ESAIM: COCV 32 (2026) 44
https://doi.org/10.1051/cocv/2026027

**ESAIM: Control, Optimisation and Calculus of Variations**
www.esaim-cocv.org

# ERROR ESTIMATES FOR A MULTIOBJECTIVE OPTIMAL CONTROL OF A POINTWISE TRACKING PROBLEM

Francisco Fuica[1,2,*] and Stefan Volkwein[3]

**Abstract.** We analyze a pointwise tracking multiobjective optimal control problem subject to the Poisson problem and bilateral control constraints. To approximate Pareto optimal points and the Pareto front numerically, we consider two different finite element-based scalarization techniques, namely the weighted-sum method and the reference point method, where in both methods many scalar-constrained optimization problems have to be solved. We prove *a priori* error estimates for both scalarizations. The underlying subproblems of either method are solved with a Barzilai-Borwein gradient method. Numerical experiments illustrate the accuracy of the proposed method.

**Mathematics Subject Classification.** 49M25, 58E17, 65N15, 65N30, 90C29.

Received May 19, 2025. Accepted April 1, 2026.

## 1. Introduction

Multiobjective optimization (see, *e.g.*, [1, 2]) plays an increasingly important role in modern applications, where several criteria are often of equal importance. The task in multiobjective optimization and multiobjective optimal control is therefore to compute the set of optimal compromises (the Pareto set) between the conflicting objectives. The advances in algorithms and the increasing interest in Pareto optimal solutions have led to a wide range of new applications related to optimal and feedback control – potentially with non-smoothness both on the level of the objectives or in the system dynamics; *cf.* [3–5], for instance. This results in new challenges such as dealing with expensive models (*e.g.*, governed by partial differential equations (PDEs)) and developing dedicated algorithms handling the non-smoothness. Since in contrast to single-objective optimization, the Pareto set generally consists of an infinite number of optimal solutions, the computational effort can quickly become challenging, which is particularly problematic when the objectives are costly to evaluate or when a solution has to be presented very quickly.

This article deals with the multiobjective optimal control of a pointwise tracking problem. Let $\Omega \subset \mathbb{R}^d$, with $d \in \{2, 3\}$, be an open, bounded and convex polytope with boundary $\partial\Omega$. We set $\mathbb{U} = L^2(\Omega)$ and $V = H_0^1(\Omega)$

*Keywords and phrases:* Multiobjective optimization of elliptic equations, pointwise tracking problems, scalarization methods, finite element method, *a priori* error estimates.

[1] Departamento de Matemática y Ciencia de la Computación, Universidad de Santiago de Chile, Santiago, Chile.
[2] Zukunftskolleg, Konstanz University, Universitätsstraße 10, 78464 Konstanz, Germany.
[3] Department of Mathematics and Statistics, Konstanz University, Universitätsstraße 10, 78464 Konstanz, Germany.

[*] Corresponding author: francisco.fuica@usach.cl

supplied with their common inner products

$$(\varphi,\psi)_{\mathbb{U}}=\int_\Omega \varphi\psi\,\mathrm{d}\boldsymbol{x} \text{ for } \varphi,\psi\in\mathbb{U} \quad\text{and}\quad (\varphi,\psi)_V=\int_\Omega \nabla\varphi\cdot\nabla\psi\,\mathrm{d}\boldsymbol{x} \text{ for } \varphi,\psi\in V.$$

The state and control Hilbert spaces are given by $\mathbb{Y}=V\cap H^2(\Omega)$ and $\mathbb{U}$, respectively, equipped with their usual topologies. Moreover, $\Omega_1=\{\boldsymbol{x}_1^1,\dots,\boldsymbol{x}_1^{n_1}\}$ and $\Omega_2=\{\boldsymbol{x}_2^1,\dots,\boldsymbol{x}_2^{n_2}\}$ are given two (non-empty) finite sets in $\Omega$, where $\Omega_1\cap\Omega_2=\emptyset$ is allowed. At the points in $\Omega_1$ and $\Omega_2$ we are given the desired states $y_k^i\in\mathbb{R}$ for $1\le i\le n_k$ and $1\le k\le 2$.

In this work, we will be interested in the study of the following bicriterial optimal control problem:

$$\min \mathcal{J}(y,u) \quad \text{subject to (s.t.)} \quad (y,u)\in\mathbb{Y}\times\mathbb{U}_{\mathsf{ad}} \text{ satisfies } -\Delta y=u \text{ in } \Omega,\quad y=0 \text{ on } \partial\Omega, \tag{$\mathbf{P}$}$$

where the objective $\mathcal{J}:\mathbb{Y}\times\mathbb{U}_{\mathsf{ad}}\to\mathbb{R}^2$ is given as

$$\mathcal{J}(y,u)=\begin{pmatrix}\mathcal{J}_1(y,u)\\ \mathcal{J}_2(y,u)\end{pmatrix}:=\frac{1}{2}\begin{pmatrix}\sum_{i=1}^{n_1}\big(y(\boldsymbol{x}_1^i)-y_1^i\big)^2+\lambda_1\|u\|_{\mathbb{U}}^2\\ \sum_{i=1}^{n_2}\big(y(\boldsymbol{x}_2^i)-y_2^i\big)^2+\lambda_2\|u\|_{\mathbb{U}}^2\end{pmatrix}, \tag{1.1}$$

with $\lambda_1,\lambda_2>0$ and the non-empty, closed, bounded, and convex set of admissible controls is defined by

$$\mathbb{U}_{\mathsf{ad}}:=\{u\in\mathbb{U}\,:\,u_{\mathsf{a}}(\boldsymbol{x})\le u(\boldsymbol{x})\le u_{\mathsf{b}}(\boldsymbol{x}) \text{ for almost all (f.a.a.) } \boldsymbol{x}\in\Omega\} \tag{1.2}$$

with $u_{\mathsf{a}},u_{\mathsf{b}}\in\mathbb{U}$ satisfying $u_{\mathsf{a}}\le u_{\mathsf{b}}$ in $\Omega$ almost everywhere (a.e.).

Error estimates for finite element (FE) discretizations of a *scalar-valued* version of the problem ($\mathbf{P}$) ($\mathcal{J}(y,u)=\mathcal{J}_1(y,u)\in\mathbb{R}$) have been previously derived in a number of works. For such a problem, in [6], the authors obtained *a priori* and *a posteriori* error estimates using continuous piecewise linear functions to approximate the optimal state and adjoint state variables and the so-called variational discretization for the optimal control variable. The *a priori* estimate derived for the approximation error of the control behaves as $\mathcal{O}(h^{2-d/2})$ [6], Theorem 3.2. Later, the authors of [7] proposed a fully discrete scheme that approximates the optimal state, adjoint state, and control variables with continuous piecewise linear functions and proved, in two dimensions, a $\mathcal{O}(h)$ convergence rate for the approximation error of the control variable [7], Theorem 5.1; error estimates for the variational discretization scheme were also analyzed [7], Theorem 5.2. In [8], based on the theory of Muckenhoupt weights and weighted Sobolev spaces, the authors analyzed a numerical scheme that discretizes the control variable with piecewise constant functions and the state and adjoint state variables with continuous piecewise linear functions. In this work, for the error approximation of the optimal control variable, the authors proved in two dimensions a nearly optimal convergence rate [8], Theorem 4.3 and in three dimensions a suboptimal convergence rate $\mathcal{O}(h^{1/2}|\log h|)$. This was later improved to $\mathcal{O}(h|\log h|)$ in [9], Theorem 6.6. In addition, the authors of [9] provide, in two dimensions, a $\mathcal{O}(h^2|\log h|^2)$ error estimate for the variational discretization [9], Theorem 7.5 and a post-processing scheme [9], Theorem 7.12. For appropriate extensions of the aforementioned results to problems involving the Stokes system, a semilinear elliptic equation, and the stationary Navier–Stokes equations, the interested reader is referred to [10–14].

In contrast to the previous advances, the analysis of approximation techniques for multiobjective optimal control problems is relatively scarce, with only a few works addressing it; see [15, 16]. To our knowledge, this is the first work that considers approximation techniques for a pointwise tracking multiobjective optimal control problem. This problem may be relevant in applications, where the state observations are performed at specific locations of the domain. One of the main challenges that arises by considering these observations in the cost function $\mathcal{J}$ is that the corresponding adjoint state variables $p_k$ ($k\in\{1,2\}$) possess reduced regularity properties: $p_k\in W_0^{1,\mathsf{p}}(\Omega)\setminus(V\cap C(\overline{\Omega}))$, where $\mathsf{p}<d/(d-1)$. This makes the analysis of *a priori* error estimates more involved.

We present the notion of *Pareto optimality* and investigate two parameter-dependent methods to derive the set of all Pareto optimal points – the so-called *Pareto front* – namely, the weighted-sum method (see, *e.g.*, [1], Chapter 3 and [2], pp. 78–85) and the reference point method (see, *e.g.*, [2], pp. 164–170 and [17]). For each one of the underlying scalar problems associated with these scalarization methods, we derive the existence of a unique solution and first-order optimality conditions. Moreover, we prove that solutions of these scalar problems are Pareto optimal for the multiobjective optimal control problem (**P**). We consider a FE discretization for each scalar problem and approximate the solution to the state and adjoint equations with continuous piecewise linear functions, whereas the control variable is approximated with piecewise constant functions. The main contribution of the present work is the proof of optimal ($d = 2$) and quasi-optimal ($d = 3$) *a priori* error estimates for the scalar problems associated with the weighted-sum method and the reference point method.

The paper is organized in the following manner: In the remainder of this section, we introduce the notation that we will use throughout the manuscript. In Section 2, we present a weak formulation for the bicriterial optimization problem (**P**) and recall notions of Pareto optimality and Pareto stationarity. Moreover, we present the weighted-sum method and the reference point method and explain how these scalarization methods can be used to characterize the Pareto stationary front. Suitable FE discretizations for the weighted-sum method and the reference point method are proposed in Section 3; *a priori* error estimates are also proved. Numerical experiments are presented in Section 4.

**Notation.** Throughout this work we use standard notation for Lebesgue and Sobolev spaces; *cf.* [18], for instance. Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach function spaces, we write $\mathcal{X} \hookrightarrow \mathcal{Y}$ to denote that $\mathcal{X}$ is continuously embedded in $\mathcal{Y}$. The norm of $\mathcal{X}$ is denoted by $\|\cdot\|_{\mathcal{X}}$. Given $p \in (1, \infty)$, we denote by $q$ its Hölder conjugate, that is, the real number such that $1/p + 1/q = 1$. In this work, vector inequalities must be understood componentwise. More precisely, given $n \in \mathbb{N}$ and $\mathfrak{a}, \mathfrak{b} \in \mathbb{R}^n$, we write $\mathfrak{a} \le \mathfrak{b}$ if $\mathfrak{a}_i \le \mathfrak{b}_i$ for all $i = 1, \ldots, n$; similarly $\mathfrak{a} \ge \mathfrak{b}$. We define the sets $\mathbb{R}^n_{\ge \mathfrak{a}} := \{\mathfrak{b} \in \mathbb{R}^n : \mathfrak{b} \ge \mathfrak{a}\}$ and $\mathbb{R}^n_{\le \mathfrak{a}} := \{\mathfrak{b} \in \mathbb{R}^n : \mathfrak{b} \le \mathfrak{a}\}$. For convenience, we shall write $\mathbb{R}^n_{\le} := \mathbb{R}^n_{\le 0}$ and $\mathbb{R}^n_{\ge} := \mathbb{R}^n_{\ge 0}$. The relation $a \lesssim b$ indicates that $a \le Cb$, with a constant $C > 0$ that depends neither on $a$, $b$ nor on the discretization parameters. The value of $C$ may change at each occurrence.

## 2. The optimal control problem

In this section, we recall basic facts for the problem (**P**). We start by briefly studying the state equation and its solution.

### 2.1. The state equation

For $u \in \mathbb{U}$, we consider the following weak formulation for the state equation: Find $y \in V$ such that

$$(y, \varphi)_V = (u, \varphi)_{\mathbb{U}} \quad \text{for all } \varphi \in V. \tag{2.1}$$

The Lax-Milgram lemma directly yields the existence of a unique solution $y \in V$ to (2.1). Moreover, the convexity of $\Omega$ implies that $y \in \mathbb{Y}$ holds; see, *e.g.*, [19], Theorems 3.2.1.2 and 4.3.1.4 for $d = 2$ and [19], Theorem 3.2.1.2 and [20], Section 4.3.1 for $d = 3$. Due to the compact embedding $H^2(\Omega) \hookrightarrow C(\overline{\Omega})$ [21], Theorem 6.3, Part III, it follows that $y \in C(\overline{\Omega})$ holds. Hence, a point evaluation of $y \in \mathbb{Y}$ is meaningful.

### 2.2. Weak formulation

We begin by defining the control-to-state map $S : \mathbb{U} \ni u \mapsto y \in \mathbb{Y}$, where $y = Su$ is the unique solution to (2.1). With $S$ at hand, we introduce the bicriterial reduced cost function $j : \mathbb{U} \to \mathbb{R}^2$ as $j(u) := \mathcal{J}(Su, u)$, that is,

$$j(u) = \begin{pmatrix} j_1(u) \\ j_2(u) \end{pmatrix} = \begin{pmatrix} \mathcal{J}_1(Su,u) \\ \mathcal{J}_2(Su,u) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \sum_{i=1}^{n_1} \left((Su)(\boldsymbol{x}_1^i) - y_1^i\right)^2 + \lambda_1 \|u\|_{\mathbb{U}}^2 \\ \sum_{i=1}^{n_2} \left((Su)(\boldsymbol{x}_2^i) - y_2^i\right)^2 + \lambda_2 \|u\|_{\mathbb{U}}^2 \end{pmatrix}. \tag{2.2}$$

We immediately note that $(Su)(\boldsymbol{x}_k^i) = y(\boldsymbol{x}_k^i)$ is meaningful for $1 \le i \le n_k$ and $k \in \{1,2\}$. Hence, the objective $j$ is also well-defined. The associated reduced optimal control problem reads:

$$\min j(u) \quad \text{s.t.} \quad u \in \mathbb{U}_{\mathsf{ad}}. \tag{$\hat{\mathbf{P}}$}$$

Problem ($\hat{\mathbf{P}}$) involves a vector-valued objective, which makes it necessary to introduce the concepts of *order relation* and *Pareto optimality* (see, *e.g.*, [1]): The point $\bar{u} \in \mathbb{U}_{\mathsf{ad}}$ is *Pareto* optimal for ($\hat{\mathbf{P}}$) if there is no other control $u \in \mathbb{U}_{\mathsf{ad}} \setminus \{\bar{u}\}$ satisfying $j(u) \le j(\bar{u})$ and $j_k(u) < j_k(\bar{u})$ for at least one $k \in \{1,2\}$. We also define the *Pareto set* and *Pareto front* as

$$\mathcal{P}_s := \{u \in \mathbb{U}_{\mathsf{ad}} \, : \, u \text{ is Pareto optimal}\} \subset \mathbb{U} \quad \text{and} \quad \mathcal{P}_f := j(\mathcal{P}_s) \subset \mathbb{R}^2,$$

respectively.

