Applied Mathematics & Optimization (2024) 90:20
https://doi.org/10.1007/s00245-024-10162-1

Bilevel Optimization of the Kantorovich Problem and Its Quadratic Regularization

Part II: Convergence Analysis

Sebastian Hillbrecht[1] · **Paul Manns**[1] · **Christian Meyer**[1]

Accepted: 24 June 2024 / Published online: 14 July 2024

Abstract

This paper is concerned with an optimization problem which is governed by the Kantorovich problem of optimal transport. More precisely, we consider a bilevel optimization problem with the underlying problem being the Kantorovich problem. This task can be reformulated as a mathematical problem with complementarity constraints in the space of regular Borel measures. Because of the non-smoothness that is induced by the complementarity constraints, problems of this type are often regularized, e.g., by an entropic regularization. However, in this paper we apply a quadratic regularization to the Kantorovich problem. By doing so, we are able to drastically reduce its dimension while preserving the sparsity structure of the optimal transportation plan as much as possible. As the title indicates, this is the second part in a series of three papers. While the existence of optimal solutions to both the bilevel Kantorovich problem and its regularized counterpart were shown in the first part, this paper deals with the (weak-$*$) convergence of solutions to the regularized bilevel problem to solutions of the original bilevel Kantorovich problem for vanishing regularization parameters.

Keywords Optimal transport · Kantorovich problem · Bilevel optimization · Quadratic regularization

Mathematics Subject Classification 49Q20 · 90C08 · 49J45

✉ Sebastian Hillbrecht
sebastian.hillbrecht@tu-dortmund.de

Paul Manns
paul.manns@tu-dortmund.de

Christian Meyer
christian2.meyer@tu-dortmund.de

[1] Lehrstuhl X, Fakultät für Mathematik, Technische Universität Dortmund, Vogelpothsweg 87, 44227 Dortmund, Germany

1 Introduction

In this paper, we consider a bilevel optimization problem with the Kantorovich problem of optimal transport as its lower-level problem. Given two non-negative marginals $\mu_1 \in \mathfrak{M}(\Omega_1)$ and $\mu_2 \in \mathfrak{M}(\Omega_2)$ which are defined on the compact domains Ω_1 and Ω_2 and have the same total mass and a continuous cost function $c \in C(\Omega_1 \times \Omega_2)$, the Kantorovich problem reads as follows:

$$\left.\begin{aligned} \inf_{\pi} \quad & \mathcal{K}(\pi) := \int_{\Omega} c(x)\, \mathrm{d}\pi(x) \\ \text{s.t.} \quad & \pi \in \Pi(\mu_1, \mu_2), \quad \pi \geq 0. \end{aligned}\right\} \tag{KP}$$

Herein, $\Pi(\mu_1, \mu_2)$ denotes the set of feasible couplings of μ_1 and μ_2 defined by

$$\Pi(\mu_1, \mu_2) := \{\pi \in \mathfrak{M}(\Omega_1 \times \Omega_2) \colon P_{i\#}\pi = \mu_i, \ i = 1, 2\}, \tag{1}$$

where P_i denotes the projection on the i-th variable and $P_{i\#}\pi$ is the push forward of P_i w.r.t. π. The Kantorovich problem is a classical, well-established model for optimizing transportation processes and we exemplarily refer to [1–4] for more details on its mathematical background.

For the definition of the bilevel problem, the cost function is set to $c = c_{\mathrm{d}} \in C(\Omega_1 \times \Omega_2)$ and we fix the second marginal $\mu_2 = \mu_2^{\mathrm{d}}$, i.e., both c_{d} and μ_2^{d} are given data. The upper-level optimization variables are π and μ_1 such that the overall bilevel problem is given by

$$\left.\begin{aligned} \inf_{\pi, \mu_1} \quad & \mathcal{J}(\pi, \mu_1) \\ \text{s.t.} \quad & \mu_1 \in \mathfrak{M}(\Omega_1), \quad \pi \in \mathfrak{M}(\Omega_1 \times \Omega_2), \\ & \mu_1 \geq 0, \quad \|\mu_1\|_{\mathfrak{M}(\Omega_1)} = \|\mu_2^{\mathrm{d}}\|_{\mathfrak{M}(\Omega_2)}, \\ & \pi \text{ solves (KP) with } \mu_2 = \mu_2^{\mathrm{d}} \text{ and } c = c_{\mathrm{d}}. \end{aligned}\right\} \tag{BK}$$

A potential application of such a bilevel problem could, for instance, be the identification of the marginal μ_1 based on measurements of the transportation process. The upper-level objective would then read $\mathcal{J}(\pi, \mu_1) = \mathcal{J}(\pi) := \|\pi - \pi^{\mathrm{d}}\|_{\mathfrak{M}(D)}$, where $D \subset \Omega_1 \times \Omega_2$ is a given observation domain and $\pi^{\mathrm{d}} \in \mathfrak{M}(D)$ denotes the measured transport plan in D, see Sect. 4.1. Another example for a problem of type (BK) reads as follows: Let $\Omega_1 = \Omega_2 =: \Omega_*$ and let a completely continuous mapping $G : \mathfrak{M}(\Omega_*) \to L^2(\Omega_*)$ be given. Then, for $\beta \in [1, \infty)$ and $\nu > 0$, we consider

$$\left.\begin{aligned} \min_{\mu_1} \quad & \frac{1}{2}\|G(\mu_1) - z\|^2_{L^2(\Omega_*)} + \nu\, W_\beta(\mu_1, \mu_2^{\mathrm{d}})^\beta \\ \text{s.t.} \quad & \mu_1 \in \mathfrak{M}(\Omega_*), \quad \mu_1 \geq 0, \quad \|\mu_1\|_{\mathfrak{M}(\Omega_*)} = 1, \end{aligned}\right\} \tag{WCP}$$

where W_β denotes the Wasserstein distance of order β. Since this is just the β^{th} root of the optimal value of the Kantorovich problem with transportation cost $c(x_1, x_2) :=$

$\|x_1 - x_2\|^\beta$, we are indeed faced with a problem of type (BK). In Sect. 4.2 below, we will investigate an example of this problem class, where G is the solution operator of an elliptic PDE.

For its numerical solution, the Kantorovich problem is frequently regularized in order to avoid the "curse of dimensionality" caused by the discretization of the transport plan on the product space $\Omega_1 \times \Omega_2$. A prominent example is the entropic regularization (see e.g. [5] for a convergence analysis in function space), which leads to the well-known Sinkhorn algorithm, cf. [6, 7]. An alternative regularization approach was proposed in [8], where the squared $L^2(\Omega_1 \times \Omega_2)$-norm of π is added to the objective of (KP), weighted with a regularization parameter $\gamma > 0$. The advantageous implications of this approach are similar to the ones caused by the entropic regularization. First, the regularized counterpart of (KP) is a strictly convex problem so that its solution is unique. Moreover, the regularization leads to a substantial reduction of the dimension since the dual problem is equivalent to a (non-smooth) system of equations in $L^2(\Omega_1) \times L^2(\Omega_2)$ instead of $L^2(\Omega_1 \times \Omega_2)$. In [8], a semi-smooth Newton-type method is employed to solve this system and the convergence for $\gamma \searrow 0$ (in a more general framework covering the quadratic regularization) is investigated in [9]. A further advantage (in comparison to the entropic regularization) of the quadratic regularization is that it better preserves the sparsity pattern of the optimal transport plan as the numerical experiments in [8] demonstrate.

In view of the success of regularization techniques for the (numerical) solution of the Kantorovich problem, it is reasonable to apply them in the bilevel context too. As the title indicates, we follow the quadratic approach from [8] and replace the Kantorovich problem in (BK) by its quadratically regularized counterpart. In this context, the following questions naturally arise:

- Does the bilevel problem (BK_ρ) and its regularized counterpart admit (globally optimal) solutions?
- Do solutions of the regularized bilevel problems (or subsequences thereof) converge to solutions of (BK_ρ) for vanishing regularization parameter $\gamma \searrow 0$?
- How can we efficiently solve (discretized versions of) the regularized bilevel problems?

While the first question is addressed in the predecessor paper [10], the present paper investigates the convergence behavior of such optimal solutions for $\gamma \searrow 0$. So far, we are only able to show that (subsequences of) optimal solutions converge (weakly-$*$) to optimal solutions of the original problem (BK_ρ) under fairly restrictive assumptions on the data μ_2^{d} and the structure of the objective. Nevertheless, at the end of the paper, we will see that there are relevant examples, where these assumptions are fulfilled. At least in finite dimensions, these assumptions can be weakened, as we show in the third part of this series of papers, see [11]. The third question, concerning an efficient and robust numerical solution of the regularized bilevel problems, is subject to future research. Albeit regularized, the bilevel problems are still non-smooth, as the necessary and sufficient optimality conditions associated with the regularized counterpart to (KP) involve the max-operator, see [8, Theorem 2.11]. At the same time, this operator promotes the desired sparsity of the solution and for this reason, a further smoothing of the max-operator should be avoided. We expect that algorithms, which are tailored to bilevel problems with non-smooth lower-level problem, behave well in this setting, see, for instance, the approaches in [12–14].

The remainder of this paper is organized as follows: After introducing some basic notation and assumptions in the rest of this introduction, we collect some known results on the Kantorovich problem, its quadratic regularization as well as the existence results from the companion paper [10] in Sect. 2. The remaining part of the paper is then devoted to the convergence analysis for vanishing regularization parameter $\gamma \searrow 0$. First, in Sect. 3.1, we show that weak-$*$ accumulation points of solutions of the regularized bilevel problems are feasible for (BK_ρ), i.e., in particular, the limit of the sequence of transport plans solves the Kantorovich problem associated with the limit of the marginals. Afterwards, in Sect. 3.2, we establish the optimality of the weak-$*$ limit under additional assumptions. The paper ends with two application-driven examples in Sect. 4, where the additional assumptions are fulfilled.

1.1 Notation and Standing Assumptions

Throughout the paper, the Euclidean norm of a vector $a \in \mathbb{R}^n$, $n \in \mathbb{N}$, is denoted by $|a|$. Moreover, the open ball in $\mathbb{R}^n$ of radius $r > 0$ centered in a is denoted by $B(a, r)$.

Domains

For $d_1, d_2 \in \mathbb{N}$, let $\Omega_1 \subset \mathbb{R}^{d_1}$ and $\Omega_2 \subset \mathbb{R}^{d_2}$ be compact sets with non-empty interior. We moreover suppose that their Cartesian product $\Omega := \Omega_1 \times \Omega_2$ coincides with the closure of its interior and has a Lipschitz boundary in the sense of [15, Def. 1.2.2.1]. The dimension of Ω is $d := d_1 + d_2$. By $\mathfrak{B}(\Omega)$ we denote the Borel σ-algebra on Ω and λ is the Lebesgue measure on $\mathfrak{B}(\Omega)$. For Ω_1 and Ω_2, $\mathfrak{B}(\Omega_i)$ and λ_i, $i = 1, 2$, are defined analogously so that $\lambda = \lambda_1 \otimes \lambda_2$. Furthermore, we abbreviate $|\Omega_1| := \lambda_1(\Omega_1)$, $|\Omega_2| := \lambda_2(\Omega_2)$, and $|\Omega| := \lambda(\Omega)$.

