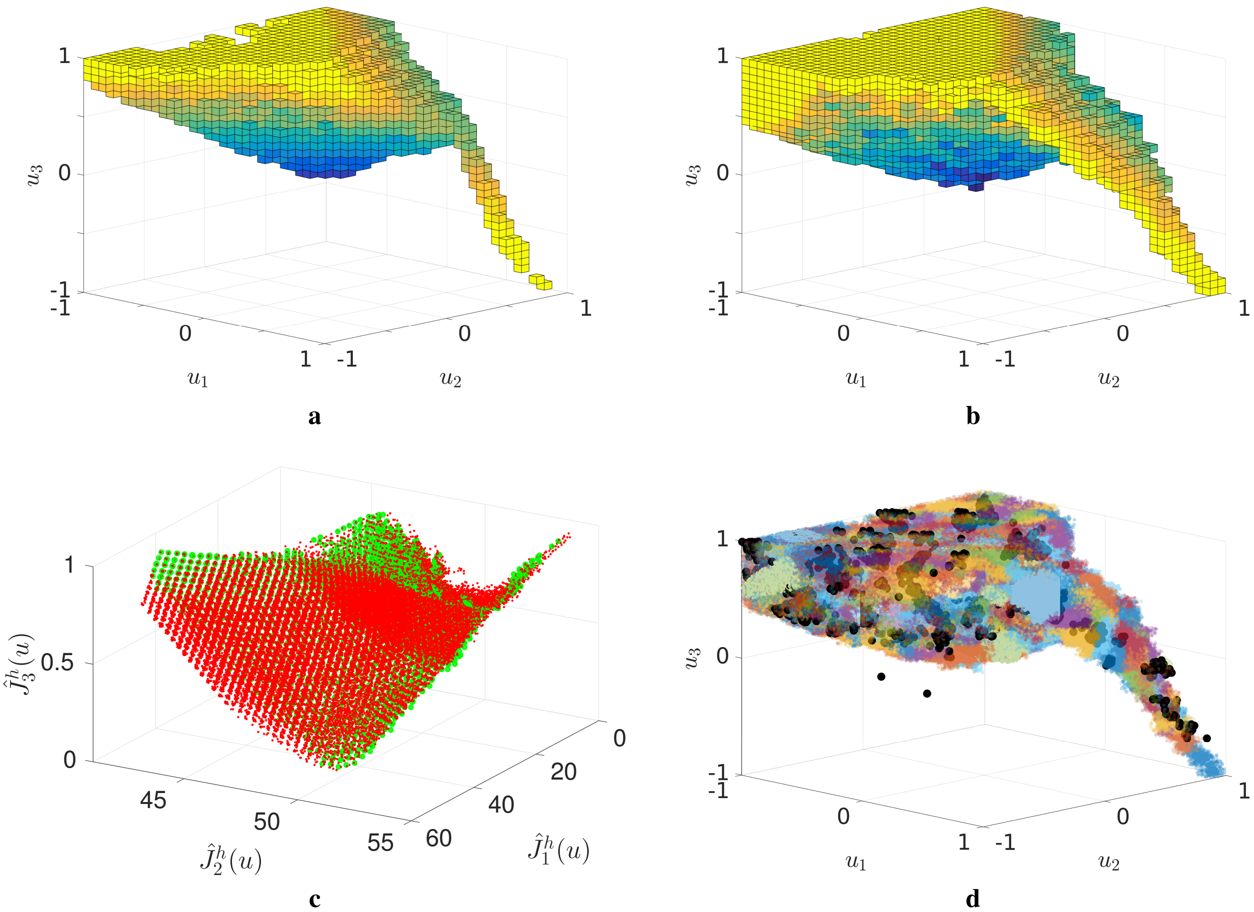

Fig. 1: **a** Pareto set of (1) after 20 subdivision steps using a FEM discretization. Projection onto the first three components of u, and u_4 is visualized by the box coloring. **b** The Pareto set based on local ROMs and the inexact sampling algorithm with $\Delta^{J}_{\mathsf{sc,max}} = (0.15, 0.15, 0)^{\top}$. **c** The corresponding Pareto fronts, where the FEM-based solution is shown in green and the POD-based solution in red. The points are the images of the box centers. **d** Assignment of sample points to ROMs. Each of the colored patches has been assigned to one of the ROMs which are represented by black dots.

to further improve the numerical efficiency and to investigate the clustering phenomenon of local ROMs in the non-smooth problem.

Acknowledgement: This work is supported by the Priority Programme SPP 1962 *Non-smooth and Complementarity-based Distributed Parameter Systems* of the German Research Foundation (DFG) and by the project *Hybrides Planungsverfahren zur energieeffizienten Wärme- und Stromversorgung von städtischen Verteilnetzen* funded by the German Ministry for Economic Affairs and Energy.

References

[1] D. Beermann, M. Dellnitz, S. Peitz, and S. Volkwein, Submitted (2017), https://spp1962.wias-berlin.de/preprints/017.pdf.
[2] M. Ehrgott, Multicriteria Optimization (Springer Berlin Heidelberg New York, 2005).
[3] W. Schilders, H. A. van der Vorst, and J. Rommes, Model Order Reduction (Springer Berlin Heidelberg, 2008).
[4] P. Holmes, J. Lumley, G. Berkooz, and C. Rowley, Turbulence, Coherent Structures, Dynamical Systems and Symmetry (Cambridge Monographs on Mechanics. Cambridge University Press, Cambridge, 2nd ed., 2012).
[5] L. Iapichino, S. Ulbrich, and S. Volkwein, Advances in Computational Mathematics (2017), DOI: 10.1007/s10444-016-9512-x.
[6] S. Peitz, S. Ober-Blöbaum, and M. Dellnitz, arXiv:1510.05819.
[7] L. Iapichino, S. Trenz, and S. Volkwein, Lecture Notes in Computational Science and Engineering **112**, 389–397 (2016).
[8] S. Banholzer, D. Beermann, and S. Volkwein, Submitted (2017), https://spp1962.wias-berlin.de/preprints/015.pdf.
[9] F. Albrecht, B. Haasdonk, S. Kaulmann, and M. Ohlberger, The localized reduced basis multiscale method, in: Proceedings of ALGORITHMY 2012, (2012), pp. 393–403.
[10] S. Peitz and M. Dellnitz, Gradient-based multiobjective optimization with uncertainties, in: NEO 2016 Proceedings (to appear, preprint: arXiv:1612.03815), (Springer, 2017).