Note that both $j_1$ and $j_2$ in (2.2) are bounded from below and strictly convex. Therefore, similar arguments to those provided in the proof of [22], Theorem 2.14 yield that the *ideal vector* $\mathsf{v}^{\mathrm{id}} := (\mathsf{v}_1^{\mathrm{id}}, \mathsf{v}_2^{\mathrm{id}})$ given as

$$\mathsf{v}_1^{\mathrm{id}} := \min\{j_1(u) \, : \, u \in \mathbb{U}_{\mathsf{ad}}\} \quad \text{and} \quad \mathsf{v}_2^{\mathrm{id}} := \min\{j_2(u) \, : \, u \in \mathbb{U}_{\mathsf{ad}}\} \tag{2.3}$$

is well-defined.

We now present an auxiliary result that establishes differentiability properties of $j_1$ and $j_2$.

**Proposition 2.1** (Differentiability of $j_1$ and $j_2$)**.** *Let $u \in \mathbb{U}$. The objective functions $j_1$ and $j_2$, defined in* (2.2)*, are Fréchet differentiable. Their derivatives at $u$ are given as*

$$j_k'(u)w = \sum_{i=1}^{n_k} \left( \left((Su)(\boldsymbol{x}_k^i) - y_k^i\right)(Sw)(\boldsymbol{x}_k^i)\right) + \lambda_k\,(u,w)_{\mathbb{U}} \quad \textit{for any } w \in \mathbb{U} \textit{ and for } k \in \{1,2\}.$$

*Proof.* The fact that $j_1$ and $j_2$ are Fréchet differentiable is an immediate consequence of the linearity and boundedness of the control-to-state map $S$. The characterization of their derivatives is obtained by straightforward computations. □

Before we continue with our analysis, we introduce, given $u \in \mathbb{U}_{\mathsf{ad}}$ and $k \in \{1,2\}$, the *adjoint state* variable $p_k \in W_0^{1,\mathsf{p}}(\Omega)$ for any $\mathsf{p} \in (1, d/d-1)$ as the unique solution to the variational problem

$$\int_\Omega \nabla\varphi \cdot \nabla p_k \, \mathrm{d}\boldsymbol{x} = \sum_{i=1}^{n_k} \left(y(\boldsymbol{x}_k^i) - y_k^i\right)\varphi(\boldsymbol{x}_k^i) \quad \text{for all } \varphi \in W_0^{1,\mathsf{q}}(\Omega), \tag{2.4}$$

where $1/\mathsf{p} + 1/\mathsf{q} = 1$ and $y = Su$ hold. Note that the integral $\int_\Omega \nabla\varphi \cdot \nabla p_k \, \mathrm{d}\boldsymbol{x}$ is well defined for $\varphi \in W_0^{1,\mathsf{q}}(\Omega)$ and $p_k \in W_0^{1,\mathsf{p}}(\Omega)$. The right-hand side in (2.4) is well defined, as $\varphi \in W_0^{1,\mathsf{q}}(\Omega) \hookrightarrow C(\overline{\Omega})$ is satisfied for $\mathsf{q} > d$ [21], Theorem 6.3, Part III. Well-posedness of the problem (2.4) follows from [23], Theorem 2.

**Lemma 2.2** (Auxiliary result)**.** *Let $u \in \mathbb{U}$, $y = Su$, and $p_k \in W_0^{1,\mathsf{p}}(\Omega)$ for any $\mathsf{p} \in (2d/2+d, d/d-1)$ be the unique solution to (2.4) for $k \in \{1,2\}$. Then,*

$$(p_k, w)_{\mathbb{U}} = \sum_{i=1}^{n_k} \left(y(\boldsymbol{x}_k^i) - y_k^i\right)(Sw)(\boldsymbol{x}_k^i) \quad \textit{for all } w \in \mathbb{U}. \tag{2.5}$$

*Proof.* The proof closely follows the argument used in [24], Theorem 1. Let $k \in \{1,2\}$ be chosen. For $\mathsf{q} \in (d, 2d/d-2)$ we have $\mathbb{Y} \hookrightarrow W_0^{1,\mathsf{q}}(\Omega)$; see, *e.g.*, [21], Theorem 4.12. Thus, it follows that $Sw \in W_0^{1,\mathsf{q}}(\Omega)$ for every $w \in \mathbb{U}$ so that we can replace $\varphi = Sw$ in (2.4) to obtain

$$(Sw, p_k)_V = \sum_{i=1}^{n_k} \left(y(\boldsymbol{x}_k^i) - y_k^i\right)(Sw)(\boldsymbol{x}_k^i). \tag{2.6}$$

Notice that $p_k \notin V$, so that we cannot choose $\varphi = p_k$ in (2.1). Let $\{p_k^n\}_{n\in\mathbb{N}} \subset C_0^\infty(\Omega)$ such that $p_k^n \to p_k$ in $W_0^{1,\mathsf{p}}(\Omega)$ with $\mathsf{p} < d/d-1$. Set, for $n \in \mathbb{N}$, $\varphi = p_k^n$ and $u = w$ in (2.1). This results in

$$(Sw, p_k^n)_V = (w, p_k^n)_{\mathbb{U}}. \tag{2.7}$$

Since $p_k^n \to p_k$ in $W_0^{1,\mathsf{p}}(\Omega) \hookrightarrow \mathbb{U}$, we observe that $|(Sw, p_k^n - p_k)_V| \to 0$ and $|(w, p_k^n - p_k)_{\mathbb{U}}| \to 0$ as $n \to \infty$. Hence, passing to the limit in (2.7) yields $(Sw, p_k)_V = (w, p_k)_{\mathbb{U}}$, which concludes the proof. ☐

## 2.3. Scalarization techniques

The search of Pareto optimal points will be carried out using two methods: the weighted-sum method and the (Euclidean) reference point method (see, *e.g.*, [25], Sect. 4.3 and [26], Sect. 3, respectively). In what follows, we describe both methods and present first-order optimality conditions for their corresponding underlying scalar problems.

### *2.3.1. Weighted-sum method (WSM)*

For weights $\alpha_1, \alpha_2 > 0$ with $\alpha_1 + \alpha_2 = 1$, we set $\alpha := (\alpha_1, \alpha_2)$ and define the (scalar) function

$$W_\alpha(u) := \alpha_1 j_1(u) + \alpha_2 j_2(u) \quad \text{for every } u \in \mathbb{U}.$$

Then, the scalar-valued problem

$$\min W_\alpha(u) \quad \text{s.t.} \quad u \in \mathbb{U}_{\mathsf{ad}} \tag{$\hat{\mathbf{P}}_\alpha$}$$

is considered. The latter is denoted as the weighted-sum problem (with positive weights $\alpha_1$ and $\alpha_2$) corresponding to ($\hat{\mathbf{P}}$). The weighted-sum method is based on solving problem ($\hat{\mathbf{P}}_\alpha$) for varying $\alpha$.

In the next result, we prove the existence of solutions for ($\hat{\mathbf{P}}_\alpha$).

**Proposition 2.3** (Existence of a solution (WSM))**.** *Let $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2_>$ such that $\alpha_1 + \alpha_2 = 1$. Then, the scalar-valued optimal control problem ($\hat{\mathbf{P}}_\alpha$) has a unique solution $\bar{u}_\alpha \in \mathbb{U}_{\mathsf{ad}}$, which is Pareto optimal for problem ($\hat{\mathbf{P}}$).*

*Proof.* The fact that $j_1$ and $j_2$ are strictly convex and weakly lower semicontinuous, in combination with the fact that $\mathbb{U}_{ad}$ is weakly sequentially compact, yields the existence of a unique optimal control $\bar{u}_\alpha$ for ($\hat{\mathbf{P}}_\alpha$).

To prove the second statement of the proposition, we proceed by contradiction and assume that $\bar{u}_\alpha$ is not Pareto optimal for ($\hat{\mathbf{P}}$). Then, there exists $\hat{u} \in \mathbb{U}_{ad}$ satisfying $j(\hat{u}) \leq j(\bar{u}_\alpha)$ and $j_k(\hat{u}) < j_k(\bar{u}_\alpha)$ for some

$k \in \{1,2\}$. Hence, $\alpha_k j_k(\hat{u}) < \alpha_k j_k(\bar{u}_\alpha)$ $(\alpha_k > 0)$, which implies that $W_\alpha(\hat{u}) < W_\alpha(\bar{u}_\alpha)$. This contradicts the optimality of $\bar{u}_\alpha$ and concludes the proof. ☐

In view of Proposition 2.1 and Lemma 2.2, we infer that

$$\begin{aligned} W'_\alpha(u)w &= \sum_{k=1}^{2} \alpha_k \left( \sum_{i=1}^{n_k} ((Su)(\boldsymbol{x}_k^i) - y_k^i)(Sw)(\boldsymbol{x}_k^i) + \lambda_k (u,w)_{\mathbb{U}} \right) \\ &= \sum_{k=1}^{2} \alpha_k (p_k + \lambda_k u, w)_{\mathbb{U}} \quad \text{for any } w \in \mathbb{U}. \end{aligned}$$

With this characterization at hand, we state the first-order sufficient optimality conditions for the WSM whose proof follows by standard arguments; see, *e.g.*, [22], Lemma 2.21.

**Theorem 2.4** (First-order optimality condition (WSM))**.** *Let $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2_>$ be such that $\alpha_1 + \alpha_2 = 1$. If $\bar{u}_\alpha \in \mathbb{U}_{\mathsf{ad}}$ is an optimal solution to ($\hat{\mathbf{P}}_\alpha$), then we have*

$$\sum_{k=1}^{2} \alpha_k (\bar{p}_{\alpha,k} + \lambda_k \bar{u}_\alpha, u - \bar{u}_\alpha)_{\mathbb{U}} \geq 0 \quad \textit{for all } u \in \mathbb{U}_{\mathsf{ad}}, \tag{2.8}$$

*where $\bar{p}_{\alpha,k}$ denotes the unique solution to (2.4) with $\bar{y}_\alpha = S\bar{u}_\alpha$.*

Define $\bar{\mathfrak{p}}_\alpha = \alpha_1 \bar{p}_{\alpha,1} + \alpha_2 \bar{p}_{\alpha,2}$ and $\Lambda_\alpha = \alpha_1\lambda_1 + \alpha_2\lambda_2$. From (2.8) (see, *e.g.*, [22], Thm. 2.28) we infer that

$$\bar{u}_\alpha(\boldsymbol{x}) = \min\left\{u_{\mathsf{b}}(\boldsymbol{x}), \max\left\{u_{\mathsf{a}}(\boldsymbol{x}), -\Lambda_\alpha^{-1}\bar{\mathfrak{p}}_\alpha(\boldsymbol{x})\right\}\right\} \quad \text{f.a.a. } \boldsymbol{x} \in \Omega. \tag{2.9}$$

We note that $\bar{\mathfrak{p}}_\alpha \in W_0^{1,\mathsf{p}}(\Omega)$, for any $\mathsf{p} \in (1, {}^{d}/_{d-1})$, solves

$$(\varphi, \bar{\mathfrak{p}}_\alpha)_V = \sum_{k=1}^{2} \alpha_k \sum_{i=1}^{n_k} \left(\bar{y}_\alpha(\boldsymbol{x}_k^i) - y_k^i\right) \varphi(\boldsymbol{x}_k^i) \quad \text{for all } \varphi \in W_0^{1,\mathsf{q}}(\Omega),$$

where ${}^1/_{\mathsf{p}} + {}^1/_{\mathsf{q}} = 1$ and $\bar{y}_\alpha = S\bar{u}_\alpha$ hold. Hence, if $u_{\mathsf{a}}, u_{\mathsf{b}}$ are real constants, then [9], Lemma 6.1 immediately yields $\bar{u}_\alpha \in W^{1,\infty}(\Omega)$.

### *2.3.2. Reference point method (RPM)*

Given a *reference point* $\zeta = (\zeta_1, \zeta_2) \in \mathcal{P}_f + \mathbb{R}^2_\leq$ we introduce the distance function $R_\zeta : \mathbb{U} \to \mathbb{R}$ by

$$R_\zeta(u) := \frac{1}{2}\left\|j(u) - \zeta\right\|^2_{\mathbb{R}^2} = \frac{1}{2}\left((j_1(u) - \zeta_1)^2 + (j_2(u) - \zeta_2)^2\right) \quad \text{for } u \in \mathbb{U},$$

which measures the Euclidean distance between $j(u)$ and the reference point $\zeta$. Moreover, if $\zeta \leq \mathsf{v}^{\mathrm{id}}$ with $\mathsf{v}^{\mathrm{id}}$ defined in (2.3), then the mapping $R_\zeta$ is strictly convex since it involves the composition between a strictly convex nondecreasing function and a strictly convex function.

The idea behind the reference point method is that by finding a point $u$ such that $j(u)$ approximates $\zeta$ as best as possible, we shall obtain a Pareto optimal point for ($\hat{\mathbf{P}}$). Therefore, we have to solve the (Euclidean) *reference point problem*

$$\min R_\zeta(u) \quad \text{s.t.} \quad u \in \mathbb{U}_{\mathsf{ad}}, \tag{$\hat{\mathbf{P}}_\zeta$}$$

which is a scalar-valued minimization problem.