Marginals

For some subset $X \subset \mathbb{R}^d$, we denote the Banach space of (signed) regular Borel measures on the measurable space $\big(X, \mathfrak{B}(X)\big)$, in which $\mathfrak{B}(X)$ refers to the Borel σ-algebra of the open sets of X, by $\mathfrak{M}(X)$. It consists of signed Borel measures $\mu : \mathfrak{B}(X) \to \mathbb{R}$ whose variation measures $|\mu|$ are (inner and outer) regular. Its norm is the total variation $\|\mu\|_{\mathfrak{M}(X)} := |\mu|(X)$. If X happens to be compact, then the Riesz-Radon theorem ensures that $\mathfrak{M}(X) \cong C(X)^*$, i.e., the dual space of the Banach space of continuous functions can be identified with the Banach space of regular Borel measures. By $P_i : \Omega_1 \times \Omega_2 \ni (x_1, x_2) \mapsto x_i \in \Omega_i$, $i = 1, 2$, we denote the projection onto the i-th variable. For the ease of notation, we will denote projections with different domains and ranges by the same symbol, i.e., e.g., $P_2 : \Omega_1 \times \Omega_1 \ni (x_1, y_1) \mapsto y_1 \in \Omega_1$. The respective domains and ranges will become clear from the context. If $\mu_1 \in \mathfrak{M}(\Omega_1)$ and $\mu_2 \in \mathfrak{M}(\Omega_2)$, the set of transport plans between the marginals μ_1 and μ_2 is given by

$$\Pi(\mu_1, \mu_2) := \{\pi \in \mathfrak{M}(\Omega) : P_{1\#}\pi = \mu_1 \text{ and } P_{2\#}\pi = \mu_2\}, \tag{2}$$

with the pushforward measure of π via the projection P_i, $i = 1, 2$, being defined by

$$P_{i\#}\pi := \pi \circ P_i^{-1} : \mathfrak{B}(\Omega_i) \to \mathbb{R}.$$

Elements from $\Pi(\mu_1, \mu_2)$ are frequently called *couplings* of μ_1 and μ_2. Note that $\Pi(\mu_1, \mu_2) = \emptyset$ if and only if $\mu_1(\Omega_1) \neq \mu_2(\Omega_2)$. Throughout the paper, $\mu_2^{\mathrm{d}} \in \mathfrak{M}(\Omega_2)$ is a fixed marginal satisfying $\mu_2^{\mathrm{d}} \geq 0$ and, in order to ease notation, $|\mu_2^{\mathrm{d}}|(\Omega_2) = 1$. The normalization condition is no restriction and can be ensured by re-scaling. In order to shorten the notation, we write $\mathfrak{P}(\Omega_i) := \{\mu \in \mathfrak{M}(\Omega_i) : \mu \geq 0,\ |\mu|(\Omega_i) = 1\}$, $i = 1, 2$, for the set of probability measures on Ω_i.

Since Ω_1, Ω_2, and Ω are compact, the pre-dual spaces of $\mathfrak{M}(\Omega_1)$, $\mathfrak{M}(\Omega_2)$, and $\mathfrak{M}(\Omega)$ are $C(\Omega_1)$, $C(\Omega_2)$, and $C(\Omega)$, respectively. We denote the associated dual pairings by $\langle \mu, v \rangle$ and it will become clear from the context, which domain this refers to.

Given a measure space $(X, \mathcal{A}, \mu)$, the Lebesgue space of p-times integrable functions is denoted by $L^p(X, \mu)$, $p \in [1, \infty)$. If $X \subset \mathbb{R}^n$, $n \in \mathbb{N}$, is a Lebesgue measurable set, $\mathcal{A} = \mathfrak{B}(X)$, and μ is the Lebesgue measure, we write $L^p(X)$.

Cost Function

The cost function is assumed to satisfy $c_{\mathrm{d}} \in W^{1,p}(\Omega)$, $p > d$, where, with a slight abuse of notation, $W^{1,p}(\Omega)$ denotes the Sobolev space on $\mathrm{int}(\Omega)$. Note that, due to the regularity of $\partial\Omega$, $W^{1,p}(\Omega)$ is compactly embedded in $C(\Omega)$, cf. e.g. [16, Theorem 6.3]. Thus, there exists a continuous representative of c_{d}, which we denote by the same symbol. We note that we will later introduce a relaxation of the bilevel Kantorovich problem by adding the cost function c as an additional optimization variable and penalizing its deviation from c_d in the $W^{1,p}$-norm. This provides additional flexibility for the construction of a recovery sequence and serves as a preparation for the finite dimensional case which is currently under investigation and will be discussed in [11].

Bilevel Objective

The functional $\mathcal{J} : \mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1) \to \mathbb{R}$ is supposed to be lower semicontinuous w.r.t. weak-$*$ convergence and bounded on bounded sets, i.e., for every $M > 0$, it holds that

$$\sup\{|\mathcal{J}(\pi, \mu_1)| : \|(\pi, \mu_1)\|_{\mathfrak{M}(\Omega)\times\mathfrak{M}(\Omega_1)} \leq M\} < \infty. \tag{3}$$

At the very end of the paper, several examples for objectives fulfilling these assumptions will be given.

2 Preliminaries and Known Results

In comparison to the bilevel problem discussed in [10], we consider a slightly different problem involving an additional constraint on the distance of $\mathrm{supp}(\mu_1)$ to the boundary

$\partial\Omega_1$. This constraint is necessary to ensure that the mass of the mollified marginals does not leave the domain in order to ensure their weak* convergence, see Lemma 3.1, but can be avoided by passing on to an equivalent problem on an enlarged domain, see Lemma 2.5 below. Our bilevel problem including the additional constraint with distance parameter $\rho \geq 0$ reads as follows:

$$\left.\begin{aligned} \inf_{\pi,\mu_1} \quad & \mathcal{J}(\pi,\mu_1) \\ \text{s.t.} \quad & \mu_1 \in \mathfrak{P}(\Omega_1), \quad \operatorname{dist}(\operatorname{supp}(\mu_1), \partial\Omega_1) \geq \rho, \\ & \pi \in \arg\min\left\{\int_\Omega c_\mathrm{d}\, \mathrm{d}\varphi \colon \varphi \in \Pi(\mu_1, \mu_2^\mathrm{d}),\ \varphi \geq 0\right\}. \end{aligned}\right\} \tag{BK$_\rho$}$$

On the one hand, choosing $\rho = 0$, we recover our original bilevel problem from [10]. If, on the other hand, ρ is too large, it may happen that there is no μ_1 that satisfies constraints in (BK$_\rho$). Therefore, in the following, we tacitly assume that we can find some $\rho > 0$ in a way that the feasible set of (BK$_\rho$) is non-empty. Let us again mention that the lower-level problem only admits solutions provided that μ_1 is a probability measure like μ_2^d. To show the existence of a solution to this bilevel problem, we need the following result:

Lemma 2.1 *The set*

$$\mathcal{M} := \{\mu_1 \in \mathfrak{M}(\Omega_1) \colon \mu_1 \geq 0,\ \operatorname{dist}(\operatorname{supp}(\mu_1), \partial\Omega_1) \geq \rho\} \tag{4}$$

is closed w.r.t. weak-$*$ convergence.

Proof Let a sequence $\{\mu_1^n\}_{n\in\mathbb{N}} \subset \mathcal{M}$ be given such that $\mu_1^n \rightharpoonup^* \mu_1$ in $\mathfrak{M}(\Omega_1)$. It is clear that weak-$*$ convergence gives $\mu_1 \geq 0$. For the remaining claim, we argue by contradiction and assume that there is an $x \in \operatorname{supp}(\mu_1)$ with

$$x \notin A := \{\xi \in \Omega_1 \colon \operatorname{dist}(\xi, \partial\Omega_1) \geq \rho\}.$$

Due to the Lipschitz continuity of the distance function, A is closed and thus there is an $r > 0$ such that $B(x,r) \cap A = \emptyset$. By Urysohn's lemma, there is thus a continuous function $\phi \in C(\Omega_1; [0,1])$ such that $\phi \equiv 0$ on A and $\phi \equiv 1$ on $\overline{B(x,r/2)} \cap \Omega_1$. Since $x \in \operatorname{supp}(\mu_1)$, this yields the desired contradiction:

$$0 < \mu_1(B(x,r/2)) \leq \int_{\Omega_1} \phi(\xi)\, \mathrm{d}\mu_1(\xi) = \lim_{n\to\infty} \int_{\Omega_1} \phi(\xi)\, \mathrm{d}\mu_1^n(\xi) = 0,$$

where the last equality follows from $\phi \equiv 0$ on $A \supset \operatorname{supp}(\mu_1^n)$. □

Proposition 2.2 *Under our standing assumptions, the bilevel Kantorovich problem (BK$_\rho$) admits at least one globally optimal solution.*

Proof In [10, Theorem 3.2], the assertion is shown by means of the stability of transport plans according to [2, Theorem 5.20] for an analogous bilevel problem without the additional constraint $\operatorname{dist}(\operatorname{supp}(\mu_1), \partial\Omega_1) \geq \rho$. Since the set $\mathcal{M}$ from (4) is weakly-$*$ closed as seen in Lemma 2.1, the proof of [10, Theorem 3.2] readily carries over to (BK_ρ). □

Since we do not only need to mollify μ_1 but also μ_2^{d}, we have to require an additional assumption on μ_2^{d} similar to the constraint on μ_1 in (BK_ρ).

Assumption 2.3 There is a constant $\rho > 0$ such that $\operatorname{dist}(\operatorname{supp}(\mu_2^{\mathrm{d}}), \partial\Omega_2) \geq \rho$.

As indicated above, by passing on to an enlarged domain, Assumption 2.3 as well as the additional constraint in (BK_ρ) can be avoided. For this purpose, let $\rho > 0$ be as above and define $\Omega_i^\rho := \Omega_i + \overline{B(0,\rho)}$, $i = 1, 2$ and $\Omega^\rho := \Omega_1^\rho \times \Omega_2^\rho$. Since Ω has a Lipschitz boundary by assumption, there exists an extension of c_{d} to Ω^ρ with the same regularity, which we denote by $c_{\mathrm{d}}^\rho \in W^{1,p}(\Omega^\rho)$, see [17, Theorem 7.25]. For the upper level objective, we assume

Assumption 2.4 The objective $\mathcal{J}$ admits an extension to $\Omega^\rho \times \Omega_1^\rho$ with the same properties as $\mathcal{J}$, i.e., there exists a functional $\mathcal{J}^\rho : \mathfrak{M}(\Omega^\rho) \times \mathfrak{M}(\Omega_1^\rho) \to \mathbb{R}$ that is lower semicontinuous w.r.t. weak-$*$ convergence and bounded on bounded sets and fulfills $\mathcal{J}^\rho(\pi, \mu_1) = \mathcal{J}(\pi|_\Omega, \mu_1|_{\Omega_1})$ for all $(\pi, \mu_1) \in \mathfrak{M}(\Omega^\rho) \times \mathfrak{M}(\Omega_1^\rho)$ with $\operatorname{supp}(\pi) \subseteq \Omega$ and $\operatorname{supp}(\mu_1) \subseteq \Omega_1$.

Herein, $\pi|_\Omega \in \mathfrak{M}(\Omega)$ and $\mu_1|_{\Omega_1} \in \mathfrak{M}(\Omega_1)$ denote the restrictions of π and μ_1. Moreover, we denote the trivial extension of μ_2^{d} to Ω_2^ρ by $\mu_2^\rho \in \mathfrak{P}(\Omega_2^\rho)$, i.e., $\mu_2^\rho(B) := \mu_2^{\mathrm{d}}(B \cap \Omega_2)$ for all $B \in \mathfrak{B}(\Omega_2^\rho)$.

Lemma 2.5 *Consider a bilevel Kantorovich problem of the form*

$$\left.\begin{aligned} &\inf_{\pi,\mu_1} \quad \mathcal{J}(\pi, \mu_1) \\ &\text{s.t.} \quad \mu_1 \in \mathfrak{P}(\Omega_1),\ \pi \in \arg\min\left\{\int_\Omega c_{\mathrm{d}}\,\mathrm{d}\varphi : \varphi \in \Pi(\mu_1, \mu_2^{\mathrm{d}}),\ \varphi \geq 0\right\}, \end{aligned}\right\} \tag{5}$$

where $\mu_2^{\mathrm{d}} \in \mathfrak{P}(\Omega_2)$ need not necessarily satisfy Assumption 2.3. If $\mathcal{J}$ satisfies Assumption 2.4, then (5) is equivalent to

$$\left.\begin{aligned} &\inf_{\pi,\mu_1} \quad \mathcal{J}^\rho(\pi, \mu_1) \\ &\text{s.t.} \quad \mu_1 \in \mathfrak{P}(\Omega_1^\rho), \quad \operatorname{dist}(\operatorname{supp}(\mu_1), \partial\Omega_1^\rho) \geq \rho, \\ &\qquad\ \pi \in \arg\min\left\{\int_{\Omega^\rho} c_{\mathrm{d}}^\rho\,\mathrm{d}\varphi : \varphi \in \Pi_\rho(\mu_1, \mu_2^\rho),\ \varphi \geq 0\right\}, \end{aligned}\right\} \tag{6}$$

where

$$\begin{aligned} \Pi_\rho(\mu_1, \mu_2^\rho) := \{\pi \in \mathfrak{M}(\Omega^\rho) : \pi(A \times \Omega_2^\rho) = \mu_1(A) \ \ \forall A \in \mathfrak{B}(\Omega_1^\rho), \\ \pi(\Omega_1^\rho \times B) = \mu_2^\rho(B) \ \ \forall B \in \mathfrak{B}(\Omega_2^\rho)\}. \end{aligned}$$

The above equivalence means that, if μ_1 and π solve (5), their extensions by zero, denoted by μ_1^ρ and π^ρ, also solve (6), whereas, if μ_1 and π solve (6), their restrictions $\mu_1|_{\Omega_1}$ and $\pi|_\Omega$ are solutions of (5), each time with the same optimal value.

Proof If μ_1 and π are feasible for (5), then μ_1^ρ and π^ρ satisfy $\operatorname{dist}(\operatorname{supp}(\mu_1^\rho), \partial\Omega_1^\rho) \geq \rho$ and $\pi^\rho \in \Pi_\rho(\mu_1^\rho, \mu_2^\rho)$ by construction. Vice versa, if μ_1 and π are feasible for (6), then $\mu_1|_{\Omega_1}$ and $\pi|_\Omega$ satisfy $\pi|_\Omega \in \Pi(\mu_1|_{\Omega_1}, \mu_2^{\mathrm{d}})$, since $\pi \in \Pi_\rho(\mu_1, \mu_2^\rho)$ implies $\operatorname{supp}(\pi) \subset \operatorname{supp}(\mu_1) \times \operatorname{supp}(\mu_2^\rho) = \operatorname{supp}(\mu_1|_{\Omega_1}) \times \operatorname{supp}(\mu_2^{\mathrm{d}}) \subset \Omega_1 \times \Omega_2$. In addition, if μ_1 satisfies the constraint on its support in (6), then

$$\int_{\Omega^\rho} c_{\mathrm{d}}^\rho \,\mathrm{d}\varphi = \int_\Omega c_{\mathrm{d}} \,\mathrm{d}\varphi \quad \forall \varphi \in \Pi_\rho(\mu_1, \mu_2^\rho),$$

where we again used that $\operatorname{supp}(\varphi) \subset \Omega_1 \times \Omega_2$ for all feasible transport plans. Thus the objectives of the lower-level problems in (5) and (6) are the same, and consequently, the same holds for the feasible sets of (5) and (6) (after extension by zero and restriction, respectively). The claim of the lemma then follows, because the upper-level objectives are identical on $\mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1)$ by Assumption 2.4. □

Lemma 2.5 shows that one can equivalently solve (6) instead of (5), provided that Assumption 2.4 holds true. This has the advantage that (6) satisfies an assumption analogous to Assumption 2.3 and guarantees the additional constraint on the distance of $\operatorname{supp}(\mu_1)$ to the boundary. Since the extensions $\mathcal{J}^\rho$ and c_{d}^ρ satisfy our standing assumptions on upper level objective and transportation costs from Sect. 1.1 on the enlarged domain by Assumption 2.4, we can therefore turn to (6) instead of (5). In Sect. 4, we will state two examples, whose upper level objectives fulfill Assumption 2.4. However, in order to keep the discussion concise, we will always focus on the case where Assumption 2.3 is satisfied in the rest of this work and we will tacitly assume it in the following.

2.1 Quadratic Regularization

We now turn to the quadratic regularization of (BK_ρ). Let us first introduce the regularized lower-level problem. Given a regularization parameter $\gamma > 0$, two marginals $\mu_1 \in L^2(\Omega_1)$, $\mu_2 \in L^2(\Omega_2)$, and a cost function $c \in L^2(\Omega)$, we consider the following regularized counterpart to (KP):

$$\left.\begin{aligned}
\inf_{\pi_\gamma} \quad & \mathcal{K}_\gamma(\pi_\gamma) := \int_\Omega c(x)\,\pi_\gamma(x)\,\mathrm{d}\lambda(x) + \tfrac{\gamma}{2}\|\pi_\gamma\|_{L^2(\Omega)}^2 \\
\text{s.t.} \quad & \pi_\gamma \in L^2(\Omega), \quad \pi_\gamma \geq 0 \quad \lambda\text{-a.e. in } \Omega, \\
& \int_{\Omega_2} \pi_\gamma(x_1, x_2)\mathrm{d}\lambda_2(x_2) = \mu_1(x_1) \quad \lambda_1\text{-a.e. in } \Omega_1, \\
& \int_{\Omega_1} \pi_\gamma(x_1, x_2)\mathrm{d}\lambda_1(x_1) = \mu_2(x_2) \quad \lambda_2\text{-a.e. in } \Omega_2.
\end{aligned}\right\} \qquad (\mathrm{KP}_\gamma)$$

Lemma 2.6 *([8, Lemma 2.1]) Problem (KP$_\gamma$) admits a solution if and only if $\mu_i \geq 0$ λ_i-a.e. in Ω_i, $i = 1, 2$, and $\|\mu_1\|_{L^1(\Omega_1)} = \|\mu_2\|_{L^1(\Omega_2)}$. If a solution exists, then it is unique.*

Thanks to the above lemma, we can define the solution operator to (KP$_\gamma$):

$$\mathcal{S}_\gamma : L^2(\Omega) \times \mathcal{M}_0(\Omega) \ni (c, \mu_1, \mu_2) \mapsto \pi_\gamma \in L^2(\Omega), \tag{7}$$

where

$$\mathcal{M}_0(\Omega) := \Big\{(\mu_1, \mu_2) \in L^2(\Omega_1) \times L^2(\Omega_2) \colon \begin{array}{l} \|\mu_1\|_{L^1(\Omega_1)} = \|\mu_2\|_{L^1(\Omega_2)}, \\ \mu_i \geq 0 \lambda_i\text{-a.e. in } \Omega_i, i = 1, 2 \end{array}\Big\}. \tag{8}$$

What is more, if there exist constants $\underline{c} > -\infty$ and $\delta > 0$ such that $c \geq \underline{c}$ λ-a.e. in Ω and $\mu_i \geq \delta$ λ_i-a.e. in Ω_i, $i = 1, 2$, then the dual problem to (KP$_\gamma$) admits a solution, too, see [8, Theorem 2.11] and Lemma 4.2 below. Similarly to the original Kantorovich problem in (KP), this dual problem leads to a significant reduction of the dimension, since it is an optimization problem in $L^2(\Omega_1) \times L^2(\Omega_2)$ rather than in $L^2(\Omega_1 \times \Omega_2)$.

In order to ensure the existence of solutions to (KP$_\gamma$) as well as the associated dual variables for two given marginals $\mu_i \in \mathfrak{M}(\Omega_i)$, $i = 1, 2$, we introduce the convolution and constant shifting operators

$$\mathcal{T}_i^\delta : \mathfrak{M}(\Omega_i) \ni \mu_i \mapsto \varphi_i^\delta * \mu_i + \frac{\delta}{|\Omega_i|} \in L^2(\Omega_i), \quad i = 1, 2. \tag{9}$$

Herein, $\delta > 0$ is a smoothing parameter, $\varphi_i^\delta \in C_c^\infty(\mathbb{R}^{d_i})$ denotes the (symmetric) standard mollifier with $\|\varphi_i^\delta\|_{L^1(\mathbb{R}^{d_i})} = 1$ and support in $\overline{B_i(0, \delta)} \subset \mathbb{R}^{d_i}$, $i = 1, 2$. As a corollary of the classical convergence result for convolution of measures on the whole space, see e.g. [18, Theorem 4.2.2], we obtain the following result. It illustrates the utility of Assumption 2.3 and the additional constraint on μ_1 in (BK$_\rho$).

Lemma 2.7 *Let $\Lambda \subset \mathbb{R}^d$, $d \in \mathbb{N}$, be compact and assume that sequences $\{\mu_n\}_{n \in \mathbb{N}} \subset \mathfrak{M}(\Lambda)$ and $\{\delta_n\}_{n \in \mathbb{N}} \subset \mathbb{R}^+$ are given such that $\mu_n \rightharpoonup^* \mu$, $\delta_n \searrow 0$, and*

$$\operatorname{dist}(\operatorname{supp}(\mu_n), \partial\Lambda) \geq \rho > 0 \quad \forall n \in \mathbb{N}. \tag{10}$$

Then $\varphi^{\delta_n} * \mu_n \rightharpoonup^* \mu$ in $\mathfrak{M}(\Lambda)$ as $n \to \infty$.

Proof Let $v \in C(\Lambda)$ be arbitrary and denote $\Lambda_\rho := \{x \in \Lambda \colon \operatorname{dist}(x, \partial\Lambda) \geq \rho\}$. Then Fubini's theorem along with (10) yields

$$\int_\Lambda v(x)(\varphi^{\delta_n} * \mu_n)(x)\,\mathrm{d}x = \int_{\Lambda_\rho} (\varphi^{\delta_n} * v)(\xi)\,\mathrm{d}\mu_n(\xi) \to \int_\Lambda v(\xi)\mathrm{d}\mu(\xi),$$

where we used the uniform convergence of $\varphi^{\delta_n} * v$ on the compact subset Λ_ρ of $\operatorname{int}(\Lambda)$. □

We note that the restrictions on Ω, namely non-empty interior and Lipschitz boundary, are not required for Lemma 2.7 because $\Lambda_\rho \subset \operatorname{int}(\Lambda)$ is ensured by $\operatorname{dist}(\operatorname{supp}\mu_n, \partial\Lambda) \geq \rho$.

According to Lemma 2.6, (KP_γ) only admits a solution if the total mass of the marginals is the same. In context of the bilevel problem (BK_γ^δ) below, this is ensured for the smoothed marginals by Assumption 2.3 and the additional constraint on $\operatorname{supp}(\mu_1)$, provided that $\delta \leq \rho$. Since these imply $\operatorname{supp}(\varphi_2^\delta * \mu_2^{\mathrm{d}}) \subset \Omega_2$ and $\operatorname{supp}(\varphi_1^\delta * \mu_1) \subset \Omega_1$, respectively, we obtain for every $\mu_1 \in \mathfrak{P}(\Omega_1)$ with $\operatorname{dist}(\operatorname{supp}(\mu_1), \partial\Omega_1) \geq \rho$ that

$$\begin{aligned} \|\mathcal{T}_1^\delta(\mu_1)\|_{L^1(\Omega_1)} &= \int_{\Omega_1} \varphi_1^\delta * \mu_1 \,\mathrm{d}\lambda_1 + \delta \\ &= \|\mu_1\|_{\mathfrak{M}(\Omega_1)} + \delta = \|\mu_2^{\mathrm{d}}\|_{\mathfrak{M}(\Omega_2)} + \delta = \|\mathcal{T}_2^\delta(\mu_2^{\mathrm{d}})\|_{L^1(\Omega_2)} \end{aligned} \tag{11}$$

and consequently, (KP_γ) is well defined with the marginals $\mathcal{T}_1^\delta(\mu_1)$ and $\mathcal{T}_2^\delta(\mu_2^{\mathrm{d}})$. We are now in the position to state the regularized version of (BK_ρ):

$$\left.\begin{aligned} \inf_{\pi_\gamma, \mu_1, c} \quad & \mathcal{J}_\gamma(\pi_\gamma, \mu_1, c) := \mathcal{J}(\pi_\gamma, \mu_1) + \frac{1}{p\gamma}\, \|c - c_{\mathrm{d}}\|^p_{W^{1,p}(\Omega)} \\ \text{s.t.} \quad & c \in W^{1,p}(\Omega), \quad \mu_1 \in \mathfrak{P}(\Omega_1), \quad \operatorname{dist}(\operatorname{supp}(\mu_1), \partial\Omega_1) \geq \rho, \\ & \pi_\gamma = \mathcal{S}_\gamma\big(c, \mathcal{T}_1^\delta(\mu_1), \mathcal{T}_2^\delta(\mu_2^{\mathrm{d}})\big). \end{aligned}\right\} \tag{BK_γ^δ}$$

Here and in the following, with a slight abuse of notation, we denote the measure induced by the L^2-function π_γ by means of the $L^2(\Omega)$-scalar product by the same symbol. In comparison to (BK_ρ), we do not only replace the lower-level Kantorovich problem by its regularized counterpart, but also add the cost function c to the set of optimization variables. This is motivated by the so-called reverse approximation property, which requires a set of optimization variables that is sufficiently rich as also observed, e.g., in case of the optimization of perfect plasticity, see [19]. For this reason, c is treated as an additional optimization variable to have more flexibility at this point. In the companion paper [11], this will be the essential tool to establish the reverse approximation property in finite dimensions. The penalty term in the upper-level objective $\mathcal{J}_\gamma$ will ensure that, in the limit, c equals the given costs c_{d}, see (14) below.