**Proposition 2.5** (Existence of a solution (RPM))**.** *Let $\zeta = (\zeta_1, \zeta_2) \in \mathcal{P}_f + \mathbb{R}^2_{\leq}$ such that $\zeta \leq \mathsf{v}^{\mathrm{id}}$. Then, the scalar-valued optimal control problem ($\hat{\mathbf{P}}_\zeta$) has a unique solution $\bar{u}_\zeta \in \mathbb{U}_{\mathsf{ad}}$, which is Pareto optimal for problem ($\hat{\mathbf{P}}$).*

*Proof.* The fact that $R_\zeta$ is strictly convex and weakly lower semicontinuous, in combination with the fact that $\mathbb{U}_{\mathsf{ad}}$ is weakly sequentially compact, yields the existence of a unique optimal control $\bar{u}_\zeta$ for ($\hat{\mathbf{P}}_\zeta$). To prove the second statement of the proposition, we proceed by contradiction and assume that $\bar{u}_\zeta$ is not Pareto optimal for ($\hat{\mathbf{P}}$). Then, there exists $\hat{u} \in \mathbb{U}_{\mathsf{ad}}$ satisfying $j(\hat{u}) \leq j(\bar{u}_\zeta)$ and $j_k(\hat{u}) < j_k(\bar{u}_\zeta)$ for some $k \in \{1, 2\}$. Hence, $0 \leq j_k(\hat{u}) - \zeta_k < j_k(\bar{u}_\zeta) - \zeta_k$, which implies that $R_\zeta(\hat{u}) < R_\zeta(\bar{u}_\zeta)$. The latter contradicts the optimality of $\bar{u}_\zeta$ and concludes the proof. ☐

For $u \in \mathbb{U}$ the use of Proposition 2.1 and Lemma 2.2 reveal that

$$\begin{aligned} R'_\zeta(u)w &= \sum_{k=1}^{2} (j_k(u) - \zeta_k) \left( \sum_{i=1}^{n_k} (Sw)(\boldsymbol{x}_k^i) \left( (Su)(\boldsymbol{x}_k^i) - y_k^i \right) + \lambda_k (u, w)_{\mathbb{U}} \right) \\ &= \sum_{k=1}^{2} (j_k(u) - \zeta_k)(p_k + \lambda_k u, w)_{\mathbb{U}} \quad \text{for any } w \in \mathbb{U}, \end{aligned}$$

from which we derive the following standard result.

**Theorem 2.6** (First-order optimality condition (RPM))**.** *Let $\zeta = (\zeta_1, \zeta_2) \in \mathcal{P}_f + \mathbb{R}^2_{\leq}$. If $\bar{u}_\zeta \in \mathbb{U}_{\mathsf{ad}}$ is an optimal solution to ($\hat{\mathbf{P}}_\zeta$), then*

$$\sum_{k=1}^{2} (j_k(\bar{u}_\zeta) - \zeta_k)(\bar{p}_{\zeta,k} + \lambda_k \bar{u}_\zeta, u - \bar{u}_\zeta)_{\mathbb{U}} \geq 0 \quad \textit{for all } u \in \mathbb{U}_{\mathsf{ad}}, \tag{2.10}$$

*where $\bar{p}_{\zeta,k}$ denotes the unique solution to (2.4) with $\bar{y}_\zeta = S\bar{u}_\zeta$.*

As in the Weighted-Sum Method (see (2.9)), we obtain a characterization for the optimal control $\bar{u}_\zeta$:

$$\bar{u}_\zeta(\boldsymbol{x}) = \min\{u_{\mathsf{b}}, \max\{u_{\mathsf{a}}, -\Lambda_\zeta^{-1} \bar{\mathfrak{p}}_\zeta(\boldsymbol{x})\}\} \quad \text{f.a.a. } \boldsymbol{x} \in \Omega,$$

where

$$\bar{\mathfrak{p}}_\zeta = (j_1(\bar{u}_\zeta) - \zeta_1)\bar{p}_{\zeta,1} + (j_2(\bar{u}_\zeta) - \zeta_2)\bar{p}_{\zeta,2}, \qquad \Lambda_\zeta = \lambda_1(j_1(\bar{u}_\zeta) - \zeta_1) + \lambda_2(j_2(\bar{u}_\zeta) - \zeta_2).$$

In particular, if $u_{\mathsf{a}}, u_{\mathsf{b}}$ are real constants, then $\bar{u}_\zeta \in W^{1,\infty}(\Omega)$ (see [9], Lem. 6.1).

## 3. The FE approximation

In this section, we present suitable FE discretizations for the scalarization techniques proposed in Section 2.3. Moreover, we prove error estimates for the error committed when approximating solutions of the scalarized problems presented in the previous section.

Let us first introduce the discrete setting in which we will operate [27, 28]. We denote by $\mathcal{T}_h = \{T\}$ a conforming partition, or mesh, of $\overline{\Omega}$ into closed simplices $T$ of size $h_T := \mathrm{diam}(T)$. Define $h := \max_{T \in \mathcal{T}_h} h_T$. We denote by $\mathbb{T} = \{\mathcal{T}_h\}_{h>0}$ a collection of conforming and quasi-uniform meshes $\mathcal{T}_h$. Given a mesh $\mathcal{T}_h \in \mathbb{T}$, we define the FE space of continuous piecewise polynomials of degree one that vanish on the boundary as

$$\mathbb{V}_h := \left\{ v^h \in C(\overline{\Omega}) : v^h|_T \in \mathbb{P}_1(T) \text{ for all } T \in \mathcal{T}_h \right\} \cap V \tag{3.1}$$

endowed with the inner product $(\cdot\,,\cdot)_V$. Let us also introduce the space

$$\mathbb{U}_h := \left\{ u^h \in L^\infty(\Omega) : u^h|_T \in \mathbb{P}_0(T) \text{ for all } T \in \mathcal{T}_h \right\} \subset \mathbb{U} \tag{3.2}$$

and the set of admissible controls

$$\mathbb{U}_{\mathsf{ad}}^h := \{ u \in \mathbb{U}_h \,:\, u_{\mathsf{a}}(\boldsymbol{x}) \leq u(\boldsymbol{x}) \leq u_{\mathsf{b}}(\boldsymbol{x}) \text{ f.a.a. } \boldsymbol{x} \in \Omega \}.$$

The FE-Galerkin approximation to (2.1) is given as follows: Find $y^h \in \mathbb{V}_h$ such that

$$(y^h, \varphi^h)_V = (u, \varphi^h)_{\mathbb{U}} \quad \text{for all } \varphi^h \in \mathbb{V}_h. \tag{3.3}$$

We introduce the discrete control-to-state map $S_h : \mathbb{U} \ni u \mapsto y^h \in \mathbb{V}_h$, where $y^h = S_h u$ is the unique solution to (3.3). The discrete function $j^h : \mathbb{U} \to \mathbb{R}^2$ is defined by

$$j^h(u) = \begin{pmatrix} j_1^h(u) \\ j_2^h(u) \end{pmatrix} := \frac{1}{2} \begin{pmatrix} \sum_{i=1}^{n_1} \left( (S_h u)(\boldsymbol{x}_1^i) - y_1^i \right)^2 + \lambda_1 \|u\|_{\mathbb{U}}^2 \\ \sum_{i=1}^{n_2} \left( (S_h u)(\boldsymbol{x}_2^i) - y_2^i \right)^2 + \lambda_2 \|u\|_{\mathbb{U}}^2 \end{pmatrix}. \tag{3.4}$$

Finally, given $u \in \mathbb{U}$ and $k \in \{1, 2\}$, we define $p_k^h \in \mathbb{V}_h$ as the unique solution to

$$(\varphi^h, p_k^h)_V = \sum_{i=1}^{n_k} \left( y^h(\boldsymbol{x}_k^i) - y_k^i \right) \varphi^h(\boldsymbol{x}_k^i) \quad \text{for all } \varphi^h \in \mathbb{V}_h, \tag{3.5}$$

where $y^h = S_h u$.

In what follows, we shall introduce FE discretizations for the scalar problems ($\hat{\mathbf{P}}_\alpha$) and ($\hat{\mathbf{P}}_\zeta$), introduced in Sections 2.3.1 and 2.3.2, respectively.

### 3.1. Discretization of WSM

For weights $\alpha_1, \alpha_2 > 0$ with $\alpha_1 + \alpha_2 = 1$, we set $\alpha := (\alpha_1, \alpha_2)$ and define the function

$$W_\alpha^h(u) := \alpha_1 j_1^h(u) + \alpha_2 j_2^h(u) \quad \text{for } u \in \mathbb{U}.$$

The proposed discrete version of the scalar-valued problem ($\hat{\mathbf{P}}_\alpha$) thus reads

$$\min W_\alpha^h(u^h) \quad \text{s.t.} \quad u^h \in \mathbb{U}_{\mathsf{ad}}^h. \tag{$\hat{\mathbf{P}}_\alpha^h$}$$

As in the continuous case, we obtain a characterization for the derivative of $W_\alpha^h$:

$$(W_\alpha^h)'(u^h) w^h = \sum_{k=1}^{2} \alpha_k (p_k^h + \lambda_k u^h, w^h)_{\mathbb{U}} \quad \text{for all } w^h \in \mathbb{U}_h.$$

We establish the existence of solutions for ($\hat{\mathbf{P}}_\alpha^h$) and first-order optimality conditions.

**Proposition 3.1** (Discrete WSM)**.** *Let $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2_>$ such that $\alpha_1 + \alpha_2 = 1$. Then, the scalar-valued optimal control problem ($\hat{\mathbf{P}}^h_\alpha$) has a unique solution $\bar{u}^h_\alpha \in \mathbb{U}^h_{\mathsf{ad}}$. Moreover, $\bar{u}^h_\alpha$ is an optimal solution to ($\hat{\mathbf{P}}^h_\alpha$) if and only if*

$$\sum_{k=1}^{2} \alpha_k(\bar{p}^h_{\alpha,k} + \lambda_k \bar{u}^h_\alpha, u^h - \bar{u}^h_\alpha)_{\mathbb{U}} \geq 0 \quad \textit{for all } u^h \in \mathbb{U}^h_{\mathsf{ad}}, \tag{3.6}$$

*where $\bar{p}^h_{\alpha,k}$ denotes the unique solution to (3.5) with $\bar{y}^h_\alpha = S_h \bar{u}^h_\alpha$.*

*Proof.* Since $j^h_1$ and $j^h_2$ are continuous and $\mathbb{U}^h_{\mathsf{ad}}$ is compact, Weierstraß theorem immediately yields the existence of a solution. Uniqueness follows from the strict convexity of $j^h_1$ and $j^h_2$. The variational inequality (3.6) follows as in the continuous case. □

We describe the algorithm proposed to solve the problem ($\hat{\mathbf{P}}^h_\alpha$), which is based on a Barzilai-Borwein gradient method [29]; see also [30], Algorithm 1.

---

**Algorithm 1:** Solving problem ($\hat{\mathbf{P}}^h_\alpha$) with $\alpha$ fixed

---

**Require:** Problem data, $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2_>$, $tol = 1$, and $u^{h,-1}, u^{h,0} \in \mathbb{U}^h_{\mathsf{ad}}$ with $u^{h,-1} \neq u^{h,0}$;
Set $l = 0$;
**while** $tol > 10^{-8}$ **do**
  Obtain $y^{h,l} \in \mathbb{V}_h$ by solving (3.3) with $u = u^{h,l}$;
  Obtain $p^{h,l}_1, p^{h,l}_2 \in \mathbb{V}_h$ by solving (3.5) with $y^h = y^{h,l}$;
  Compute step size

$$t_l = \frac{((W^h_\alpha)'(u^{h,l}) - (W^h_\alpha)'(u^{h,l-1}), (W^h_\alpha)'(u^{h,l}) - (W^h_\alpha)'(u^{h,l-1}))_{\mathbb{U}}}{((W^h_\alpha)'(u^{h,l}) - (W_{\alpha,h})'(u^{h,l-1}), u^{h,l} - u^{h,l-1})_{\mathbb{U}}};$$

  Set $u^{h,l+1} = \min\{u_{\mathsf{b}}, \max\{u_{\mathsf{a}}, u^{h,l} - {}^1\!/_{t_l}(W^h_\alpha)'(u^{h,l})\}\}$ and $u_{\text{ref}} = \min\{u_{\mathsf{b}}, \max\{u_{\mathsf{a}}, u^{h,l} - (W^h_\alpha)'(u^{h,l})\}\}$ ;
  Set $tol = \|u^{h,l+1} - u_{\text{ref}}\|_{\mathbb{U}}$ and $l = l + 1$;
**end while**
**return** Approximative optimal solution $u^{h,l}$ and its objective value $j^h(u^{h,l})$.

---

In Algorithm 2, we show how we compute the discrete approximation of the set of Pareto points and the corresponding front.

---

**Algorithm 2:** Weighted-sum method

---

**Require:** Number $\ell_{\max} \in \mathbb{N}$ of stationary points, problem data, and $0 < \varepsilon \ll 1$:
Set $\mathcal{P}^h_s = \mathcal{P}^h_f = \emptyset$;
**for** $\ell = 1, \ldots, \ell_{\mathsf{max}}$ **do**
  Set $\alpha_2 = \varepsilon + {}^{(\ell-1)}\!/_{(\ell_{\mathsf{max}}-1)}$;
  Solve problem ($\hat{\mathbf{P}}^h_\alpha$) using Algorithm 1 with weight $\alpha(\ell) = (1 - \alpha_2, \alpha_2)$, save solution $\bar{u}^h_{\alpha(\ell)}$ and its objective value $j^h(\bar{u}^h_{\alpha(\ell)})$;
  Set $\mathcal{P}^h_s = \mathcal{P}^h_s \cup \{\bar{u}^h_{\alpha(\ell)}\}$ and $\mathcal{P}^h_f = \mathcal{P}^h_f \cup \{j^h(\bar{u}^h_{\alpha(\ell)})\}$;
**end for**
**return** Discrete approximations $\mathcal{P}^h_s$ and $\mathcal{P}^h_f$ of Pareto stationary points and front, respectively.

---

*3.1.1. Error estimates for WSM*

Let $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2_>$ be fixed with $\alpha_1 + \alpha_2 = 1$. In this section, we prove convergence rates for $\|\bar{u}_\alpha - \bar{u}_\alpha^h\|_{\mathbb{U}}$, where $\bar{u}_\alpha$ and $\bar{u}_\alpha^h$ denote the unique solutions to ($\hat{\mathbf{P}}_\alpha$) and ($\hat{\mathbf{P}}_\alpha^h$), respectively.