Proposition 2.8 *Let $\gamma > 0$ and $\delta \in (0, \rho]$ be fixed. There exists at least one globally optimal solution to the regularized bilevel Kantorovich problem (BK_γ^δ).*

Proof The existence of solutions for a slightly different problem has been shown in [10, Theorem 4.7], which differs from (BK_γ^δ) as follows: First, the bilevel problem in [10] does not contain the additional constraint on $\operatorname{supp}(\mu_1)$, but, similarly to the proof of Proposition 2.2, this constraint can easily be incorporated into the existence proof using Lemma 2.1. Secondly, the bilevel problem is posed in $\Omega_1^\delta \times \Omega_2^\delta$ with $\Omega_i^\delta := \Omega_i + \overline{B(0,\delta)}$, $i = 1, 2$. This ensures that the marginals $\mathcal{T}_1^\delta(\mu_1)$ and $\mathcal{T}_2^\delta(\mu_2^{\mathrm{d}})$ have the same total mass. In our case, however, this is guaranteed by Assumption 2.3 and the constraint on $\operatorname{supp}(\mu_1)$ together with $\delta \leq \rho$, see (11). With the equality of the total mass of the marginals at hand, the remaining part of the existence proof is then completely along the lines of [10, Theorem 4.7]. □

Remark 2.9 The restrictions on Ω other than compactness, in particular the Lipschitz boundary, are only required in order to ensure the (compact) embedding $W^{1,p}(\Omega) \hookrightarrow C(\Omega)$ that is used in Proposition 2.2 ([10, Theorem 4.7]). If one can drop the term $\frac{1}{p\gamma}\|c - c_\mathrm{d}\|^p_{W^{1,p}(\Omega)}$ and c as an optimization variable, all arguments up to this point are valid under the assumption that Ω_1 and Ω_2 are compact.

3 Convergence for Vanishing Regularization

In the following, we investigate the behavior of optimal solutions of the regularized bilevel problem ($\mathrm{BK}_\gamma^\delta$) for regularization parameters γ, δ tending to zero. For this purpose, assume that we are given sequences of non-negative regularization and smoothing parameters $\{\gamma_n\}_{n\in\mathbb{N}}, \{\delta_n\}_{n\in\mathbb{N}} \subset \mathbb{R}_+$ satisfying $\gamma_n, \delta_n \searrow 0$ as well as

$$0 < \delta_n \le \rho \quad \text{for all } n \in \mathbb{N} \quad \text{and} \quad \frac{\gamma_n}{\delta_n^d} \to 0 \text{ as } n \to \infty. \tag{12}$$

The reason for the coupling of the parameters δ_n and γ_n will become clear in the proof of Proposition 3.3 below. To shorten the notation, we write (BK_n) instead of ($\mathrm{BK}_{\gamma_n}^{\delta_n}$) for the regularized bilevel problem associated with γ_n and δ_n. Similarly, from now on, we equip all entities and variables that depend on either γ_n or δ_n (or both) only with the index n, i.e., e.g. $\mathcal{S}_n{:=}\mathcal{S}_{\gamma_n}$, $\mathcal{T}_1^n{:=}\mathcal{T}_1^{\delta_n}$ and so on. For each $n \in \mathbb{N}$, Proposition 2.8 ensures the existence of a solution $(\bar\pi_n, \bar\mu_1^n, \bar c_n)$ to (BK_n). Owing to the feasibility of $\bar\mu_1^n$, we find that $\|\bar\mu_1^n\|_{\mathfrak{M}(\Omega_1)} = 1$. Moreover, the constraints in (KP_γ) imply

$$\begin{aligned} \|\bar\pi_n\|_{\mathfrak{M}(\Omega)} &= \int_{\Omega_1}\int_{\Omega_2} \bar\pi_n \,\mathrm{d}\lambda_2\,\mathrm{d}\lambda_1 \\ &= \int_{\Omega_1} \mathcal{T}_1^n(\bar\mu_1^n)\,\mathrm{d}\lambda_1 = \|\varphi_1^n\|_{L^1(\mathbb{R}^{d_1})}\|\bar\mu_1^n\|_{\mathfrak{M}(\Omega_1)} + \delta_n \le 1 + \rho \end{aligned} \tag{13}$$

for all $n \in \mathbb{N}$, where we make use of $\bar\pi_n \ge 0$, (11), and (12). Hence, there is a subsequence (denoted by the same symbol to ease notation) such that

$$(\bar\pi_n, \bar\mu_1^n) \rightharpoonup^* (\bar\pi, \bar\mu_1) \quad \text{in } \mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1) \quad \text{as } n \to \infty.$$

Furthermore, take an arbitrary, but fixed $\mu_1 \in \mathfrak{P}(\Omega_1)$ and consider the regularized optimal transport plans $\pi_n = \mathcal{S}_n(c_\mathrm{d}, \mathcal{T}_1^n(\mu_1), \mathcal{T}_2^n(\mu_2^\mathrm{d}))$ for $n \in \mathbb{N}$. Then, the triple $(\pi_n, \mu_1, c_\mathrm{d})$ is feasible for (BK_n) and $(\pi_n, \mu_1)_{n\in\mathbb{N}}$ is bounded in $\mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1)$, cf. (13). The optimality of $(\bar\pi_n, \bar\mu_1^n, \bar c_n)$ for (BK_n) thus yields

$$\|\bar c_n - c_\mathrm{d}\|^p_{W^{1,p}(\Omega)} \le p\,\gamma_n\big(\mathcal{J}(\pi_n, \mu_1) - \mathcal{J}(\bar\pi_n, \bar\mu_1^n)\big) \le \gamma_n\, C$$

for all $n \in \mathbb{N}$ with some constant $C > 0$, because $|\mathcal{J}|$ is bounded on bounded sets by assumption. Hence, we obtain for the whole sequence $\{\bar{c}_n\}_{n\in\mathbb{N}}$ (and not just the subsequence) that

$$\bar{c}_n \to c_{\mathrm{d}} \quad \text{in } W^{1,p}(\Omega) \text{ as } n \to \infty. \tag{14}$$

Now that we have found an accumulation point $(\bar{\pi}, \bar{\mu}_1, c_{\mathrm{d}})$ of the sequence of regularized solutions $\{(\bar{\pi}_n, \bar{\mu}_1^n, \bar{c}_n)\}_{n\in\mathbb{N}}$, we aim to show its optimality for the original bilevel Kantorovich problem (BK$_\rho$). We start with the feasibility of $(\bar{\pi}, \bar{\mu}_1)$ for (BK$_\rho$) in the next subsection.

3.1 Feasibility of the Limit Plan

This subsection is devoted to showing that the accumulation point $(\bar{\pi}, \bar{\mu}_1)$ is feasible for the bilevel problem (BK$_\rho$). This, in particular, requires to show that it is not only feasible for (KP) but also optimal w.r.t. the cost c_d. We start with its feasibility in the following lemma.

Lemma 3.1 *Let $\{(\pi_n, \mu_1^n, c_n)\}_{n\in\mathbb{N}} \subset \mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1) \times W^{1,p}(\Omega)$ be a sequence of feasible points for the regularized bilevel problems (BK$_n$), $n \in \mathbb{N}$. If $(\pi, \mu_1) \in \mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1)$ is a weak-$*$ accumulation point of $\{(\pi_n, \mu_1^n)\}$, then π is a non-negative coupling between μ_1 and μ_2^{d}, i.e., $\pi \in \Pi(\mu_1, \mu_2^{\mathrm{d}})$.*

Proof In order to avoid double subscripts, we assume w.l.o.g. that the whole sequence converges. The non-negativity of π_n carries over to the weak-$*$ limit π. It remains to show that π is a coupling of μ_1 and μ_2^{d}. For this purpose, let $\phi_1 \in C(\Omega_1)$ be arbitrary but fixed. Then, the equality constraints in (KP$_\gamma$) imply

$$\begin{aligned}
\langle \pi, \phi_1 \circ P_1 \rangle &= \lim_{n\to\infty} \langle \pi_n, \phi_1 \circ P_1 \rangle \\
&= \lim_{n\to\infty} \int_{\Omega_1} \phi_1(x_1) \int_{\Omega_2^n} \pi_n(x_1, x_2)\, \mathrm{d}\lambda_2(x_2)\, \mathrm{d}\lambda_1(x_1) \\
&= \lim_{n\to\infty} \int_{\Omega_1} \phi_1(x_1)\, \mathcal{T}_1^n(\mu_1^n)(x_1)\, \mathrm{d}\lambda_1(x_1) \\
&= \lim_{n\to\infty} \int_{\Omega_1} \phi_1(x_1)(\varphi_1^n * \mu_1^n)(x_1)\, \mathrm{d}\lambda_1(x_1) + \frac{\delta_n}{|\Omega_1|} \int_{\Omega_1} \phi_1(x_1)\, \mathrm{d}\lambda_1(x_1) \\
&= \langle \mu_1, \phi_1 \rangle,
\end{aligned}$$

where we use Lemma 2.7 and the additional constraint on $\operatorname{supp}(\mu_1)$ in (KP$_\gamma$) for the passage to the limit. Since ϕ_1 was arbitrary, this implies $P_{1\#}\pi = \mu_1$ as desired. An analogous argument for an arbitrary $\phi_2 \in C(\Omega_2)$ shows $P_{2\#}\pi = \mu_2^{\mathrm{d}}$. □

We now come to a result which is required to show that the weak* cluster point $\bar{\pi}$ is optimal for (KP). Its proof is based on the gluing lemma for measures and the equivalence of convergence in the Wasserstein 1-metric and weak* convergence of measures on compact sets.

Lemma 3.2 *Let $\{\mu_1^n\}_{n\in\mathbb{N}} \subset \mathfrak{P}(\Omega_1)$ be given such that $\mu_1^n \rightharpoonup^* \mu_1$. Moreover, let $\pi \in \Pi(\mu_1, \mu_2^{\mathrm{d}})$ be a non-negative coupling between μ_1 and μ_2^{d}. Then, there exists a sequence of non-negative couplings $\pi_n \in \Pi(\mu_1^n, \mu_2^{\mathrm{d}})$ that converges weakly-$*$ to π.*

Proof The proof relies on the gluing lemma in combination with the equivalence of weak-$*$ convergence and convergence in the Wasserstein-1-distance on compact domains. We first note that, according to [2, Theorem 4.1], for each $n \in \mathbb{N}$, there exists an optimal coupling $\theta_n \in \Pi(\mu_1^n, \mu_1)$ between μ_1^n and μ_1 with respect to the metric cost $|x_1 - y_1|$ on $\Omega_1 \times \Omega_1$. Furthermore, thanks to the gluing lemma, see, e.g., [4, Lemma 5.5], there exist non-negative measures $\sigma_n \in \mathfrak{M}(\Omega_1 \times \Omega_1 \times \Omega_2)$ such that $P_{12\#}\sigma_n = \theta_n$ and $P_{23\#}\sigma_n = \pi$ for all $n \in \mathbb{N}$. Here and in the following,

$$P_{jk}\colon \Omega_1 \times \Omega_1 \times \Omega_2 \to \Omega_1 \times \Omega_\ell, \quad j, k = 1, 2, 3,\ j < k,\ \ell = k - 1$$

refers to the projection onto the coordinates j and k. Using this projection, we define

$$\pi_n := P_{13\#}\sigma_n \in \mathfrak{M}(\Omega_1 \times \Omega_2).$$

Then, by construction, we obtain for all $B_1 \in \mathfrak{B}(\Omega_1)$

$$\begin{aligned}(P_{1\#}\pi_n)(B_1) &= \sigma_n\big(P_{13}^{-1}(B_1 \times \Omega_2)\big)\\ &= \sigma_n(B_1 \times \Omega_1 \times \Omega_2)\\ &= \sigma_n\big(P_{12}^{-1}(B_1 \times \Omega_1)\big) = (P_{1\#}\theta_n)(B_1) = \mu_1^n(B_1)\end{aligned}$$

and analogously, for all $B_2 \in \mathfrak{B}(\Omega_2)$,

$$(P_{2\#}\pi_n)(B_2) = \sigma_n(\Omega_1 \times \Omega_1 \times B_2) = (P_{2\#}\pi)(B_2) = \mu_2^{\mathrm{d}}(B_2)$$

so that $\pi_n \in \Pi(\mu_1^n, \mu_2^{\mathrm{d}})$ as desired. Moreover, the non-negativity of π_n directly follows from the non-negativity of σ_n.

To show the weak-$*$ convergence, we borrow an argument from the proof of [20, Theorem 3.1]. For this purpose, define the mapping

$$P_{1323}\colon \Omega_1 \times \Omega_1 \times \Omega_2 \ni (x_1, y_1, x_2) \mapsto \big((x_1, x_2), (y_1, x_2)\big) \in \Omega \times \Omega,$$

as well as $\zeta := P_{1323\#}\sigma_n$. We observe that $\zeta \in \mathfrak{M}(\Omega \times \Omega)$ and

$$(P_{1\#}\zeta)(B) = \zeta(B \times \Omega) = \sigma_n\big(P_{1323}^{-1}(B \times \Omega)\big) = \sigma_n\big(P_{13}^{-1}(B)\big) = \pi_n(B)$$

as well as

$$\big(P_{2\#}\zeta\big)(B) = \zeta(\Omega \times B) = \sigma_n\big(P_{1323}^{-1}(\Omega \times B)\big) = \sigma_n\big(P_{23}^{-1}(B)\big) = \pi(B)$$

for all $B \in \mathfrak{B}(\Omega)$ so that $\zeta \in \Pi(\pi_n, \pi)$. Again, the non-negativity of ζ directly follows from the non-negativity of σ_n. Now we have everything at hand to estimate the Wasserstein-1-distance of π_n and π:

$$
\begin{aligned}
0 \leq W_1(\pi_n, \pi) &= \inf_{0 \leq \theta \in \Pi(\pi_n, \pi)} \int_{\Omega \times \Omega} |x - y| \, \mathrm{d}\theta(x, y) \\
&\leq \int_{\Omega \times \Omega} |x - y| \, \mathrm{d}\zeta(x, y) \\
&\leq \int_{\Omega \times \Omega} |(x_1, x_2) - (y_1, y_2)| \, \mathrm{d}(P_{1323\#}\sigma_n)((x_1, x_2), (y_1, y_2)) \\
&= \int_{\Omega_1 \times \Omega_1 \times \Omega_2} |x_1 - y_1| \, \mathrm{d}\sigma_n(x_1, y_1, x_2) \\
&= \int_{\Omega_1 \times \Omega_1} |x_1 - y_1| \, \mathrm{d}(P_{12\#}\sigma_n)(x_1, y_1) \\
&= \int_{\Omega_1 \times \Omega_1} |x_1 - y_1| \, \mathrm{d}\theta_n = W_1(\mu_1^n, \mu_1) \to 0 \quad \text{as } n \to \infty,
\end{aligned}
$$

where we use $\mu_1^n \rightharpoonup^* \mu_1$ by assumption and the equivalence of weak-$*$ convergence and convergence in the Wasserstein-1-distance on compact domains according to [2, Theorem 6.9]. Using this equivalence once more finally yields $\pi_n \rightharpoonup^* \pi$ in $\mathfrak{M}(\Omega)$ as $n \to \infty$. □

Having everything together, we can now prove the feasibility of cluster points of the sequence of regularized solutions for the non-regularized bilevel problem.

Proposition 3.3 *Assume that the vanishing sequences $\{\gamma_n\}_{n\in\mathbb{N}}$ and $\{\delta_n\}_{n\in\mathbb{N}}$ satisfy (12). Let $\{(\pi_n, \mu_1^n, c_n)\}_{n\in\mathbb{N}} \subset \mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1) \times W^{1,p}(\Omega)$ be a sequence of feasible points for the regularized bilevel problems (BK_n), $n \in \mathbb{N}$. If $(\pi, \mu_1, c_\mathrm{d}) \in \mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1) \times W^{1,p}(\Omega)$ is an accumulation point of this sequence w.r.t. weak-$*$ convergence in $\mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1)$ and weak convergence in $W^{1,p}(\Omega)$, then (π, μ_1) is feasible for (BK_ρ), i.e., $\mu_1 \in \mathfrak{P}(\Omega_1)$, $\mathrm{dist}(\mathrm{supp}(\mu_1), \partial\Omega_1) \geq \rho$, and π is optimal for (KP) with respect to the marginals μ_1 and μ_2^d as well as the cost function c_d.*

Proof In order to avoid double subscripts, we again assume w.l.o.g. that the whole sequence converges. Since $\mathfrak{P}(\Omega_1)$ as well as the set $\mathcal{M}$ from (4) are weakly-$*$ closed, see Lemma 2.1, the properties for μ_1 follow immediately.

As we have already seen in Lemma 3.1, π is feasible for the Kantorovich problem (KP) with respect to μ_1 and μ_2^d. So, it suffices to show the optimality of π for (KP). To this end, recall the lower-level problems from the feasible sets of (BK_n) that are solved by π_n:

$$\left.\begin{aligned} \min_{\pi} \quad & \mathcal{K}_n(\pi) := \int_\Omega c_n\, \pi \,\mathrm{d}\lambda + \tfrac{\gamma_n}{2} \|\pi\|^2_{L^2(\Omega)} \\ \text{s.t.} \quad & \pi \in L^2(\Omega), \quad \pi \geq 0 \quad \lambda\text{-a.e. in } \Omega, \\ & \int_{\Omega_2} \pi(x_1, x_2) \mathrm{d}\lambda_2(x_2) = \mathcal{T}_1^n(\mu_1^n)(x_1) \quad \lambda_1\text{-a.e. in } \Omega_1, \\ & \int_{\Omega_1} \pi(x_1, x_2) \mathrm{d}\lambda_1(x_1) = \mathcal{T}_2^n(\mu_2^{\mathrm{d}})(x_2) \quad \lambda_2\text{-a.e. in } \Omega_2. \end{aligned}\right\} \tag{KP$_n$}$$

By [2, Theorem 4.1], we know that there is at least one solution of the Kantorovich problem (KP) (associated with μ_1, μ_2^{d}, and the limiting cost function c_{d}). Let us consider an arbitrary solution $\pi^* \in \mathfrak{M}(\Omega)$ of (KP). Owing to Lemma 3.2, there exists a sequence of non-negative couplings $\{\pi_n^*\}_{n\in\mathbb{N}}$ between μ_1^n and μ_2^{d} that converges weakly-$*$ to π^*. We then define

$$\begin{aligned} &\varphi_n(x_1, x_2) := \varphi_1^n(x_1)\, \varphi_2^n(x_2), \quad (x_1, x_2) \in \Omega, \\ \text{and} \quad &\vartheta_n^* := \varphi_n * \pi_n^* + \frac{\delta_n}{|\Omega_1|\,|\Omega_2|} = \int_\Omega \varphi_n(\xi - \cdot\,)\, \mathrm{d}\pi_n^*(\xi) + \frac{\delta_n}{|\Omega_1|\,|\Omega_2|} \in L^2(\Omega). \end{aligned}$$

Then, the non-negativity of π_n^* implies the positivity of ϑ_n^*. Moreover, the definition of ϑ_n^* in combination with Fubini's theorem yields

$$\int_{\Omega_2} \vartheta_n^* \mathrm{d}\lambda_2 = \varphi_1^n * \mu_1^n + \frac{\delta_n}{|\Omega_1|} = \mathcal{T}_1^n(\mu_1^n), \tag{15}$$

$$\int_{\Omega_1} \vartheta_n^* \mathrm{d}\lambda_1 = \varphi_2^n * \mu_2^{\mathrm{d}} + \frac{\delta_n}{|\Omega_2|} = \mathcal{T}_2^n(\mu_2^{\mathrm{d}}) \tag{16}$$

so that ϑ_n^* is feasible for (KP$_n$). For the objective of the Kantorovich problem, the optimality of π_n for (KP$_n$) yields

$$\begin{aligned} \langle c_{\mathrm{d}}, \pi \rangle = \lim_{n\to\infty} \langle c_n, \pi_n \rangle &\leq \liminf_{n\to\infty} \int_\Omega c_n\, \pi_n \,\mathrm{d}\lambda + \tfrac{\gamma_n}{2} \|\pi_n\|^2_{L^2(\Omega)} \\ &\leq \limsup_{n\to\infty} \int_\Omega c_n\, \vartheta_n^* \,\mathrm{d}\lambda + \tfrac{\gamma_n}{2} \|\vartheta_n^*\|^2_{L^2(\Omega)}. \end{aligned} \tag{17}$$

Let us investigate the two addends on the right hand side of this inequality separately. Since $c_n \rightharpoonup c_{\mathrm{d}}$ in $W^{1,p}(\Omega)$ and the embedding $W^{1,p}(\Omega) \hookrightarrow C(\Omega)$ is compact due to $p > d$, c_n converges uniformly to c_{d} in Ω. Let us define $\Omega_\rho := \{x \in \Omega\colon\ \mathrm{dist}(x, \partial\Omega) \geq \rho\}$. Then, the uniform convergence of both c_n and the convolution in compact subsets of $\mathrm{int}(\Omega)$ yields that

$$\max_{x\in\Omega_\rho} \left| \int_\Omega c_n(\xi)\, \varphi(x - \xi)\, \mathrm{d}\lambda(\xi) - c_{\mathrm{d}}(x) \right| \;\to\; 0.$$

Since $\text{supp}(\pi_n^*) \subset \text{supp}(\mu_1^n) \times \text{supp}(\mu_2^{\text{d}}) \subset \Omega_\rho$, this in combination with the definition of ϑ_n^* and the weak-$*$ convergence of π_n^* implies

$$\int_\Omega c_n \, \vartheta_n^* \, \mathrm{d}\lambda = \int_{\Omega_\rho} \int_\Omega c_n(x) \, \varphi(x-\xi) \mathrm{d}\lambda(x) \, \mathrm{d}\pi_n^*(\xi) + \frac{\delta_n}{|\Omega_1| \, |\Omega_2|} \int_\Omega c_n(x) \, \mathrm{d}\lambda(x) \to \langle c_{\text{d}}, \pi^* \rangle. \tag{18}$$

For the second addend on the right hand side of (17), we obtain

$$\begin{aligned} \|\vartheta_n^*\|_{L^2(\Omega)}^2 &\le 2 \int_\Omega \left(\int_\Omega \varphi(x-\xi) \, \mathrm{d}\pi_n^*(\xi)^2 \right) \mathrm{d}\lambda(x) + 2\delta_n^2 \\ &\le 2\|\varphi_n\|_{L^2(B(0,\delta_n))}^2 \|\pi_n^*\|_{\mathfrak{M}(\Omega)}^2 + 2\delta_n^2, \end{aligned}$$

where the L^2-norm of the standard mollifier is estimated by

$$\begin{aligned} \|\varphi_n\|_{L^2(B(0,\delta_n))}^2 &= \|\varphi_1^n\|_{L^2(B_1(0,\delta_n))}^2 \, \|\varphi_2^n\|_{L^2(B_2(0,\delta_n))}^2 \\ &\le \prod_{i=1}^2 \|\varphi_i^n\|_{L^\infty(B_i(0,\delta_n))} \, \|\varphi_i^n\|_{L^1(B_i(0,\delta_n))} \le C \, \delta_n^{-d_1-d_2} = C \, \delta_n^{-d} \end{aligned}$$

with a constant $C > 0$. In view of the coupling of γ_n and δ_n in (12) and $\|\pi_n^*\|_{\mathfrak{M}(\Omega)} = 1$ for all $n \in \mathbb{N}$, we thus arrive at

$$\frac{\gamma_n}{2} \, \|\vartheta_n^*\|_{L^2(\Omega)}^2 \le C \Big(\frac{\gamma_n}{\delta_n^d} + \gamma_n \, \delta_n^2 \Big) \to 0.$$

Inserting this together with (18) in (17) implies $\langle c_{\text{d}}, \pi \rangle \le \langle c_{\text{d}}, \pi^* \rangle$ and, since π is feasible for (KP) associated with μ_1, μ_2^{d}, and c_{d}, as seen above, while π^* is optimal for that problem, π is optimal, too, and thus (π, μ_1) is feasible for (BK_ρ) as claimed. □

Note that we only obtain the feasibility of the sequence ϑ_n^* in the above proof, if the marginals are smooth and shifted, see (15) and (16). Therefore, even if the given data μ_2^{d} is a function in $L^2(\Omega_2)$, we cannot use it in the regularized problems, but need to consider $\mathcal{T}_2^n(\mu_2^{\text{d}})$ instead.