We start by introducing the discrete auxiliary variable $\hat{y}^h \in \mathbb{V}_h$ as the unique solution to

$$(\hat{y}^h, \varphi^h)_V = (\bar{u}_\alpha, \varphi^h)_{\mathbb{U}} \quad \text{for all } \varphi^h \in \mathbb{V}_h. \tag{3.7}$$

We also define, for $k \in \{1,2\}$, the variable $\hat{p}_k^h \in \mathbb{V}_h$ solution to

$$(\varphi^h, \hat{p}_k^h)_V = \sum_{i=1}^{n_k} \left(\hat{y}^h(\boldsymbol{x}_k^i) - y_k^i\right) \varphi^h(\boldsymbol{x}_k^i) \quad \text{for all } \varphi^h \in \mathbb{V}_h. \tag{3.8}$$

**Proposition 3.2** (Auxiliary estimate)**.** *Let $\bar{p}_{\alpha,k}$, with $k \in \{1,2\}$, be the unique solution to (2.4) associated with the optimal solution $\bar{u}_\alpha$ to ($\hat{\mathbf{P}}_\alpha$), and let $\hat{p}_k^h$, with $k \in \{1,2\}$, be the unique solution (3.8). Assume that $u_{\mathsf{a}}, u_{\mathsf{b}}$ are real constants. Then, there exists an $h_* > 0$ such that*

$$\sum_{k=1}^{2} \|\bar{p}_{\alpha,k} - \hat{p}_k^h\|_{L^1(\Omega)} \lesssim h^2 |\log h|^2 \quad \textit{for all} \ \ 0 < h < h_*.$$

*Proof.* For simplicity, we only prove the estimate for the term $\|\bar{p}_{\alpha,1} - \hat{p}_1^h\|_{L^1(\Omega)}$. The case $k = 2$ follows by similar arguments. Let $\tilde{p}^h \in \mathbb{V}_h$ be the unique solution to

$$(\varphi^h, \tilde{p}^h)_V = \sum_{i=1}^{n_1} \left(\bar{y}_\alpha(\boldsymbol{x}_1^i) - y_1^i\right) \varphi^h(\boldsymbol{x}_1^i) \quad \text{for all } \varphi^h \in \mathbb{V}_h. \tag{3.9}$$

We note that $\tilde{p}^h$ corresponds to the FE approximation of $\bar{p}_{\alpha,1}$ (solution to (2.4) with $k = 1$) in $\mathbb{V}_h$. An application of the triangle inequality and [9], Lemma 5.3 results in

$$\|\bar{p}_{\alpha,1} - \hat{p}_1^h\|_{L^1(\Omega)} \le \|\bar{p}_{\alpha,1} - \tilde{p}^h\|_{L^1(\Omega)} + \|\tilde{p}^h - \hat{p}_1^h\|_{L^1(\Omega)} \lesssim h^2|\log h|^2 + \|\tilde{p}^h - \hat{p}_1^h\|_{L^1(\Omega)}.$$

To control $\|\tilde{p}^h - \hat{p}_1^h\|_{L^1(\Omega)}$, we introduce the variable $\chi \in W_0^{1,\mathsf{p}}(\Omega)$ with $\mathsf{p} < d/d{-}1$ as the unique solution to

$$(\varphi, \chi)_V = \sum_{i=1}^{n_1} \left(\bar{y}_\alpha(\boldsymbol{x}_1^i) - \hat{y}^h(\boldsymbol{x}_1^i)\right) \varphi(\boldsymbol{x}_1^i) \quad \text{for all } \varphi \in W_0^{1,\mathsf{q}}(\Omega).$$

We note that $\tilde{p}^h - \hat{p}_1^h$ is the unique FE approximation of $\chi$ in $\mathbb{V}_h$. Therefore, an application of [9], Lemma 5.3, the embedding $W_0^{1,\mathsf{p}}(\Omega) \hookrightarrow L^1(\Omega)$, and the stability estimate $\|\nabla\chi\|_{L^{\mathsf{p}}(\Omega;\mathbb{R}^d)} \lesssim \sum_{i=1}^{n_1} |\bar{y}_\alpha(\boldsymbol{x}_1^i) - \hat{y}^h(\boldsymbol{x}_1^i)|$ (*cf.* [23], Thm. 1) yield

$$\begin{aligned} \|\tilde{p}^h - \hat{p}_1^h\|_{L^1(\Omega)} &\le \|(\tilde{p}^h - \hat{p}_1^h) - \chi\|_{L^1(\Omega)} + \|\chi\|_{L^1(\Omega)} \\ &\lesssim h^2|\log h|^2 + \|\nabla\chi\|_{L^{\mathsf{p}}(\Omega;\mathbb{R}^d)} \lesssim h^2|\log h|^2 + \sum_{i=1}^{n_1} |\bar{y}_\alpha(\boldsymbol{x}_1^i) - \hat{y}^h(\boldsymbol{x}_1^i)|. \end{aligned} \tag{3.10}$$

To estimate $\sum_{i=1}^{n_1} |\bar{y}_\alpha(\boldsymbol{x}_1^i) - \hat{y}^h(\boldsymbol{x}_1^i)|$ in (3.10), we proceed as follows. Since $\Omega_1$ is finite, there exist sets $\omega_1, \widehat{\omega}_1$ such that $\Omega_1 \subset \omega_1 \Subset \widehat{\omega}_1 \Subset \Omega$ with $\widehat{\omega}_1$ smooth. At the same time, since $u_{\mathsf{a}}, u_{\mathsf{b}} \in \mathbb{R}$, we have in particular $\bar{u}_\alpha \in L^\infty(\Omega)$.

These ingredients, in combination with the fact that $\hat{y}^h$ corresponds to the unique FE approximation of $\bar{y}_\alpha$ in $\mathbb{V}_h$, allow us to use [9], Lemma 4.4 (i) to obtain

$$\sum_{i=1}^{n_1} |\bar{y}_\alpha(\boldsymbol{x}_1^i) - \hat{y}^h(\boldsymbol{x}_1^i)| \leq n_1 \|\bar{y}_\alpha - \hat{y}^h\|_{L^\infty(\omega_1)} \lesssim h^2 |\log h|^2, \qquad 0 < h < h_*.$$

This concludes the proof. $\square$

With these ingredients at hand, we present the following error estimate.

**Theorem 3.3** (Error estimate: WSM)**.** *Let $\bar{u}_\alpha$ and $\bar{u}_\alpha^h$ be the unique solutions to ($\hat{\mathbf{P}}_\alpha$) and ($\hat{\mathbf{P}}_\alpha^h$), respectively. Assume that $u_{\mathsf{a}}, u_{\mathsf{b}}$ are real constants. Then, there exists $h_* > 0$ such that*

$$\left(\sum_{k=1}^{2} \alpha_k \lambda_k\right)^{\frac{1}{2}} \|\bar{u}_\alpha - \bar{u}_\alpha^h\|_{\mathbb{U}} \lesssim h|\log h| \quad \textit{for all } 0 < h < h_*.$$

*Proof.* Let $\pi_0 : \mathbb{U} \to \mathbb{U}_h$ be the $\mathbb{U}$-orthogonal projection operator. We consider $u = \bar{u}_\alpha^h$ in (2.8) and $u^h = \pi_0 \bar{u}_\alpha$ in (3.6), and add the obtained inequalities. This results in

$$\begin{aligned}\sum_{k=1}^{2} \alpha_k \lambda_k \|\bar{u}_\alpha - \bar{u}_\alpha^h\|_{\mathbb{U}}^2 &\leq \sum_{k=1}^{2} \alpha_k (\bar{p}_{\alpha,k} - \bar{p}_{\alpha,k}^h, \bar{u}_\alpha^h - \bar{u}_\alpha)_{\mathbb{U}} + \sum_{k=1}^{2} \alpha_k (\bar{p}_{\alpha,k}^h + \lambda_k \bar{u}_\alpha^h, \pi_0 \bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}} \\ &=: \mathsf{I} + \mathsf{II}.\end{aligned} \tag{3.11}$$

Let us bound

$$\mathsf{I} = \alpha_1 (\bar{p}_{\alpha,1} - \bar{p}_{\alpha,1}^h, \bar{u}_\alpha^h - \bar{u}_\alpha)_{\mathbb{U}} + \alpha_2 (\bar{p}_{\alpha,2} - \bar{p}_{\alpha,2}^h, \bar{u}_\alpha^h - \bar{u}_\alpha)_{\mathbb{U}} =: \mathsf{I}_1 + \mathsf{I}_2. \tag{3.12}$$

For clarity, we concentrate only on $\mathsf{I}_1$, since $\mathsf{I}_2$ can be treated analogously. We first utilize the auxiliary variable $\hat{p}_1^h \in \mathbb{V}_h$, introduced in (3.8), and write

$$\mathsf{I}_1 = \alpha_1 (\hat{p}_1^h - \bar{p}_{\alpha,1}^h, \bar{u}_\alpha^h - \bar{u}_\alpha)_{\mathbb{U}} + \alpha_1 (\bar{p}_{\alpha,1} - \hat{p}_1^h, \bar{u}_\alpha^h - \bar{u}_\alpha)_{\mathbb{U}}.$$

Recall that $\hat{p}_1^h - \bar{p}_{\alpha,1}^h \in \mathbb{V}_h$ and $\bar{y}_\alpha^h - \hat{y}^h \in \mathbb{V}_h$ solve

$$\begin{aligned}(\varphi^h, \hat{p}_1^h - \bar{p}_{\alpha,1}^h)_V &= \sum_{i=1}^{n_1} \left(\hat{y}^h(\boldsymbol{x}_1^i) - \bar{y}_\alpha^h(\boldsymbol{x}_1^i)\right) \varphi^h(\boldsymbol{x}_1^i) \qquad && \text{for all } \varphi^h \in \mathbb{V}_h, \\ (\bar{y}_\alpha^h - \hat{y}^h, \varphi^h)_V &= (\bar{u}_\alpha^h - \bar{u}_\alpha, \varphi^h)_{\mathbb{U}} && \text{for all } \varphi^h \in \mathbb{V}_h,\end{aligned}$$

respectively, with $\hat{y}^h \in \mathbb{V}_h$ defined in (3.7). Hence, choosing $\varphi^h = \bar{y}_\alpha^h - \hat{y}^h$ in the first equation and $\varphi^h = \hat{p}_1^h - \bar{p}_{\alpha,1}^h$ in the second equation, we infer that

$$\alpha_1 (\hat{p}_1^h - \bar{p}_{\alpha,1}^h, \bar{u}_\alpha^h - \bar{u}_\alpha)_{\mathbb{U}} = -\alpha_1 \sum_{i=1}^{n_1} \left(\hat{y}^h(\boldsymbol{x}_1^i) - \bar{y}_\alpha^h(\boldsymbol{x}_1^i)\right)^2 \leq 0. \tag{3.13}$$

Consequently, $\mathsf{I}_1 \leq \alpha_1(\bar{p}_{\alpha,1} - \hat{p}_1^h, \bar{u}_\alpha^h - \bar{u}_\alpha)_{\mathbb{U}}$. To control the latter, we apply Hölder's inequality, Proposition 3.2, and the fact that $\bar{u}_\alpha, \bar{u}_\alpha^h \in L^\infty(\Omega)$:

$$\mathsf{I}_1 \leq \alpha_1 \|\bar{p}_{\alpha,1} - \hat{p}_1^h\|_{L^1(\Omega)} \|\bar{u}_\alpha^h - \bar{u}_\alpha\|_{L^\infty(\Omega)} \lesssim h^2 |\log h|^2, \qquad 0 < h < h_*. \tag{3.14}$$

Therefore, in view of (3.14) and its analogous version for $\mathsf{I}_2$, (3.12), and (3.11), we conclude that

$$\left( \sum_{k=1}^{2} \alpha_k \lambda_k \right) \|\bar{u}_\alpha - \bar{u}_\alpha^h\|_{\mathbb{U}}^2 \leq C_1 h^2 |\log h|^2 + \mathsf{II} \tag{3.15}$$

for a constant $C_1 > 0$ and all $0 < h < h_*$. To estimate $\mathsf{II}$ we write

$$\mathsf{II} = \alpha_1(\bar{p}_{\alpha,1}^h + \lambda_1 \bar{u}_\alpha^h, \pi_0 \bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}} + \alpha_2(\bar{p}_{\alpha,2}^h + \lambda_2 \bar{u}_\alpha^h, \pi_0 \bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}} =: \mathsf{II}_1 + \mathsf{II}_2. \tag{3.16}$$

In what follows, we only estimate $\mathsf{II}_1$ since the estimation of $\mathsf{II}_2$ follows analogous steps. First, we note that

$$\begin{aligned} \mathsf{II}_1 = & \, \alpha_1(\bar{p}_{\alpha,1} + \lambda_1 \bar{u}_\alpha, \pi_0 \bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}} + \alpha_1(\hat{p}_1^h - \bar{p}_{\alpha,1}, \pi_0 \bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}} \\ & + \alpha_1(\bar{p}_{\alpha,1}^h - \hat{p}_1^h, \pi_0 \bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}} + \alpha_1 \lambda_1 (\bar{u}_\alpha^h - \bar{u}_\alpha, \pi_0 \bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}}. \end{aligned} \tag{3.17}$$

To control the term $\alpha_1(\bar{p}_{\alpha,1} + \lambda_1 \bar{u}_\alpha, \pi_0 \bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}}$ in (3.17), we use the orthogonality of $\pi_0$ to infer that

$$\begin{aligned} \alpha_1(\bar{p}_{\alpha,1} + \lambda_1 \bar{u}_\alpha, \pi_0 \bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}} &= \alpha_1((\bar{p}_{\alpha,1} + \lambda_1 \bar{u}_\alpha) - \pi_0(\bar{p}_{\alpha,1} + \lambda_1 \bar{u}_\alpha), \pi_0 \bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}} \\ &= \alpha_1(\bar{p}_{\alpha,1} - \pi_0 \bar{p}_{\alpha,1}, \pi_0 \bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}} - \lambda_1 \|\pi_0 \bar{u}_\alpha - \bar{u}_\alpha\|_{\mathbb{U}}^2 \\ &\leq \alpha_1(\bar{p}_{\alpha,1} - \pi_0 \bar{p}_{\alpha,1}, \pi_0 \bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}}. \end{aligned}$$

Then, in view of the $W^{1,\infty}(\Omega)$-regularity of $\bar{u}_\alpha$, which stems from (2.9) and [9], Lemma 6.1, and using that $\|\pi_0 \bar{p}_{\alpha,1} - \bar{p}_{\alpha,1}\|_{L^1(\Omega)} \lesssim h \|\nabla \bar{p}_{\alpha,1}\|_{L^1(\Omega;\mathbb{R}^d)}$ (*cf.* [28], Prop. 1.135) we obtain

$$\begin{aligned} \alpha_1(\bar{p}_{\alpha,1} + \lambda_1 \bar{u}_\alpha, \pi_0 \bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}} &\leq \alpha_1 \|\pi_0 \bar{p}_{\alpha,1} - \bar{p}_{\alpha,1}\|_{L^1(\Omega)} \|\pi_0 \bar{u}_\alpha - \bar{u}_\alpha\|_{L^\infty(\Omega)} \\ &\lesssim h^2 \|\nabla \bar{p}_{\alpha,1}\|_{L^1(\Omega;\mathbb{R}^d)} \|\nabla \bar{u}_\alpha\|_{L^\infty(\Omega;\mathbb{R}^d)} \lesssim h^2. \end{aligned}$$

The second term on the right-hand side of (3.17) is bounded in view of Proposition 3.2 and the $W^{1,\infty}(\Omega)$-regularity of $\bar{u}_\alpha$. These arguments yield

$$\alpha_1(\hat{p}_1^h - \bar{p}_{\alpha,1}, \pi_0 \bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}} \leq \frac{\alpha_1}{2} \|\hat{p}_1^h - \bar{p}_{\alpha,1}\|_{L^1(\Omega)}^2 + \frac{\alpha_1}{2} \|\pi_0 \bar{u}_\alpha - \bar{u}_\alpha\|_{L^\infty(\Omega)}^2 \lesssim h^4 |\log h|^4 + h^2 \lesssim h^2.$$