Recall the sequence of solutions $(\bar{\pi}_n, \bar{\mu}_1^n, \bar{c}_n)$ to the regularized bilevel problems from the beginning of this section. We already know that this sequence admits a weak-$*$ accumulation point. To be more precise $(\bar{\pi}_n, \bar{\mu}_1^n) \rightharpoonup^* (\bar{\pi}, \bar{\mu}_1)$ in $\mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1)$ after possibly restricting to a subsequence, while the whole sequence of cost functions $\bar{c}_n$ converges strongly in $W^{1,p}(\Omega)$ to c_{d}. Thus, according to Proposition 3.3, the weak-$*$ limit $(\bar{\pi}, \bar{\mu}_1)$ is feasible for (BK_ρ). As this observation holds for every accumulation point, we immediately obtain the following

Corollary 3.4 *Assume that the vanishing sequences $\{\gamma_n\}_{n\in\mathbb{N}}$ and $\{\delta_n\}_{n\in\mathbb{N}}$ satisfy (12). Suppose moreover that (π^*, μ_1^*) is a solution to the bilevel problem (BK_ρ) that admits* *a* recovery sequence *in the following sense: There is a sequence* $(\pi_n^*, \mu_{1,n}^*, c_n^*)_{n\in\mathbb{N}} \subset \mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1) \times W^{1,p}(\Omega)$ *satisfying*

(i) $(\pi_n^*, \mu_{1,n}^*, c_n^*)$ *is feasible for (BK$_n$) for all* $n \in \mathbb{N}$,
(ii) $\limsup_{n\to\infty} \mathcal{J}_n(\pi_n^*, \mu_{1,n}^*, c_n^*) \le \mathcal{J}(\pi^*, \mu_1^*)$.

Then, every weak-∗ accumulation point $(\bar\pi, \bar\mu_1)$ *of the sequence of solutions* $\{(\bar\pi_n, \bar\mu_1^n)\}_{n\in\mathbb{N}}$ *to the regularized bilevel problem is also a solution to (BK$_\rho$).*

Proof As the feasibility of $(\bar\pi, \bar\mu_1)$ has already been established, we only need to prove its optimality, which is a consequence of the existence of a recovery sequence and the weak-∗ lower semicontinuity of $\mathcal{J}$. To this end, we index the weakly-∗ convergent subsequence by the symbol n.

$$\begin{aligned}
\mathcal{J}(\bar\pi, \bar\mu_1) &\le \liminf_{n\to\infty} \mathcal{J}(\bar\pi_n, \bar\mu_1^n) \\
&\le \liminf_{n\to\infty} \mathcal{J}(\bar\pi_n, \bar\mu_1^n) + \frac{1}{p\,\gamma_n}\|\bar c_n - c_{\mathrm{d}}\|^p_{W^{1,p}(\Omega)} \\
&= \liminf_{n\to\infty} \mathcal{J}_n(\bar\pi_n, \bar\mu_1^n, \bar c_n) \\
&\le \limsup_{n\to\infty} \mathcal{J}_n(\pi_n^*, \mu_{1,n}^*, c_n^*) \le \mathcal{J}(\pi^*, \mu_1^*).
\end{aligned} \tag{19}$$

The optimality of (π^*, μ_1^*) gives the result. □

Remark 3.5 Note that the first two inequalities in (19) in essence correspond to the liminf-condition of Γ-convergence of the bilevel objective. To be more precise, if we denote the feasible set of (BK$_n$) by $\mathcal{F}_n \subset \mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1) \times W^{1,p}(\Omega)$ and define

$$\widetilde{\mathcal{J}}_n(\pi, \mu_1, c) := \mathcal{J}_n(\pi, \mu_1, c) + I_{\mathcal{F}_n}(\pi, \mu_1, c),$$

where $I_{\mathcal{F}_n}$ denotes the indicator functional of $\mathcal{F}_n$, then the weak-∗ lower semicontinuity of $\mathcal{J}$ along with Proposition 3.3 gives that, for every sequence $\{(\pi_n, \mu_{1,n}, c_n)\}$ converging weakly-∗ to a point $(\pi, \mu_1, c_{\mathrm{d}})$, there holds

$$\widetilde{\mathcal{J}}(\pi, \mu_1) \le \liminf_{n\to\infty} \widetilde{\mathcal{J}}_n(\pi_n, \mu_{1,n}, c_n),$$

where $\widetilde{\mathcal{J}}(\pi, \mu_1) := \mathcal{J}(\pi, \mu_1) + I_{\mathcal{F}}(\pi, \mu_1)$ and $\mathcal{F}$ denotes the feasible set of (BK$_\rho$). Note moreover that, since we are only interested in the convergence of minimizers, we do not need a recovery sequence realizing the limsup-inequality for every tuple $(\pi, \mu_1) \in \mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1)$, but just for one global minimizer of (BK$_\rho$). Thus the assumptions (i) and (ii) in Corollary 3.4 do not guarantee Γ-convergence of $\tilde{\mathcal{J}}_n$ to $\tilde{\mathcal{J}}$ (w.r.t. weak-∗ convergence), but are sufficient for our purpose.

The crucial task is now of course to establish the existence of a recovery sequence satisfying the conditions (i) and (ii) in Corollary 3.4. So far, unfortunately, we are not able to guarantee the existence of such a sequence in the general setting without further assumptions. If, however, $\mu_2^{\mathrm{d}} \ll \lambda^{d_2}$, $c(x_1, x_2) = h(x_1 - x_2)$ with a strictly convex function h, and $\mathcal{J}$ is even weak-∗ continuous, then a recovery sequence can be constructed as we will see in the next section.

3.2 Reverse Approximation in Case of Strictly Convex Costs and an Absolutely Continuous Marginal

We now describe a scenario in which we can explicitly construct a recovery sequence in the sense of Corollary 3.4.

Theorem 3.6 *Suppose that, in addition to our standing assumptions, the following hold true:*

1. $\Omega_1 = \Omega_2 =: \Omega_*$ *(such that* $d_1 = d_2 =: d_*$*),*
2. $\mu_2^{\mathrm{d}} \ll \lambda_*$ *with* $\lambda_* := \lambda_1 = \lambda_2$*,*
3. $c_{\mathrm{d}}(x_1, x_2) = h(x_1 - x_2)$ *with a function* $h\colon \mathbb{R}^{d_*} \to \mathbb{R}$ *that is strictly convex and even symmetric, i.e.,* $h(-\xi) = h(\xi)$ *for all* $\xi \in \mathbb{R}^{d_*}$*,*
4. *Additionally to its lower semicontinuity,* $\mathcal{J}$ *shall be upper semicontinuous w.r.t. weak-∗ convergence in the first variable, i.e., if* $\pi_n \rightharpoonup^* \pi$ *in* $\mathfrak{M}(\Omega)$*, then, for every* $\mu_1 \in \mathfrak{M}(\Omega_1)$*, there holds*
$$\limsup_{n\to\infty} \mathcal{J}(\pi_n, \mu_1) \le \mathcal{J}(\pi, \mu_1).$$
In other words, $\mathcal{J}$ *shall be continuous in its first variable w.r.t. weak* convergence.*
5. *The null sequences* $\{\gamma_n\}_{n\in\mathbb{N}}$ *and* $\{\delta_n\}_{n\in\mathbb{N}}$ *satisfy (12).*

Then, a sequence of solutions $\{(\bar{\pi}_n, \bar{\mu}_1^n, \bar{c}_n)\}_{n\in\mathbb{N}}$ *of the regularized bilevel problems* (BK_n)*,* $n \in \mathbb{N}$*, converges (up to subsequences) to an optimal solution of* (BK_ρ)*.*

Proof The assertion is a direct consequence of Corollary 3.4 and the uniqueness of the transport plan under the additional assumptions. Let (π^*, μ_1^*) be an optimal solution of the bilevel problem (BK_ρ). According to the generalized version of Brenier's theorem in [2, Theorem 2.44], the additional assumptions 1–3 ensure the existence of an optimal transport map, which in turn yields a unique optimal transport plan that solves the Kantorovich problem with marginals $(\mu_2^{\mathrm{d}}, \mu_1^*)$ and cost $c_{\mathrm{d}}(x_1, x_2)$, i.e.,

$$\min \Big\{ \int_\Omega c_{\mathrm{d}}(x_2, x_1)\, \mathrm{d}\varphi(x_1, x_2) \colon \varphi \in \Pi(\mu_2^{\mathrm{d}}, \mu_1^*),\ \varphi \ge 0 \Big\}.$$

For reasons of symmetry, π is a solution to the above problem if and only if π' defined by $\pi'(B_1 \times B_2) = \pi(B_2 \times B_1)$, $B_1, B_2 \in \mathcal{B}(\Omega_*)$, solves

$$\min \Big\{ \int_\Omega c_{\mathrm{d}}(x_1, x_2)\, \mathrm{d}\varphi(x_1, x_2) \colon \varphi \in \Pi(\mu_1^*, \mu_2^{\mathrm{d}}),\ \varphi \ge 0 \Big\} \tag{KP*}$$

such that the solution set of (KP*) is a singleton, too. Therefore, π^* is the only solution of (KP*).

Let us define the sequence

$$\mu_{1,n}^* := \mu_1^*, \quad c_n^* := c_{\mathrm{d}}, \quad \pi_n^* := \mathcal{S}_n\big(c_{\mathrm{d}}, \mathcal{T}_1^n(\mu_1^*), \mathcal{T}_2^n(\mu_2^{\mathrm{d}})\big), \quad n \in \mathbb{N}. \tag{20}$$

By construction, $(\pi_n^*, \mu_{1,n}^*, c_n^*)$ is feasible for (BK$_n$) for every $n \in \mathbb{N}$. Since the sequence $\{\pi_n^*\}$ is bounded, there is a weakly-$*$ convergent subsequence and, according to Proposition 3.3, its limit, together with μ_1^*, is feasible for (BK$_\rho$), i.e., it is a solution of (KP*). Thus, by uniqueness, said weak-$*$ limit equals π^* and a classical argument by contradiction yields that the whole sequence $(\pi_n^*, \mu_{1,n}^*, c_n^*)$ converges (weakly-$*$) to $(\pi^*, \mu_1^*, c_{\mathrm{d}})$. The presupposed weak-$*$ continuity of $\mathcal{J}$ finally yields that

$$\limsup_{n\to\infty} \mathcal{J}_n(\pi_n^*, \mu_{1,n}^*, c_n^*) = \limsup_{n\to\infty} \mathcal{J}(\pi_n^*, \mu_1^*) \leq \mathcal{J}(\pi^*, \mu_1^*) \quad \text{as } n \to \infty$$

such that $\{(\pi_n^*, \mu_{1,n}^*, c_n^*)\}_{n\in\mathbb{N}}$ is the desired recovery sequence. Corollary 3.4 then yields the claim. □

Remark 3.7 An inspection of the proof of Theorem 3.6 shows that, under the assumptions of this theorem, there is no need for c as additional optimization variable to construct the recovery sequence, as c_n^* is set to c_{d} in (20). Accordingly, the assertion of Theorem 3.6 remains true, if one considers

$$\left.\begin{aligned} \inf_{\pi_\gamma, \mu_1} \quad & \mathcal{J}(\pi_\gamma, \mu_1) \\ \text{s.t.} \quad & \mu_1 \in \mathfrak{P}(\Omega_1), \quad \operatorname{dist}(\operatorname{supp}(\mu_1), \partial\Omega_1) \geq \rho, \\ & \pi_\gamma = \mathcal{S}_\gamma\big(c_{\mathrm{d}}, \mathcal{T}_1^\delta(\mu_1), \mathcal{T}_2^\delta(\mu_2^{\mathrm{d}})\big). \end{aligned}\right\} \tag{BK$^\delta_{\gamma,c_{\mathrm{d}}}$}$$

instead of (BK$_\gamma^\delta$). In the finite-dimensional setting, however, it is exactly the additional optimization variable c, which enables the construction of a recovery sequence, see [11].