To estimate the third term on the right-hand side of (3.17), we use Hölder's inequality and similar arguments to those that lead to (3.10), to arrive at

$$\begin{aligned} \alpha_1(\bar{p}_{\alpha,1}^h - \hat{p}_1^h, \pi_0 \bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}} &\leq \alpha_1 \|\bar{p}_{\alpha,1}^h - \hat{p}_1^h\|_{L^1(\Omega)} \|\pi_0 \bar{u}_\alpha - \bar{u}_\alpha\|_{L^\infty(\Omega)} \\ &\lesssim \left( h^2 |\log h|^2 + \sum_{i=1}^{n_1} |\bar{y}_\alpha^h(\boldsymbol{x}_1^i) - \hat{y}^h(\boldsymbol{x}_1^i)| \right) \|\pi_0 \bar{u}_\alpha - \bar{u}_\alpha\|_{L^\infty(\Omega)}. \end{aligned} \tag{3.18}$$

We let $\phi \in \mathbb{Y}$ be the unique solution to $(\phi, \varphi)_V = (\bar{u}_\alpha^h - \bar{u}_\alpha, \varphi)_{\mathbb{U}}$ for all $\varphi \in V$, and immediately note that $\bar{y}_\alpha^h - \hat{y}^h$ corresponds to the unique FE approximation of $\phi$ in $\mathbb{V}_h$. On the other hand, since the set $\Omega_1 = \{\boldsymbol{x}_1^1, \ldots, \boldsymbol{x}_1^{n_1}\}$ is finite, there exist sets $\omega_1, \widehat{\omega}_1$ such that $\Omega_1 \subset \omega_1 \Subset \widehat{\omega}_1 \Subset \Omega$ with $\widehat{\omega}_1$ smooth. Hence, the application of [9], Lemma 4.4 (i) and the use of the stability estimate $\|\phi\|_{L^\infty(\omega_1)} \leq \|\phi\|_{L^\infty(\Omega)} \lesssim \|\bar{u}_\alpha^h - \bar{u}_\alpha\|_{\mathbb{U}}$ imply that for a potentially smaller $h_* > 0$

$$\sum_{i=1}^{n_1} |\bar{y}_\alpha(\boldsymbol{x}_1^i) - \hat{y}^h(\boldsymbol{x}_1^i)| \lesssim \|(\bar{y}_\alpha - \hat{y}^h) - \psi\|_{L^\infty(\omega_1)} + \|\psi\|_{L^\infty(\Omega)} \lesssim h^2|\log h|^2 + \|\bar{u}_\alpha^h - \bar{u}_\alpha\|_{\mathbb{U}}, \qquad 0 < h < h_*.$$

Using this estimate, the $W^{1,\infty}(\Omega)$-regularity of $\bar{u}_\alpha$, and Young's inequality in (3.18), we conclude that

$$\alpha_1(\bar{p}_{\alpha,1}^h - \hat{p}_1^h, \pi_0\bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}} \lesssim h^3|\log h|^2 + h\|\bar{u}_\alpha^h - \bar{u}_\alpha\|_{\mathbb{U}} \leq C_2 h^3|\log h|^2 + C_3 h^2 + \frac{\alpha_1\lambda_1}{4}\|\bar{u}_\alpha^h - \bar{u}_\alpha\|_{\mathbb{U}}^2$$

for constants $C_2, C_3 > 0$. Finally, the term $\alpha_1\lambda_1(\bar{u}_\alpha^h - \bar{u}_\alpha, \pi_0\bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}}$ in (3.17) is bounded using Cauchy-Schwarz inequality and estimate [28], Proposition 1.135 as follows

$$\begin{aligned}\alpha_1\lambda_1(\bar{u}_\alpha^h - \bar{u}_\alpha, \pi_0\bar{u}_\alpha - \bar{u}_\alpha)_{\mathbb{U}} &\leq \alpha_1\lambda_1\|\pi_0\bar{u}_\alpha - \bar{u}_\alpha\|_{\mathbb{U}}^2 + \frac{\alpha_1\lambda_1}{4}\|\bar{u}_\alpha^h - \bar{u}_\alpha\|_{\mathbb{U}}^2 \\ &\leq C_4 h^2 + \frac{\alpha_1\lambda_1}{4}\|\bar{u}_\alpha^h - \bar{u}_\alpha\|_{\mathbb{U}}^2\end{aligned}$$

for a constant $C_4 > 0$. Therefore, the combination of all the estimates derived for the terms on the right-hand side of (3.17) gives as a result that

$$\mathsf{II}_1 \leq Ch^2 + \frac{\alpha_1\lambda_1}{2}\|\bar{u}_\alpha^h - \bar{u}_\alpha\|_{\mathbb{U}}^2 \quad (C > 0).$$

Similarly, we have that $\mathsf{II}_2 \leq Ch^2 + {}^{\alpha_2\lambda_2}/{}_2\|\bar{u}_\alpha^h - \bar{u}_\alpha\|_{\mathbb{U}}^2$. We conclude the desired error estimate using the estimates obtained for $\mathsf{II}_1$ and $\mathsf{II}_2$ in (3.16), and using the resulting estimate in (3.15). □

We now improve the error estimate of Theorem 3.3 in two dimensions. Its proof utilizes an FE estimate for the specific case $d = 2$; *cf.*, [31, 32] and [23].

**Theorem 3.4** (Improved error estimate: WSM)**.** *Let $d = 2$. In the framework of Theorem 3.3, we have the following optimal error estimate:*

$$\left(\sum_{k=1}^{2}\alpha_k\lambda_k\right)^{\frac{1}{2}}\|\bar{u}_\alpha - \bar{u}_\alpha^h\|_{\mathbb{U}} \lesssim h \quad \textit{for all } 0 < h < h_*.$$

*Proof.* The proof relies on a more careful estimate for $\mathsf{I}$ in (3.12). To control this term, we concentrate, again, only on $\mathsf{I}_1$. Writing $\mathsf{I}_1 = \alpha_1(\hat{p}_1^h - \bar{p}_{\alpha,1}^h, \bar{u}_\alpha^h - \bar{u}_\alpha)_{\mathbb{U}} + \alpha_1(\bar{p}_{\alpha,1} - \hat{p}_1^h, \bar{u}_\alpha^h - \bar{u}_\alpha)_{\mathbb{U}}$, using the estimate (3.13) and Cauchy-Schwarz inequality, we obtain

$$\mathsf{I}_1 \leq \alpha_1(\bar{p}_{\alpha,1} - \hat{p}_1^h, \bar{u}_\alpha^h - \bar{u}_\alpha)_{\mathbb{U}} \leq \alpha_1\|\bar{p}_{\alpha,1} - \hat{p}_1^h\|_{\mathbb{U}}\|\bar{u}_\alpha^h - \bar{u}_\alpha\|_{\mathbb{U}}.$$

On the other hand, the use of arguments similar to those that lead to the result of Proposition 3.2 in combination with the error estimate from [23], Theorem 3 gives

$$\|\bar{p}_{\alpha,1} - \hat{p}_1^h\|_{\mathbb{U}} \lesssim h, \tag{3.19}$$

upon using that $d = 2$. Hence, Young's inequality and the error estimate (3.19) allow us to conclude that

$$\mathsf{I}_1 \leq \alpha_1\lambda_1\lambda_1^{-1}\|\bar{p}_{\alpha,1} - \hat{p}_1^h\|_{\mathbb{U}}\|\bar{u}_\alpha^h - \bar{u}_\alpha\|_{\mathbb{U}} \leq Ch^2 + \frac{\alpha_1\lambda_1}{4}\|\bar{u}_\alpha^h - \bar{u}_\alpha\|_{\mathbb{U}}^2, \quad (C > 0).$$

Using, this estimate and its analogous version for $\mathsf{I}_2$, (3.12), and (3.11), we obtain

$$\left(\frac{3}{4}\sum_{k=1}^{2}\alpha_k\lambda_k\right)\|\bar{u}_\alpha - \bar{u}_\alpha^h\|_{\mathbb{U}}^2 \leq Ch^2 + \mathsf{II}.$$

The term $\mathsf{II}$ is bounded as in the proof of Theorem 3.3. This concludes the proof. □

### 3.2. Discretization of RPM

Given $\zeta = (\zeta_1, \zeta_2) \in \mathcal{P}_f + \mathbb{R}^2_{\leq}$, we introduce the function $R_\zeta^h(u) := {}^1\!/_2\left((j_1^h(u) - \zeta_1)^2 + (j_2^h(u) - \zeta_2)^2\right)$ for every $u \in \mathbb{U}$. The discrete version of the scalar-valued problem ($\hat{\mathbf{P}}_\zeta$) then reads

$$\min R_\zeta^h(u^h) \quad \text{s.t.} \quad u^h \in \mathbb{U}_{\mathsf{ad}}^h. \qquad (\hat{\mathbf{P}}_\zeta^h)$$

We have the following characterization for $(R_\zeta^h)'(u^h)$

$$(R_\zeta^h)'(u^h)w^h = \sum_{k=1}^{2}(j_k^h(u^h) - \zeta_k)(p_k^h + \lambda_k u^h, w^h)_{\mathbb{U}} \quad \text{for all } w^h \in \mathbb{U}_h.$$

With this characterization at hand, we present the following result, the proof of which follows as in Proposition 3.1.

**Proposition 3.5** (Discrete RPM)**.** *Let $\zeta = (\zeta_1, \zeta_2) \in \mathcal{P}_f + \mathbb{R}^2_{\leq}$. Then, the scalar-valued optimal control problem ($\hat{\mathbf{P}}_\zeta^h$) has a unique solution $\bar{u}_\zeta^h \in \mathbb{U}_{\mathsf{ad}}^h$. Moreover, $\bar{u}_\zeta^h$ is an optimal solution to ($\hat{\mathbf{P}}_\zeta^h$) if and only if*

$$\sum_{k=1}^{2}(j_k^h(\bar{u}_\zeta^h) - \zeta_k)(\bar{p}_{\zeta,k}^h + \lambda_k\bar{u}_\zeta^h, u^h - \bar{u}_\zeta^h)_{\mathbb{U}} \geq 0 \quad \textit{for all } u^h \in \mathbb{U}_{\mathsf{ad}}^h, \qquad (3.20)$$

*where $\bar{p}_{\zeta,k}^h$ denotes the unique solution to (3.5) with $\bar{y}_\zeta^h = S_h\bar{u}_\zeta^h$.*

Given $\zeta$ fixed, we use Algorithm 3 to solve problem ($\hat{\mathbf{P}}_\zeta^h$). This algorithm is also based on a Barzilai-Borwein gradient method.

**Algorithm 3:** Solving problem ($\hat{\mathbf{P}}^h_\zeta$) with $\zeta$ fixed

**Require:** Problem data, $\zeta = (\zeta_1, \zeta_2)$ fixed, $tol = 1$, and $u^{h,-1}, u^{h,0} \in \mathbb{U}^h_{\mathsf{ad}}$ with $u^{h,-1} \neq u^{h,0}$;

Set $l = 0$;
**while** $tol > 10^{-8}$ **do**
  Obtain $y^{h,l} \in \mathbb{V}_h$ by solving (3.3) with $u = u^{h,l}$;
  Obtain $p_1^{h,l}, p_2^{h,l} \in \mathbb{V}_h$ by solving (3.5) with $y^h = y^{h,l}$;
  Compute step size

$$t_l = \frac{((R^h_\zeta)'(u^{h,l}) - (R^h_\zeta)'(u^{h,l-1}), (R^h_\zeta)'(u^{h,l}) - (R^h_\zeta)'(u^{h,l-1}))_{\mathbb{U}}}{((R^h_\zeta)'(u^{h,l}) - (R^h_\zeta)'(u^{h,l-1}), u^{h,l} - u^{h,l-1})_{\mathbb{U}}};$$

  Set $u^{h,l+1} = \min\{u_{\mathsf{b}}, \max\{u_{\mathsf{a}}, u^{h,l} - {}^1/_{t_l}(R^h_\zeta)'(u^{h,l})\}\}$ and
  $u_{\mathrm{ref}} = \min\{u_{\mathsf{b}}, \max\{u_{\mathsf{a}}, u^{h,l} - (R^h_\zeta)'(u^{h,l})\}\}$;
  Set $tol = \|u^{h,l+1} - u_{\mathrm{ref}}\|_{\mathbb{U}}$ and $l = l + 1$;
**end while**
**return** Approximative optimal solution $u^{h,l}$ and its objective value $j^h(u^{h,l})$.

In Algorithm 2 we show how we compute the discrete approximation of the set of Pareto points and the corresponding front.

**Algorithm 4:** Reference point method

**Require:** Number $\ell_{\max} \in \mathbb{N}$ of stationary points, problem data, parameters $\eta^\perp, \eta^\parallel > 0$, and parameter $\varepsilon \ll 1$:

Compute $\bar{u}^{h,\mathrm{init}}_\alpha$ solution to ($\hat{\mathbf{P}}^h_\alpha$) with $\alpha = (1 - \varepsilon, \varepsilon)$ (initial point);
Compute $\bar{u}^{h,\mathrm{end}}_\alpha$ solution to ($\hat{\mathbf{P}}^h_\alpha$) with $\alpha = (\varepsilon, 1 - \varepsilon)$ (ending point);
Set $\mathcal{P}^h_s = \{\bar{u}^{h,\mathrm{init}}_\alpha\} \cup \{\bar{u}^{h,\mathrm{end}}_\alpha\}$, $\mathcal{P}^h_f = \{j^h(\bar{u}^{h,\mathrm{init}}_\alpha)\} \cup \{j^h(\bar{u}^{h,\mathrm{end}}_\alpha)\}$ and $\ell = 1$;
**while** $\zeta_1^\ell < j_1^h(\bar{u}^{h,\mathrm{end}}_\alpha)$ and $\ell \leq \ell_{\max} - 1$ **do**
  Solve problem ($\hat{\mathbf{P}}^h_\zeta$) using Algorithm 3 with reference point $\zeta^\ell$, save the solution $\bar{u}^{h,\ell}_\zeta$ and its evaluation $j^h(\bar{u}^{h,\ell}_\zeta)$;
  Set $\mathcal{P}^h_s = \mathcal{P}^h_s \cup \{\bar{u}^{h,\ell}_\zeta\}$, $\mathcal{P}^h_f = \mathcal{P}^h_f \cup \{j^h(\bar{u}^{h,\ell}_\zeta)\}$, and $\ell = \ell + 1$;
**end while**
**return** Discrete approximations $\mathcal{P}^h_s$ and $\mathcal{P}^h_f$ of Pareto stationary points and front.