Consequently, one may drop the additional optimization over c and the term $\frac{1}{\gamma p}\|c - c_{\mathrm{d}}\|^p_{W^{1,p}(\Omega)}$ in the setting of Theorem 3.6. Then, the restrictions on the domain that ensure the compact embedding $W^{1,p}(\Omega) \hookrightarrow C(\Omega)$ are no longer necessary and one may restrict to Ω_1 and Ω_2 being compact, see also Remark 2.9.

4 Two Application-Driven Examples

The additional assumptions in Theorem 3.6 are certainly rather restrictive, in particular condition 4. Let us therefore end our considerations by giving two examples fulfilling these assumptions.

4.1 Marginal Identification Problem

As an example for a bilevel Kantorovich problem, we have already mentioned the problem of identifying the marginal μ_1 based on measurements of the transport plan in an observation domain $D \subset \Omega$ in Sect. 1. The most natural choice for the upper-level objective probably reads

$$\mathcal{J}(\pi, \mu_1) := \|\pi - \pi^{\mathrm{d}}\|_{\mathfrak{M}(D)} + \nu\, \|\mu_1 - \mu_1^{\mathrm{d}}\|_{\mathfrak{M}(\Omega_*)}, \tag{21}$$

where $\pi^{\mathrm{d}} \in \mathfrak{M}(D)$ denotes the measurement of the transport plan, while $\mu_1^{\mathrm{d}} \in \mathfrak{P}(\Omega_*)$ is a guess for the unknown marginal μ_1. Moreover, $\nu \geq 0$ is a given weighting parameter. However, an objective of this form does not satisfy condition 4 in Theorem 3.6. To ensure this condition, let us assume that D is an open and bounded domain with a Lipschitz boundary. Then, thanks to $p > d$ by our standing assumptions, the embedding $W_0^{1,p}(D) \hookrightarrow C_0(D)$ is compact and so, by Schauder's theorem, $\mathfrak{M}(D)$ embeds compactly in $W^{-1,p'}(D) := W_0^{1,p}(D)^*$, where, as usual, $p' = p/(p-1)$ denotes the conjugate exponent. Note that the restriction operator from $\mathfrak{M}(\Omega)$ to $\mathfrak{M}(D)$ is weak-$*$ continuous, since D is open by assumption. Therefore, for a given $\pi^{\mathrm{d}} \in \mathfrak{M}(D)$, an objective of the form

$$\widetilde{\mathcal{J}}(\pi, \mu_1) := \|\pi - \pi^{\mathrm{d}}\|_{W^{-1,p'}(D)}^{p'} + \nu\, \|\mu_1 - \mu_1^{\mathrm{d}}\|_{\mathfrak{M}(\Omega_*)} \tag{22}$$

fulfills condition 4. A natural extension of $\widetilde{\mathcal{J}}$ fulfilling Assumption 2.4 (in case that Assumption 2.3 is not fulfilled) reads

$$\widetilde{\mathcal{J}}^{\rho}(\pi, \mu_1) := \|\pi - \pi^{\mathrm{d}}\|_{W^{-1,p'}(D)}^{p'} + \nu\, \|\mu_1 - \mu_1^{\mathrm{d},\rho}\|_{\mathfrak{M}(\Omega_*^{\rho})},$$

where Ω_*^{ρ} again denotes the enlarged domain and $\mu_1^{\mathrm{d},\rho}$ is the extension of μ_1^{d} by zero.

The $W^{-1,p'}(D)$-norm can be evaluated with the help of the p-Laplacian. For this purpose, denote by $\psi \in W_0^{1,p}(D)$ the unique solution of

$$-\operatorname{div}(|\nabla\psi|^{p-2}\nabla\psi) = \pi - \pi^{\mathrm{d}} \quad \text{in } W^{-1,p'}(D). \tag{23}$$

Then, it is easily seen that

$$\begin{aligned}
\|\pi - \pi^{\mathrm{d}}\|_{W^{-1,p'}(D)} &= \sup_{v \in W_0^{1,p}(D)} \frac{\langle \pi - \pi^{\mathrm{d}}, v\rangle}{\|\nabla v\|_{L^p(D;\mathbb{R}^d)}} \\
&= \sup_{v \in W_0^{1,p}(D)} \frac{\int_D |\nabla\psi|^{p-2}\nabla\psi \cdot \nabla v \,\mathrm{d}\lambda}{\|\nabla v\|_{L^p(D;\mathbb{R}^d)}} = \|\nabla\psi\|_{L^p(D;\mathbb{R}^d)}^{p-1}
\end{aligned}$$

and hence, the objective from (22) becomes $\|\nabla\psi\|_{L^p(D;\mathbb{R}^d)}^{p} + \nu\, \|\mu_1 - \mu_1^{\mathrm{d}}\|_{\mathfrak{M}(\Omega_*)}$. If one aims to avoid the p-Laplace equation, one can resort to an equivalent norm based on the Poisson equation on D. To this end, let $\eta \in \mathfrak{M}(D)$ be given and consider

$$\varphi \in W_0^{1,p'}(D), \quad -\triangle\varphi = \eta \quad \text{in } W^{-1,p'}(D), \tag{24}$$

where $\langle -\triangle\varphi, v\rangle := \int_D \nabla\varphi \cdot \nabla v \,\mathrm{d}\lambda$, $\varphi \in W_0^{1,p'}(D)$, $v \in W_0^{1,p}(D)$, denotes the Laplace operator.

Lemma 4.1 *Let D be a bounded domain of class C^1. Then there exists an exponent $p > d$ such that, for every $\eta \in \mathfrak{M}(D)$, there exists a unique solution $\varphi \in W_0^{1,p'}(D)$ of*

(24). The associated solution operator denoted by $G\colon \mathfrak{M}(D) \to W_0^{1,p'}(D)$ is linear and compact.

If $d \leq 3$, the result also applies if D is only a bounded Lipschitz domain.

Proof According to [21, Theorem 4.6], there exists a $p > d$ such that the Poisson equation (24) admits a unique solution in $W_0^{1,p}(D)$ for right hand sides in $W^{-1,p}(D)$.[1] The same assertion for Lipschitz domains can be found in [22] for the case $d = 2$ and [23] for $d = 3$. By duality, there is thus a unique solution in $\varphi \in W_0^{1,p'}(D)$ to the state equation in (24) for every right hand side in $W^{-1,p'}(D)$. Due to the continuous embedding $W_0^{1,p}(D) \hookrightarrow C_0(D)$ already mentioned above, there holds $\mathfrak{M}(D) \hookrightarrow W^{-1,p'}(D)$ and we obtain the existence and uniqueness of $\varphi \in W_0^{1,p'}(D)$ for every $\eta \in \mathfrak{M}(D)$. The associated solution operator G is clearly linear and, by Banach's inverse theorem, continuous. The compactness of G follows from the compactness of the embedding $W_0^{1,p}(D) \hookrightarrow C_0(D)$. □

Unfortunately, the $W^{-1,p'}(D)$-norm cannot be expressed by means of the solution operator G, but it holds that

$$\|\eta\|_{W^{-1,p'}(D)} \leq \|\nabla\varphi\|_{L^{p'}(\Omega;\mathbb{R}^d)} \leq \|G\| \, \|\eta\|_{W^{-1,p'}(D)}$$

and therefore,

$$\mathfrak{M}(D) \ni \eta \mapsto \|\nabla G\eta\|_{L^{p'}(\Omega;\mathbb{R}^d)}$$

defines a norm equivalent to the $W^{-1,p'}(D)$-norm. Since G is compact, an objective of the form

$$\widehat{\mathcal{J}}(\pi, \mu_1) := \|\nabla G(\pi - \pi^{\mathrm{d}})\|^{p'}_{L^{p'}(D;\mathbb{R}^d)} + \nu \, \|\mu_1 - \mu_1^{\mathrm{d}}\|_{\mathfrak{M}(\Omega_*)} \tag{25}$$

satisfies the condition 4 in Theorem 3.6. Thus, if we consider the bilevel Kantorovich problem (BK_ρ) with $\widetilde{\mathcal{J}}$ or $\widehat{\mathcal{J}}$, Theorem 3.6 applies. Let us shortly turn to the associated regularized problems. For this purpose, recall the following result from [8]:

Lemma 4.2 *([8, Theorem 2.11]) Consider the regularized Kantorovich problem (KP_γ) with marginals $\mu_i \in L^2(\Omega_i)$, $i = 1, 2$, and a cost function $c \in L^2(\Omega)$. Assume that the marginals satisfy $\|\mu_1\|_{L^1(\Omega_1)} = \|\mu_2\|_{L^1(\Omega_2)}$ and $\mu_i \geq \delta \; \lambda_i$-a.e. in Ω_i, $i = 1, 2$, with a constant $\delta > 0$. Moreover, assume that there exists a constant $\underline{c} > -\infty$ such that $c \geq \underline{c}$ λ-a.e. in Ω. Then, $\pi_\gamma \in L^2(\Omega)$ is a solution of (KP_γ) if and only if there*

[1] On C^1-domains, this even holds for every $p < \infty$, see [21, Theorem 4.6].

exist functions $\alpha_1 \in L^2(\Omega_1)$ *and* $\alpha_2 \in L^2(\Omega_2)$ *satisfying*

$$\pi_\gamma - \frac{1}{\gamma}(\alpha_1 \oplus \alpha_2 - c)_+ = 0 \qquad \lambda\text{-a.e. in } \Omega, \tag{26a}$$

$$\int_{\Omega_2} \pi_\gamma(x_1, x_2) \mathrm{d}\lambda_2(x_2) = \mu_1(x_1) \qquad \lambda_1\text{-a.e. in } \Omega_1, \tag{26b}$$

$$\int_{\Omega_1} \pi_\gamma(x_1, x_2) \mathrm{d}\lambda_1(x_1) = \mu_2(x_2) \qquad \lambda_2\text{-a.e. in } \Omega_2. \tag{26c}$$

Herein, $(\alpha_1 \oplus \alpha_2)(x_1, x_2) := \alpha_1(x_1) + \alpha_2(x_2)$ λ*-a.e. in* Ω *refers to the direct sum of* $\alpha_1 \in L^2(\Omega_1)$ *and* $\alpha_2 \in L^2(\Omega_2)$*, while, for given* $u \in L^2(\Omega)$, $(u)_+(x) := \max\{u(x); 0\}$ λ*-a.e. in* Ω *denotes the pointwise maximum.*

With this result at hand, we can erase the regularized optimal transport plan π_γ from ($\mathrm{BK}^\delta_{\gamma,c_\mathrm{d}}$) by using the necessary and sufficient conditions in (26). Let us exemplarily consider ($\mathrm{BK}^\delta_{\gamma,c_\mathrm{d}}$) with the objective from (25). The corresponding regularized bilevel problem then reads as follows:

$$\left.\begin{aligned}
\min\ & \|\nabla\varphi\|^{p'}_{L^{p'}(D;\mathbb{R}^d)} + \nu\, \|\mu_1 - \mu_1^\mathrm{d}\|_{\mathfrak{M}(\Omega_*)} \\
\text{s.t. } & \alpha_1, \alpha_2 \in L^2(\Omega_*), \quad \mu_1 \in \mathfrak{P}(\Omega_*), \quad \varphi \in W_0^{1,p'}(D), \\
& \mathrm{dist}(\mathrm{supp}(\mu_1), \partial\Omega_*) \geq \rho, \\
& -\triangle\varphi = \tfrac{1}{\gamma}(\alpha_1 \oplus \alpha_2 - c_\mathrm{d})_+ - \pi^\mathrm{d} \quad \text{in } W^{-1,p'}(D), \\
& \int_{\Omega_*} (\alpha_1 \oplus \alpha_2 - c_\mathrm{d})_+(x_1, x_2) \mathrm{d}\lambda_*(x_2) = \gamma\, \mathcal{T}_1^\delta(\mu_1)(x_1) \quad \text{a.e. in } \Omega_*, \\
& \int_{\Omega_*} (\alpha_1 \oplus \alpha_2 - c_\mathrm{d})_+(x_1, x_2) \mathrm{d}\lambda_*(x_1) = \gamma\, \mathcal{T}_2^\delta(\mu_2^\mathrm{d})(x_2) \quad \text{a.e. in } \Omega_*.
\end{aligned}\right\} \tag{27}$$

Remark 4.3 If one replaces the lower-level Kantorovich problem as in (BK_ρ) by its necessary and sufficient (and thus equivalent) optimality conditions, then an optimization problem with complementarity constraints (MPCC) in $\mathfrak{M}(\Omega)$ is obtained, see [10, Section 3]. Problems of this type are challenging, since standard constraint qualifications are typically violated, even in finite dimensions. For this reason, regularization and relaxation techniques are frequently applied, we only refer to [24] and the references therein. In light of the above reformulation based on Lemma 4.2, the quadratic regularization of the Kantorovich problem can be interpreted as a relaxation of the complementarity constraints, too, since it is precisely the Moreau–Yosida regularization of the dual Kantorovich problem as follows the considerations in [8, Section 2.2]. Indeed the Moreau–Yosida regularization of inequality constraints relaxes the complementarity constraints, but the associated regularized optimization problems typically still contain a (moderately) non-smooth term, which is also observed here, in form of the max-function involved in (27). Even though it complicates the (numerical) solution of (27), this is a desirable feature, as the max-function promotes the sparsity of the optimal transport plan, cf. the numerical experiments in [8, Section 4.3]. A similar observation is also made in the context of optimal control of VIs such as the obstacle problem, where the max-operator that arises from the Moreau–Yosida regularization of the complementarity system can be further smoothed (see e.g. [25]) or tackled by a semi-smooth Newton method (cf. [26]).