An important question to tackle when using the RPM is how to choose the reference points $\zeta$ in the numerical implementation. We perform this following the approach of [4], Section 3.2. To make matter precise, let $\ell_{\max}$ be the maximal number of Pareto stationary points in the numerical implementation and let $u^{h,0}$ denote an initial starting point being a stationary point of the weighted-sum problem with weights $\alpha_1 = 1 - \alpha_2$ and $\alpha_2 = \varepsilon \ll 1$. Then, the first reference point $\zeta^1$ (corresponding to the second point on the front) is chosen as

$$\zeta^1 = j^h(u^{h,0})^\intercal - (\mathfrak{h}^\perp, \mathfrak{h}^\parallel), \tag{3.21}$$

where $\mathfrak{h}^\perp, \mathfrak{h}^\parallel > 0$ are suitable scaling parameters. Then, for $\ell = 1, \ldots, \ell_{\max} - 2$ we take

$$\zeta^{\ell+1} = j^h(u^{h,\ell})^\intercal + \mathfrak{h}^\parallel \frac{\eta^\parallel}{\|\eta^\parallel\|_{\mathbb{R}^2}} + \mathfrak{h}^\perp \frac{\eta^\perp}{\|\eta^\perp\|_{\mathbb{R}^2}}, \tag{3.22}$$

with $\eta^\perp = \zeta^\ell - j^h(u^{h,\ell})$ and $\eta^\parallel = (-\eta_2^\perp, \eta_1^\perp)$. Note that due to the strong weighting of $j_1^h(u^{h,0})$, the Pareto front is approximately vertical in the area of the first reference point, which motivates the initial choice $\eta^\parallel = (0,-1)$ and $\eta^\perp = (-1,0)$ in (3.21). Using this update technique, we end up with the reference point method stated in Algorithm 4.

### *3.2.1. Error estimates for RPM*

Let $\zeta = (\zeta_1, \zeta_2) \in \mathcal{P}_f + \mathbb{R}^2_\leq$ be fixed. In this section, we prove convergence rates for $\|\bar{u}_\zeta - \bar{u}_\zeta^h\|_{\mathbb{U}}$, where $\bar{u}_\zeta$ and $\bar{u}_\zeta^h$ denote the unique solutions to ($\hat{\mathbf{P}}_\zeta$) and ($\hat{\mathbf{P}}_\zeta^h$), respectively.

**Lemma 3.6** (Auxiliary estimate)**.** *Let $v, w \in \mathbb{U}_{\mathsf{ad}}$. Assume that $u_{\mathsf{a}}, u_{\mathsf{b}}$ are real constants. Then,*

$$\sum_{k=1}^{2} |j_k(v) - j_k^h(w)| \lesssim h^2 |\log h|^2 + \|v - w\|_{\mathbb{U}} \quad \textit{for all } h < h_*.$$

*Proof.* We begin by estimating the term $|j_1(v) - j_1^h(w)|$. Let $y_v, y_w \in V$ be the unique solutions to

$$(y_v, \varphi)_V = (v, \varphi)_{\mathbb{U}} \quad \text{for all } \varphi \in V,$$
$$(y_w, \varphi)_V = (w, \varphi)_{\mathbb{U}} \quad \text{for all } \varphi \in V,$$

and let $y_v^h, y_w^h \in \mathbb{V}_h$ be their corresponding FE approximations, *i.e.*,

$$(y_v^h, \varphi^h)_V = (v, \varphi^h)_{\mathbb{U}} \quad \text{for all } \varphi^h \in \mathbb{V}_h,$$
$$(y_w^h, \varphi^h)_V = (w, \varphi^h)_{\mathbb{U}} \quad \text{for all } \varphi^h \in \mathbb{V}_h.$$

Then, we have that

$$|j_1(v) - j_1^h(w)| \leq \frac{1}{2} \sum_{i=1}^{n_1} |y_v(\boldsymbol{x}_1^i) - y_w^h(\boldsymbol{x}_1^i)||y_v(\boldsymbol{x}_1^i) + y_w^h(\boldsymbol{x}_1^i) - 2y_1^i| + \frac{\lambda_1}{2}|(v - w, v + w)_{\mathbb{U}}|.$$

Using the fact that $v, w \in \mathbb{U}_{\mathsf{ad}}$ and that (see [9], Lem. 4.4 (i))

$$|y_w^h(\boldsymbol{x}_1^i)| \leq |y_w^h(\boldsymbol{x}_1^i) - y_w(\boldsymbol{x}_1^i)| + \|y_w\|_{L^\infty(\Omega)} \lesssim h^2 |\log h|^2 + \|w\|_{\mathbb{U}} \leq C, \quad i \in \{1, \ldots, n_1\},$$

we infer the estimate

$$|j_1(v) - j_1^h(w)| \lesssim \sum_{i=1}^{n_1} |y_v(\boldsymbol{x}_1^i) - y_w^h(\boldsymbol{x}_1^i)| + \|v - w\|_{\mathbb{U}}, \tag{3.23}$$

with a hidden constant that is independent of the discretization parameter. The term $\sum_{i=1}^{n_1} |y_v(\boldsymbol{x}_1^i) - y_w^h(\boldsymbol{x}_1^i)|$ in (3.23) is controlled using $y_w$, the stability estimate $\|y_v - y_w\|_{L^\infty(\Omega)} \lesssim \|v - w\|_{\mathbb{U}}$, and [9], Lemma 4.4 (i):

$$\sum_{i=1}^{n_1} |y_v(\boldsymbol{x}_1^i) - y_w^h(\boldsymbol{x}_1^i)| \leq n_1 \|y_v - y_w\|_{L^\infty(\Omega)} + \sum_{i=1}^{n_1} |y_w(\boldsymbol{x}_1^i) - y_w^h(\boldsymbol{x}_1^i)| \lesssim \|v - w\|_{\mathbb{U}} + h^2 |\log h|^2.$$

Using this estimate in (3.23) we obtain $|j_1(v) - j_1^h(w)| \lesssim h^2 |\log h|^2 + \|v - w\|_{\mathbb{U}}$. The estimation of $|j_2(v) - j_2^h(w)|$ follows analogous arguments. This concludes the proof. $\square$

Before presenting the next result, we introduce $\tilde{y}^h \in \mathbb{V}_h$ as the unique solution to

$$(\tilde{y}^h, \varphi^h)_V = (\bar{u}_\zeta, \varphi^h)_{\mathbb{U}} \quad \text{for all } \varphi^h \in \mathbb{V}_h, \tag{3.24}$$

and, for $k \in \{1,2\}$, the auxiliary variable $\tilde{p}_k^h \in \mathbb{V}_h$, solution to

$$(\varphi^h, \tilde{p}_k^h)_V = \sum_{i=1}^{n_k} \left(\tilde{y}^h(\boldsymbol{x}_k^i) - y_k^i\right) \varphi^h(\boldsymbol{x}_k^i) \quad \text{for all } \varphi^h \in \mathbb{V}_h. \tag{3.25}$$

The following error estimate is based on first-order optimality (*cf.*, *e.g.*, [33]) and FE error estimation.

**Theorem 3.7** (Error estimate: RPM)**.** *Let $\bar{u}_\zeta$ and $\bar{u}_\zeta^h$ be the unique solutions to ($\hat{\mathbf{P}}_\zeta$) and ($\hat{\mathbf{P}}_\zeta^h$), respectively. Assume that $u_{\mathsf{a}}, u_{\mathsf{b}} \in \mathbb{R}$. Then, we have the error estimate*

$$\sum_{k=1}^{2} \lambda_k \left( \frac{3 j_k(\bar{u}_\zeta) + j_k^h(\bar{u}_\zeta^h)}{2} - 2\zeta_k \right) \|\bar{u}_\zeta - \bar{u}_\zeta^h\|_{\mathbb{U}}^2 \lesssim h^2 |\log h|^2 \quad \textit{for all } h < h_*.$$

*In particular, if $h < h_*$ is small enough, then we have $\|\bar{u}_\zeta - \bar{u}_\zeta^h\|_{\mathbb{U}} \lesssim h|\log h|$.*

*Proof.* Let $\pi_0 : \mathbb{U} \to \mathbb{U}_h$ be the $\mathbb{U}$-orthogonal projection operator. We consider $u = \bar{u}_\zeta^h$ in (2.10) and $u^h = \pi_0 \bar{u}_\zeta$ in (3.20), and add the obtained inequalities. This results in

$$\begin{aligned} &\sum_{k=1}^{2} (j_k(\bar{u}_\zeta) - \zeta_k)(\bar{p}_{\zeta,k} + \lambda_k \bar{u}_\zeta, \bar{u}_\zeta^h - \bar{u}_\zeta)_{\mathbb{U}} + \sum_{k=1}^{2} (j_k^h(\bar{u}_\zeta^h) - \zeta_k)(\bar{p}_{\zeta,k}^h + \lambda_k \bar{u}_\zeta^h, \bar{u}_\zeta - \bar{u}_\zeta^h)_{\mathbb{U}} \\ &\quad + \sum_{k=1}^{2} (j_k^h(\bar{u}_\zeta^h) - \zeta_k)(\bar{p}_{\zeta,k}^h + \lambda_k \bar{u}_\zeta^h, \pi_0 \bar{u}_\zeta - \bar{u}_\zeta)_{\mathbb{U}} \geq 0. \end{aligned} \tag{3.26}$$

The fact that the cost functions $j_k$ and $j_k^h$ ($k \in \{1,2\}$) are quadratic implies the following Taylor expansions:

$$\begin{aligned} j_k(\bar{u}_\zeta^h) &= j_k(\bar{u}_\zeta) + (\bar{p}_{\zeta,k} + \lambda_k \bar{u}_\zeta, \bar{u}_\zeta^h - \bar{u}_\zeta)_{\mathbb{U}} + \frac{1}{2} \left( \sum_{i=1}^{n_k} (S(\bar{u}_\zeta - \bar{u}_\zeta^h)(\boldsymbol{x}_k^i))^2 + \lambda_k \|\bar{u}_\zeta - \bar{u}_\zeta^h\|_{\mathbb{U}}^2 \right), \\ j_k^h(\bar{u}_\zeta) &= j_k^h(\bar{u}_\zeta^h) + (\bar{p}_{\zeta,k}^h + \lambda_k \bar{u}_\zeta^h, \bar{u}_\zeta - \bar{u}_\zeta^h)_{\mathbb{U}} + \frac{1}{2} \left( \sum_{i=1}^{n_k} (S_h(\bar{u}_\zeta - \bar{u}_\zeta^h)(\boldsymbol{x}_k^i))^2 + \lambda_k \|\bar{u}_\zeta - \bar{u}_\zeta^h\|_{\mathbb{U}}^2 \right). \end{aligned}$$

Using these expansions in (3.26), it follows that

$$\begin{aligned} &\sum_{k=1}^{2} (j_k(\bar{u}_\zeta) - \zeta_k) \left( j_k(\bar{u}_\zeta^h) - j_k(\bar{u}_\zeta) - \tfrac{1}{2} \sum_{i=1}^{n_k} (S(\bar{u}_\zeta - \bar{u}_\zeta^h)(\boldsymbol{x}_k^i))^2 - \tfrac{\lambda_k}{2} \|\bar{u}_\zeta - \bar{u}_\zeta^h\|_{\mathbb{U}}^2 \right) \\ &+ \sum_{k=1}^{2} (j_k^h(\bar{u}_\zeta^h) - \zeta_k) \left( j_k^h(\bar{u}_\zeta) - j_k^h(\bar{u}_\zeta^h) - \tfrac{1}{2} \sum_{i=1}^{n_k} (S_h(\bar{u}_\zeta - \bar{u}_\zeta^h)(\boldsymbol{x}_k^i))^2 - \tfrac{\lambda_k}{2} \|\bar{u}_\zeta - \bar{u}_\zeta^h\|_{\mathbb{U}}^2 \right) \\ &+ \sum_{k=1}^{2} (j_k^h(\bar{u}_\zeta^h) - \zeta_k)(\bar{p}_{\zeta,k}^h + \lambda_k \bar{u}_\zeta^h, \pi_0 \bar{u}_\zeta - \bar{u}_\zeta)_{\mathbb{U}} \geq 0. \end{aligned}$$

Hence, using the identity

$$\begin{aligned}
&\sum_{k=1}^{2}(j_k(\bar{u}_\zeta)-\zeta_k)(j_k(\bar{u}_\zeta^h)-j_k(\bar{u}_\zeta))+\sum_{k=1}^{2}(j_k^h(\bar{u}_\zeta^h)-\zeta_k)(j_k^h(\bar{u}_\zeta)-j_k^h(\bar{u}_\zeta^h))\\
&\quad=\sum_{k=1}^{2}(j_k(\bar{u}_\zeta)-\zeta_k)(j_k(\bar{u}_\zeta^h)-j_k^h(\bar{u}_\zeta^h))+\sum_{k=1}^{2}(j_k(\bar{u}_\zeta)-\zeta_k)(j_k^h(\bar{u}_\zeta^h)-j_k(\bar{u}_\zeta))\\
&\quad+\sum_{k=1}^{2}(j_k^h(\bar{u}_\zeta^h)-\zeta_k)(j_k^h(\bar{u}_\zeta)-j_k^h(\bar{u}_\zeta^h))\\
&\quad=\sum_{k=1}^{2}(j_k(\bar{u}_\zeta)-\zeta_k)(j_k(\bar{u}_\zeta^h)-j_k^h(\bar{u}_\zeta^h))-\sum_{k=1}^{2}(j_k(\bar{u}_\zeta)-j_k^h(\bar{u}_\zeta^h))^2\\
&\quad+\sum_{k=1}^{2}(j_k^h(\bar{u}_\zeta^h)-\zeta_k)(j_k^h(\bar{u}_\zeta)-j_k(\bar{u}_\zeta)),
\end{aligned}$$

we arrive at

$$\begin{aligned}
&\sum_{k=1}^{2}\frac{(j_k(\bar{u}_\zeta)-\zeta_k)}{2}\left(\sum_{i=1}^{n_k}(S(\bar{u}_\zeta-\bar{u}_\zeta^h)(\boldsymbol{x}_k^i))^2+\lambda_k\|\bar{u}_\zeta-\bar{u}_\zeta^h\|_{\mathbb{U}}^2\right)+\sum_{k=1}^{2}(j_k(\bar{u}_\zeta)-j_k^h(\bar{u}_\zeta^h))^2\\
&\quad+\sum_{k=1}^{2}\frac{(j_k^h(\bar{u}_\zeta^h)-\zeta_k)}{2}\left(\sum_{i=1}^{n_k}(S_h(\bar{u}_\zeta-\bar{u}_\zeta^h)(\boldsymbol{x}_k^i))^2+\lambda_k\|\bar{u}_\zeta-\bar{u}_\zeta^h\|_{\mathbb{U}}^2\right)\\
&\leq\sum_{k=1}^{2}(j_k(\bar{u}_\zeta)-\zeta_k)(j_k(\bar{u}_\zeta^h)-j_k^h(\bar{u}_\zeta^h))+\sum_{k=1}^{2}(j_k^h(\bar{u}_\zeta^h)-\zeta_k)(j_k^h(\bar{u}_\zeta)-j_k(\bar{u}_\zeta))\\
&\quad+\sum_{k=1}^{2}(j_k^h(\bar{u}_\zeta^h)-\zeta_k)(\bar{p}_{\zeta,k}^h+\lambda_k\bar{u}_\zeta^h,\pi_0\bar{u}_\zeta-\bar{u}_\zeta)_{\mathbb{U}}.
\end{aligned}\tag{3.27}$$

Thus, using in (3.27) the identity

$$\begin{aligned}
j_k(\bar{u}_\zeta^h)-j_k^h(\bar{u}_\zeta^h)&=j_k(\bar{u}_\zeta)+(\bar{p}_{\zeta,k}+\lambda_k\bar{u}_\zeta,\bar{u}_\zeta^h-\bar{u}_\zeta)_{\mathbb{U}}+\frac{1}{2}\sum_{i=1}^{n_k}(S(\bar{u}_\zeta-\bar{u}_\zeta^h)(\boldsymbol{x}_k^i))^2+\frac{\lambda_k}{2}\|\bar{u}_\zeta-\bar{u}_\zeta^h\|_{\mathbb{U}}^2\\
&\quad-j_k^h(\bar{u}_\zeta)-(\tilde{p}_k^h+\lambda_k\bar{u}_\zeta^h,\bar{u}_\zeta^h-\bar{u}_\zeta)_{\mathbb{U}}-\frac{1}{2}\sum_{i=1}^{n_k}(S_h(\bar{u}_\zeta-\bar{u}_\zeta^h)(\boldsymbol{x}_k^i))^2-\frac{\lambda_k}{2}\|\bar{u}_\zeta-\bar{u}_\zeta^h\|_{\mathbb{U}}^2\\
&=(j_k(\bar{u}_\zeta)-j_k^h(\bar{u}_\zeta))+(\bar{p}_{\zeta,k}-\tilde{p}_k^h,\bar{u}_\zeta^h-\bar{u}_\zeta)_{\mathbb{U}}-\lambda_k\|\bar{u}_\zeta-\bar{u}_\zeta^h\|_{\mathbb{U}}^2\\
&\quad+\frac{1}{2}\sum_{i=1}^{n_k}(S(\bar{u}_\zeta-\bar{u}_\zeta^h)(\boldsymbol{x}_k^i))^2-\frac{1}{2}\sum_{i=1}^{n_k}(S_h(\bar{u}_\zeta-\bar{u}_\zeta^h)(\boldsymbol{x}_k^i))^2,
\end{aligned}$$

which follows from a Taylor expansion, we conclude that

$$\begin{aligned}
&\left(\sum_{k=1}^{2}\left(\frac{3j_k(\bar u_\zeta)+j_k^h(\bar u_\zeta^h)}{2}-2\zeta_k\right)\lambda_k\right)\|\bar u_\zeta-\bar u_\zeta^h\|_{\mathbb U}^2+\sum_{k=1}^{2}(j_k(\bar u_\zeta)-j_k^h(\bar u_\zeta^h))^2\\
&\quad+\sum_{k=1}^{2}\left(\frac{j_k(\bar u_\zeta)+j_k^h(\bar u_\zeta^h)}{2}-\zeta_k\right)\left(\sum_{i=1}^{n_k}(S_h(\bar u_\zeta-\bar u_\zeta^h)(\boldsymbol{x}_k^i))^2\right)\\
&\le\sum_{k=1}^{2}(j_k(\bar u_\zeta)-j_k^h(\bar u_\zeta^h))(j_k(\bar u_\zeta)-j_k^h(\bar u_\zeta))+\sum_{k=1}^{2}(j_k(\bar u_\zeta)-\zeta_k)(\bar p_{\zeta,k}-\tilde p_k^h,\bar u_\zeta^h-\bar u_\zeta)_{\mathbb U}\\
&\quad+\sum_{k=1}^{2}(j_k^h(\bar u_\zeta^h)-\zeta_k)(\bar p_{\zeta,k}^h+\lambda_k\bar u_\zeta^h,\pi_0\bar u_\zeta-\bar u_\zeta)_{\mathbb U}=:\mathsf{I}+\mathsf{II}+\mathsf{III}.
\end{aligned}\tag{3.28}$$

We recall that the auxiliary term $\tilde p_k^h$ ($k\in\{1,2\}$) is defined in (3.25). We immediately note that, for $h>0$ sufficiently small, the inequality $j_k^h(\bar u_\zeta^h)-\zeta_k>0$ ($k\in\{1,2\}$) holds; this follows from the fact that $\|\bar u_\zeta-\bar u_\zeta^h\|_{\mathbb U}\to 0$ when $h\to 0$ in combination with Lemma 3.6. Consequently, all the terms on the left-hand side of (3.28) are nonnegative when $h>0$ is small enough.

We now estimate $\mathsf{I}$, $\mathsf{II}$, and $\mathsf{III}$. To estimate $\mathsf{I}$ we use Lemma 3.6 and Young's inequality to arrive at

$$\mathsf{I}\lesssim(h^2|\log h|^2+\|\bar u_\zeta-\bar u_\zeta^h\|_{\mathbb U})h^2|\log h|^2\lesssim(1+\varepsilon^{-1})h^4|\log h|^4+\varepsilon\|\bar u_\zeta-\bar u_\zeta^h\|_{\mathbb U}^2,$$

with $\varepsilon>0$ arbitrary. To control the term $\mathsf{II}$, we use Hölder's inequality, an analogous version of Proposition 3.2 for $\|\bar p_{\zeta,k}-\tilde p_k^h\|_{L^1(\Omega)}$, and the fact that $\bar u_\zeta^h,\bar u_\zeta\in\mathbb U_{\mathsf{ad}}$ to obtain

$$\mathsf{II}\lesssim\|\bar p_{\zeta,k}-\tilde p_k^h\|_{L^1(\Omega)}\|\bar u_\zeta^h-\bar u_\zeta\|_{L^\infty(\Omega)}\lesssim h^2|\log h|^2.$$

The term $\mathsf{III}$ can be bounded as the term $\mathsf{II}$ in the proof of Theorem 3.3 (*cf.* (3.16)). Thus, we obtain that

$$\mathsf{III}\lesssim\varepsilon^{-1}h^2+\varepsilon\|\bar u_\zeta^h-\bar u_\zeta\|_{\mathbb U}^2,\tag{3.29}$$

with $\varepsilon>0$ arbitrary, upon using that $\sum_{k=1}^{2}|j_k^h(\bar u_\zeta^h)-\zeta_k|\le C$, for $C>0$.

We conclude the desired bound by replacing the estimates obtained for $\mathsf{I}$, $\mathsf{II}$, and $\mathsf{III}$ in (3.28) and taking $\varepsilon>0$ small enough. □

As in the discrete approximation of the weighted-sum method (see Thm. 3.4), we can improve the error estimate of Theorem 3.7 in two dimensions.

**Theorem 3.8** (Improved error estimate: RPM)**.** *Let $d=2$. In the framework of Theorem 3.7, we have the following optimal error estimate:*

$$\|\bar u_\zeta-\bar u_\zeta^h\|_{\mathbb U}\lesssim h\quad\text{for all } 0<h<h_*.$$

*Proof.* The proof relies on different estimates for $\mathsf{I}$ and $\mathsf{II}$ in (3.28). First, we note that similar arguments to the ones that lead to Lemma 3.6 in combination with the error estimate from [23], Theorem 3 give

$$\sum_{k=1}^{2}|j_k(v)-j_k^h(w)|\lesssim h+\|v-w\|_{\mathbb U}.\tag{3.30}$$

TABLE 1. Approximation error $\|\bar{u}_\alpha - \bar{u}_\alpha^h\|_{\mathbb{U}}$ considering the four different choices for $\alpha$.

| $h$ | $\alpha = (0.2, 0.8)$ | $\alpha = (0.4, 0.6)$ | $\alpha = (0.6, 0.4)$ | $\alpha = (0.8, 0.2)$ |
|---|---|---|---|---|
| $2^{-2}$ | 0.727125 | 0.994289 | 1.312741 | 1.580559 |
| $2^{-3}$ | 0.399550 | 0.555188 | 0.704527 | 0.790838 |
| $2^{-4}$ | 0.209558 | 0.300159 | 0.353305 | 0.389034 |
| $2^{-5}$ | 0.107604 | 0.155751 | 0.173634 | 0.193365 |
| Rate of convergence | 0.92 | 0.89 | 0.97 | 1.01 |

Then, we bound $\mathsf{I}$ using the estimate (3.30) as follows:

$$\mathsf{I} \lesssim (h + \|\bar{u}_\zeta - \bar{u}_\zeta^h\|_{\mathbb{U}}^2)h \lesssim (1 + \varepsilon^{-1})h^2 + \varepsilon\|\bar{u}_\zeta - \bar{u}_\zeta^h\|_{\mathbb{U}}^2,$$

with $\varepsilon > 0$ arbitrary. For the term $\mathsf{II}$ in (3.28), we use the error bound $\|\bar{p}_{\zeta,k} - \tilde{p}_k^h\|_{\mathbb{U}} \lesssim h$ and obtain

$$\mathsf{II} \lesssim \|\bar{p}_{\zeta,k} - \tilde{p}_k^h\|_{\mathbb{U}}\|\bar{u}_\zeta^h - \bar{u}_\zeta\|_{\mathbb{U}} \lesssim \varepsilon^{-1}h^2 + \varepsilon\|\bar{u}_\zeta - \bar{u}_\zeta^h\|_{\mathbb{U}}^2,$$

with $\varepsilon > 0$ arbitrary. Therefore, the use of the previous two estimates and (3.29) in (3.28), and taking $\varepsilon > 0$ and $h > 0$ sufficiently small, yields the desired result. ☐

## 4. NUMERICAL EXAMPLES

In this section, we present some numerical experiments that show the performance of the scalarization methods presented in Section 2.3 when approximating Pareto stationary points and front. The experiments were carried out with a code implemented in `MATLAB© (R2024a)`.

In both numerical examples, we consider the domains $\Omega = (0,1)^2$, $\Omega_1 = \{(0.75, 0.25)\}$, $\Omega_2 = \{(0.25, 0.75)\}$, the desired states $y_1^1 = 6$, $y_2^1 = -2$, and the bilateral bounds $u_{\mathsf{a}} = -7$, $u_{\mathsf{b}} = 15$. For the regularization parameters $\lambda_1$ and $\lambda_2$ in (1.1), we consider four different configurations: $(\lambda_1, \lambda_2) = (1, 1)$, $(\lambda_1, \lambda_2) = (1, 0.1)$, $(\lambda_1, \lambda_2) = (0.1, 1)$, and $(\lambda_1, \lambda_2) = (0.1, 0.1)$.

In the absence of an exact solution, we compute the error committed in the approximation by taking as a reference solution $\bar{u}_\alpha$ (resp. $\bar{u}_\zeta$) the discrete optimal control $\bar{u}_\alpha^h$ (resp. $\bar{u}_\zeta^h$) obtained on a finer triangulation $\mathcal{T}_h$: the mesh $\mathcal{T}_h$ is such that $h = 2^{-8}$.

We finally mention that no significant differences are observed in the numerical performance of the two methods when computing a point on the Pareto front for fixed values of $\alpha$ and $\zeta$. This is a consequence of the structure of Algorithms 1 and 3. Therefore, when a similar number of points is used to approximate the Pareto front, neither method significantly outperforms the other.

### 4.1. Example 1 (WSM)

We present the results obtained for this example in Figures 1–2, and Table 1. In Figure 1, the Pareto fronts and the approximation error $\|\mathcal{J}(\bar{y}_\alpha^h, \bar{u}_\alpha^h) - \mathcal{J}(\bar{y}_\alpha, \bar{u}_\alpha)\|_2$ are shown with 50 different values of $\alpha$ considering mesh refinement using the weighted-sum method, and for the four different configurations of $(\lambda_1, \lambda_2)$ mentioned above. For these four configurations, we observe that the approximate Pareto front seems to converge to the continuous one as the discretization parameter $h$ decreases. We also observe that, even for the same problem data, the shape of the curves that represent the Pareto fronts changes as the values of $\lambda_1$ and $\lambda_2$ change. In Figure 2, we show the approximate optimal control $\bar{u}_\alpha^h$ obtained with $\lambda_1 = \lambda_2 = 0.1$, $h = 2^{-6}$, and for four different instances of $\alpha$, namely $\alpha = (0.2, 0.8)$, $\alpha = (0.4, 0.6)$, $\alpha = (0.6, 0.4)$, and $\alpha = (0.8, 0.2)$. Finally, in Table 1, we present the approximation error $\|\bar{u}_\alpha - \bar{u}_\alpha^h\|_{\mathbb{U}}$ for different cases of $\alpha$. We observe that optimal experimental rates of convergence $\mathcal{O}(h)$ are attained for the four different choices of $\alpha$, which is in agreement with Theorem 3.4.

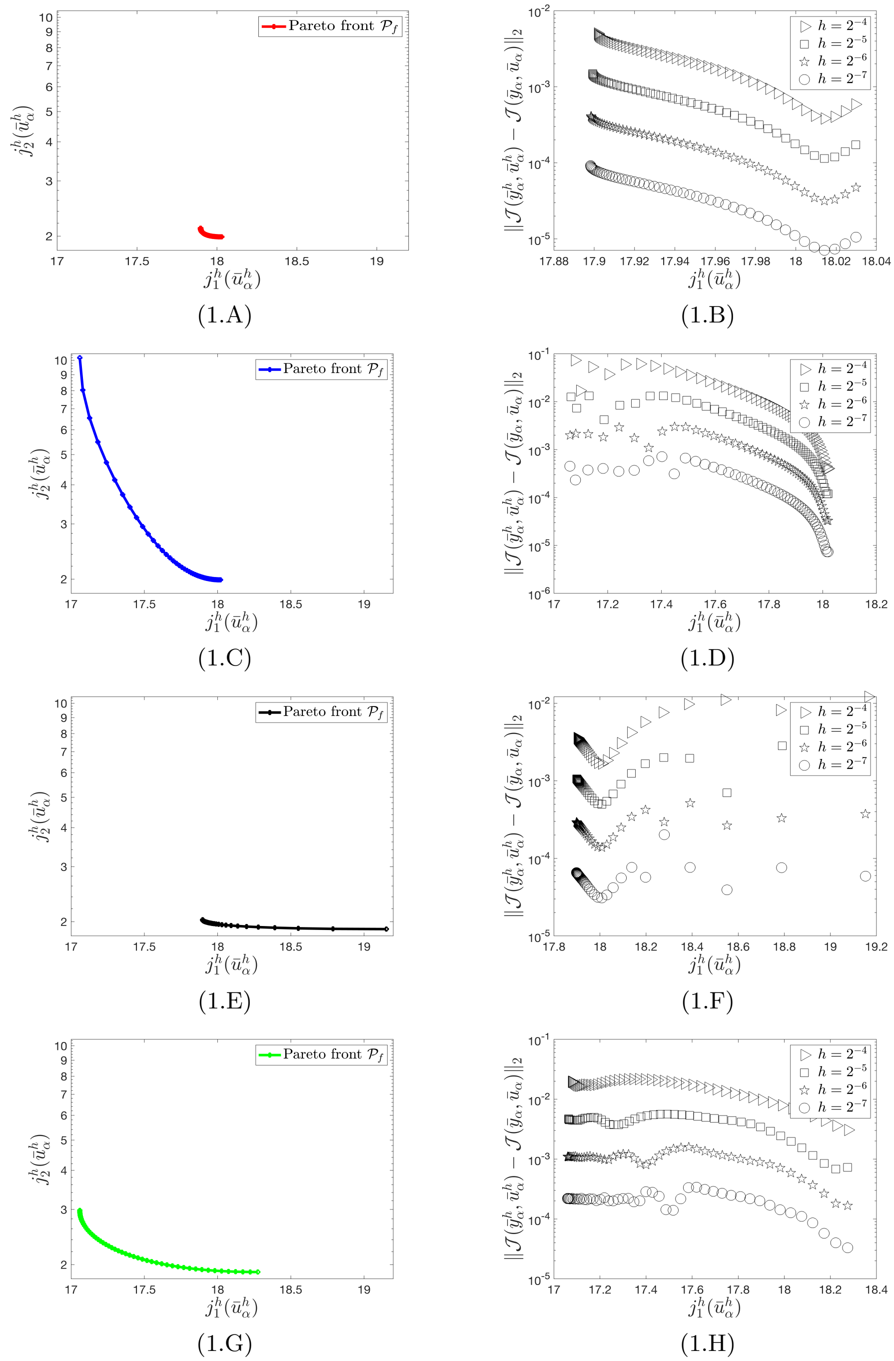
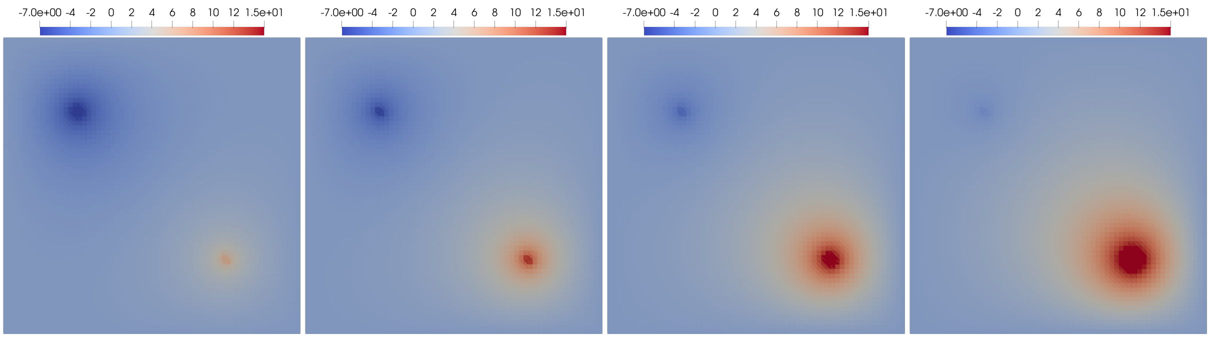
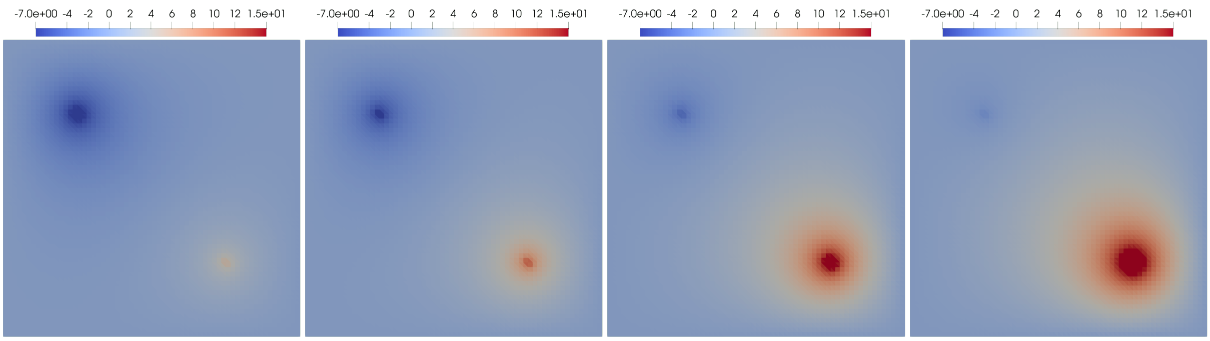



FIGURE 1. Pareto front and approximation error $\|\mathcal{J}(\bar{y}_\alpha^h, \bar{u}_\alpha^h) - \mathcal{J}(\bar{y}_\alpha, \bar{u}_\alpha)\|_2$ with 50 different values of $\alpha$ considering mesh refinement under the WSM for the cases $\lambda_1 = \lambda_2 = 1$ (1.A)–(1.B), $\lambda_1 = 0.1$ and $\lambda_2 = 1$ (1.C)–(1.D), $\lambda_1 = 1$ and $\lambda_2 = 0.1$ (1.E)–(1.F), and $\lambda_1 = \lambda_2 = 0.1$ (1.G)–(1.H).




(2.A) (2.B) (2.C) (2.D)

FIGURE 2. Approximate optimal control $\bar{u}_\alpha^h$ when $\alpha = (0.2, 0.8)$ (2.A), $\alpha = (0.4, 0.6)$ (2.B), $\alpha = (0.6, 0.4)$ (2.C), and $\alpha = (0.8, 0.2)$ (2.D). Here, $h = 2^{-6}$ and $\lambda_1 = \lambda_2 = 0.1$.




(3.A) (3.B) (3.C) (3.D)

FIGURE 3. Approximate optimal control $\bar{u}_\zeta^h$ for $\zeta^9$ (3.A), $\zeta^7$ (3.B), $\zeta^4$ (3.C), and $\zeta^2$ (3.D). Here, $h = 2^{-6}$ and $\lambda_1 = \lambda_2 = 0.1$.

TABLE 2. Approximation error $\|\bar{u}_\zeta - \bar{u}_\zeta^h\|_{\mathbb{U}}$ considering the four different choices for $\zeta$ mentioned in (4.1).

| $h$ | $\zeta^2$ | $\zeta^4$ | $\zeta^7$ | $\zeta^9$ |
|---|---|---|---|---|
| $2^{-2}$ | 1.583765 | 1.338101 | 1.002442 | 0.928546 |
| $2^{-3}$ | 0.799229 | 0.711743 | 0.489412 | 0.375962 |
| $2^{-4}$ | 0.390953 | 0.353613 | 0.266316 | 0.195219 |
| $2^{-5}$ | 0.193401 | 0.173748 | 0.139464 | 0.096843 |
| Rate of convergence | 1.01 | 0.98 | 0.94 | 1.07 |

## 4.2. Example 2 (RPM)

For the reference point method, we choose in (3.21) the scaling parameters $\mathfrak{h}^\perp = \mathfrak{h}^\parallel = 0.2$. We note that the values of the points $\zeta$ change in each refinement, due to their implicit dependency on $h$; see (3.22). In this example, we show the error in the steps $\zeta^2, \zeta^4, \zeta^7$ and $\zeta^9$, whose values on the finest mesh ($h = 2^{-8}$) are

$$\zeta^2 \approx (16.89, 2.58), \qquad \zeta^4 \approx (17.04, 2.21), \qquad \zeta^7 \approx (17.49, 1.82), \qquad \zeta^9 \approx (17.88, 1.71). \tag{4.1}$$

We present the results obtained for this example in Figure 3 and Table 2. In fact, we get similar results to those shown for the WSM. In particular, we observe optimal experimental rates of convergence $\mathcal{O}(h)$ for all the different cases of $\zeta$, which is in agreement with Theorem 3.8.

## Funding

This project has received funding by the Federal Ministry of Education and Research (BMBF) and the Baden-Württemberg Ministry of Science as part of the Excellence Strategy of the German Federal and State Governments. In addition, the first author has been supported by ANID through FONDECYT postdoctoral project 3230126.

## Data availability statement

The data in this article are available upon reasonable request to the corresponding author.

## References

[1] M. Ehrgott, *Multicriteria Optimization*, 2nd edn. Springer-Verlag, Berlin (2005).
[2] K. Miettinen, *Nonlinear Multiobjective Optimization*, vol. 112. Springer, New York (1998).
[3] D. Beermann, M. Dellnitz, S. Peitz and S. Volkwein, POD-based multiobjective optimal control of PDEs with non-smooth objectives. *PAMM* **17** (2017) 51–54.
[4] M. Bernreuther, G. Müller and S. Volkwein, Efficient scalarization in multiobjective optimal control of a nonsmooth PDE. *Comput. Optim. Appl.* **83** (2022) 435–464.
[5] K. Sonntag, B. Gebken, G. Müller, S. Peitz and S. Volkwein, A descent method for nonsmooth multiobjective optimization in Hilbert spaces. *J. Optim. Theory Appl.* **203** (2024) 455–487.
[6] L. Chang, W. Gong and N. Yan, Numerical analysis for the approximation of optimal control problems with pointwise observations. *Math. Methods Appl. Sci.* **38** (2015) 4502–4520.
[7] C. Brett, A. Dedner and C. Elliott, Optimal control of elliptic PDEs at points. *IMA J. Numer. Anal.* **36** (2016) 1015–1050.
[8] H. Antil, E. Otárola and A.J. Salgado, Some applications of weighted norm inequalities to the error analysis of PDE-constrained optimization problems. *IMA J. Numer. Anal.* **38** (2018) 852–883.
[9] N. Behringer, D. Meidner and B. Vexler, Finite element error estimates for optimal control problems with pointwise tracking. *Pure Appl. Funct. Anal.* **4** (2019) 177–204.
[10] A. Allendes, F. Fuica and E. Otárola, Error estimates for a pointwise tracking optimal control problem of a semilinear elliptic equation. *SIAM J. Control Optim.* **60** (2022) 1763–1790.
[11] A. Allendes, F. Fuica, E. Otárola and D. Quero, An adaptive FEM for the pointwise tracking optimal control problem of the Stokes equations. *SIAM J. Sci. Comput.* **41** (2019) A2967–A2998.
[12] N. Behringer, Improved error estimates for optimal control of the Stokes problem with pointwise tracking in three dimensions. *Math. Control Relat. Fields* **11** (2021) 313–328.
[13] F. Fuica and E. Otárola, A pointwise tracking optimal control problem for the stationary Navier–Stokes equations. *J. Math. Anal. Appl.* **558** (2026) Paper No. 130343, 29.
[14] F. Fuica, E. Otárola and D. Quero, Error estimates for optimal control problems involving the Stokes system and Dirac measures. *Appl. Math. Optim.* **84** (2021) 1717–1750.
[15] S. Banholzer, *POD-based Bicriterial Optimal Control of Convection-Diffusion Equations*. Master's thesis, University of Konstanz, Germany (2017).
[16] S. Banholzer, *ROM-Based Multiobjective Optimization with PDE Constraints*. PhD thesis, University of Konstanz (2021).
[17] C. Romaus, J. Bocker, K. Witting, A. Seifried and O. Znamenshchykov. Optimal energy management for a hybrid energy storage system combining batteries and double layer capacitors, in *2009 IEEE Energy Conversion Congress and Exposition* (2009) 1640–1647.
[18] L.C. Evans, *Partial Differential Equations*, vol. 19 of *Graduate Studies in Mathematics*, 2nd edn. American Mathematical Society, Providence, RI (2010).
[19] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, volume 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA (1985).

[20] V. Maz'ya and J. Rossmann, *Elliptic Equations in Polyhedral Domains*, vol. 162 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI (2010).

[21] R.A. Adams and J.J.F. Fournier, *Sobolev Spaces*, vol. 140 of *Pure and Applied Mathematics (Amsterdam)*, 2nd edn. Elsevier/Academic Press, Amsterdam (2003).

[22] F. Tröltzsch, *Optimal Control of Partial Differential Equations*, vol. 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI (2010).

[23] E. Casas, $L^2$ estimates for the finite element method for the Dirichlet problem with singular data. *Numer. Math.* **47** (1985) 627–632.

[24] A. Allendes, E. Otárola, R. Rankin and A.J. Salgado, Adaptive finite element methods for an optimal control problem involving Dirac measures. *Numer. Math.* **137** (2017) 159–197.

[25] S. Peitz, S. Ober-Blöbaum and M. Dellnitz, Multiobjective optimal control methods for the Navier–Stokes equations using reduced order modeling. *Acta Appl. Math.* **161** (2019) 171–199.

[26] S. Banholzer, D. Beermann and S. Volkwein, POD-based bicriterial optimal control by the reference point method. *IFAC-PapersOnLine* **49** (2016) 210–215.

[27] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods*, vol. 15 of *Texts in Applied Mathematics*, 3rd edn. Springer, New York (2008).

[28] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, vol. 159 of *Applied Mathematical Sciences*. Springer-Verlag, New York (2004).

[29] J. Barzilai and J.M. Borwein, Two-point step size gradient methods. *IMA J. Numer. Anal.* **8** (1988) 141–148.

[30] B. Azmi and K. Kunisch, Analysis of the Barzilai–Borwein step-sizes for problems in Hilbert spaces. *J. Optim. Theory Appl.* **185** (2020) 819–844.

[31] R. Scott, Finite element convergence for singular data. *Numer. Math.* **21** (1973) 317–327.

[32] R. Scott, Optimal $l^\infty$ estimates for the finite element method on irregular meshes. *Math. Computat.* **30** (1976) 681–697.

[33] F. Tröltzsch and S. Volkwein, POD a-posteriori error estimates for linear-quadratic optimal control problems. *Computat. Optim. Appl.* **44** (2009) 83–115.