Concerning the convergence of solutions to (27) for regularization parameters tending to zero, Theorem 3.6 and Remark 3.7 imply the following result:

Corollary 4.4 *Let $D \subset \Omega$ satisfy the assumptions from Lemma 4.1 and let $\pi^{\mathrm{d}} \in \mathfrak{M}(D)$ be given. Moreover, assume that $\mu_2^{\mathrm{d}} \in L^1(\Omega_*)$ and that the transportation costs c_{d} fulfill condition 3 of Theorem 3.6. Then, for every sequence $\{(\gamma_n, \delta_n)\}_{n\in\mathbb{N}}$ tending to zero and fulfilling (12), there exists a subsequence of solutions to (27) denoted by $(\bar{\mu}_1^n, \bar{\alpha}_1^n, \bar{\alpha}_2^n, \bar{\varphi}_n)$ such that*

$$\bar{\mu}_1^n \rightharpoonup^* \bar{\mu}_1 \text{ in } \mathfrak{M}(\Omega_*),\ \ \frac{1}{\gamma}(\bar{\alpha}_1^n \oplus \bar{\alpha}_2^n - c_{\mathrm{d}})_+ \rightharpoonup^* \bar{\pi} \text{ in } \mathfrak{M}(\Omega),\ \ \bar{\varphi}_n \to \bar{\varphi} \text{ in } W_0^{1,p'}(D)$$

and the limit $(\bar{\mu}_1, \bar{\pi}, \bar{\varphi})$ is a solution to the bilevel Kantorovich problem with the objective from (25).

Though the transport plan has been erased from the regularized bilevel problem, the quadratic regularization does not completely resolve the "curse of dimensionality" associated with the Kantorovich problem, since it involves a PDE on D, which is a subset of Ω and thus $\dim(D) = d_1 + d_2$. By contrast, the next example provides a substantial reduction of the dimension for the price however that one looses convexity.

4.2 Optimal Control in Wasserstein Spaces

For our second example of a problem fulfilling the assumptions of Theorem 3.6, we suppose that $\Omega_* := \Omega_1 = \Omega_2$ satisfies the same assumptions as $\Omega = \Omega_* \times \Omega_*$, i.e., it shall coincide with the closure of its interior and shall have a Lipschitz boundary. As in case of Ω, we write $W_0^{1,q}(\Omega_*) := W_0^{1,q}(\mathrm{int}(\Omega_*))$ for the Sobolev space with vanishing trace. In order to have Lemma 4.1 at our disposal, we moreover assume that $d_* \le 3$. Furthermore, we fix $\nu > 0$, $\beta > 1$, and $q > d_*$ and set $q' = q/(q-1)$. We then consider the following *elliptic optimal control problem*:

$$\left.\begin{aligned} \min_{y,\mu_1}\ \ & \tfrac{1}{2}\,\|y - y_{\mathrm{d}}\|_{L^2(\Omega_*)}^2 + \nu\, W_\beta(\mu_1, \mu_2^{\mathrm{d}})^\beta \\ \text{s.t.}\ \ & y \in W_0^{1,q'}(\Omega_*),\quad \mu_1 \in \mathfrak{P}(\Omega_*), \\ & \mathrm{dist}(\mathrm{supp}(\mu_1), \partial\Omega_*) \ge \rho,\quad -\triangle y = \mu_1 \ \text{ in } W^{-1,q'}(\Omega_*), \end{aligned}\right\} \tag{OCP}$$

where $W_\beta(\mu_1, \mu_2^{\mathrm{d}})$ denotes the Wasserstein-β-distance between μ_1 and μ_2^{d} given by

$$W_\beta(\mu_1, \mu_2^{\mathrm{d}}) := \min\left\{\int_\Omega |x_1 - x_2|^\beta \,\mathrm{d}\varphi(x_1, x_2) \colon \varphi \in \Pi(\mu_1, \mu_2^{\mathrm{d}}),\ \varphi \ge 0\right\}^{\frac{1}{\beta}}. \tag{28}$$

Moreover, $y_{\mathrm{d}} \in L^2(\Omega_*)$ is a given desired state. Note that, again, due to $q > d_*$, there holds $\mathfrak{M}(\operatorname{int}(\Omega_*)) \hookrightarrow W^{-1,q'}(\Omega_*)$ with compact embedding such that the right hand side in the Poisson equation in (OCP) is well defined. Thanks to the weak-$*$ continuity of the restriction from $\mathfrak{M}(\Omega_*)$ to $\mathfrak{M}(\operatorname{int}(\Omega_*))$, the embedding $\mathfrak{M}(\Omega_*)$ in $W^{-1,q'}(\Omega_*)$ compact, too.

Remark 4.5 Depending on the application background, it might be favorable to measure the distance of the control μ_1 to a given prior μ_2^{d} in the Wasserstein distance instead of taking, e.g., the total variation norm $|\mu_1 - \mu_2^{\mathrm{d}}|(\Omega_*)$. Optimal control problems in measure spaces with the total variation as control costs have intensively been studied in literature, we only refer to [27–29] and the references therein. However, the total variation might be a too strong norm for several applications, see, e.g., [30–32] for beneficial properties of the Wasserstein distance.

Due to $d_* \leq 3$ and our regularity assumptions on Ω_*, Lemma 4.1 is applicable and guarantees the existence of a unique solution y of the Poisson equation in (OCP) for every $\mu_1 \in \mathfrak{M}(\Omega_*)$. The associated solution operator is again denoted by $G\colon \mathfrak{M}(\Omega_*) \to W_0^{1,q'}(\Omega_*)$. Given this solution operator, we can erase the state y from (OCP), which, together with (28), leads to the following reformulation:

$$
(\text{OCP}) \iff \begin{cases} \min\limits_{\pi,\mu_1} & \mathcal{J}(\pi,\mu_1) := \frac{1}{2}\,\|G\mu_1 - y_{\mathrm{d}}\|^2_{L^2(\Omega_*)} + \nu \displaystyle\int_\Omega |x_1 - x_2|^\beta \,\mathrm{d}\pi(x_1,x_2) \\ \text{s.t.} & \mu_1 \in \mathfrak{P}(\Omega_*), \quad \operatorname{dist}(\operatorname{supp}(\mu_1), \partial\Omega_*) \geq \rho, \\ & \pi \in \operatorname{arg\,min}\left\{ \displaystyle\int_\Omega |x_1 - x_2|^\beta \,\mathrm{d}\varphi(x_1,x_2) \colon \varphi \in \Pi(\mu_1, \mu_2^{\mathrm{d}}),\ \varphi \geq 0 \right\} \end{cases}
$$

$$
\iff \begin{cases} \min\limits_{\pi} & \frac{1}{2}\,\|G(P_{1\#}\pi) - y_{\mathrm{d}}\|^2_{L^2(\Omega_*)} + \nu \displaystyle\int_\Omega |x_1 - x_2|^\beta \,\mathrm{d}\pi(x_1,x_2) \\ \text{s.t.} & \pi \in \mathfrak{M}(\Omega), \quad \operatorname{dist}(\operatorname{supp}(P_{1\#}\pi), \partial\Omega_*) \geq \rho, \\ & \pi \in \Pi(P_{1\#}\pi, \mu_2^{\mathrm{d}}), \quad \pi \geq 0. \end{cases}
$$

Again, as in case of the example in Sect. 4.1, there is a natural extension of the objective to the enlarged domain satisfying Assumption 2.4, which reads

$$
\mathcal{J}^\rho(\pi,\mu_1) = \frac{1}{2}\,\|G\mu_1 - y_{\mathrm{d}}\|^2_{L^2(\Omega_*)} + \nu \int_{\Omega^\rho} |x_1 - x_2|^\beta \,\mathrm{d}\pi(x_1,x_2).
$$

Note that the solution operator G is clearly also compact from Ω_*^ρ to $W_0^{1,q'}(\Omega_*)$ so that $\mathcal{J}^\rho$ is indeed weakly-$*$ lower semicontinuous (actually it is even weakly-$*$ continuous) and it clearly coincides with $\mathcal{J}$, if $\operatorname{supp}(\pi) \subset \Omega$ and $\operatorname{supp}(\mu_1) \subset \Omega_*$.

While the first reformulation of (OCP) is a bilevel problem of the form (BK_ρ), the second one is (astonishingly) a convex problem, which is of course a favorable feature. However, as the transport plan is the optimization variable, we have to deal with a problem in $\Omega = \Omega_* \times \Omega_*$. Using Lemma 4.2, the quadratic regularization allows us to avoid this "curse of dimensionality". Abbreviating the transportation costs associated with the Wasserstein-β-distance by c_d, i.e., $c_d(x_1, x_2) = |x_1 - x_2|^\beta$, the regularized counterpart to (OCP) reads:

$$\left.\begin{aligned}
\min \quad & \tfrac{1}{2}\,\|y - y_d\|^2_{L^2(\Omega_*)} + \frac{\nu}{\gamma}\int_\Omega c_d\,(\alpha_1 \oplus \alpha_2 - c_d)_+\,d\lambda \\
\text{s.t.} \quad & y \in W_0^{1,q'}(\Omega_*), \quad \alpha_1, \alpha_2 \in L^2(\Omega_*), \quad \mu_1 \in \mathfrak{P}(\Omega_*), \\
& \operatorname{dist}(\operatorname{supp}(\mu_1), \partial\Omega_*) \geq \rho, \quad -\triangle y = \mu_1 \ \text{ in } W^{-1,q'}(\Omega_*), \\
& \int_{\Omega_*} (\alpha_1 \oplus \alpha_2 - c_d)_+(x_1, x_2)\,d\lambda_*(x_2) = \gamma\,\mathcal{T}_1^\delta(\mu_1)(x_1) \ \text{ a.e. in } \Omega_*, \\
& \int_{\Omega_*} (\alpha_1 \oplus \alpha_2 - c_d)_+(x_1, x_2)\,d\lambda_*(x_1) = \gamma\,\mathcal{T}_2^\delta(\mu_2^d)(x_2) \ \text{ a.e. in } \Omega_*.
\end{aligned}\right\} \qquad (\mathrm{OCP}_\gamma^\delta)$$

We observe that this problem no longer contains any variable or constraint in Ω, but only quantities and constraints in Ω_*. However, the price we have to pay for this reduction of the dimension is a loss of convexity, since ($\mathrm{OCP}_\gamma^\delta$) is no longer a convex problem due to the equality constraints including the max-function.

Remark 4.6 We point out that alternative regularization procedures like the entropic regularization lead to similar non-convex equality constraints, see [5, Theorem 4.8]. Alternatively, one might replace the Kantorovich problem in the bilevel formulation of (OCP) by its dual problem without any further regularization. At first glance, this seems to be promising, since the dual Kantorovich problem is posed in $C(\Omega_*) \times C(\Omega_*)$ instead of $\mathfrak{M}(\Omega_* \times \Omega_*)$ indicating the desired reduction of the dimension, see [2, Theorem 5.10] for the derivation of the dual Kantorovich problem. However, the bilevel problem then becomes a min-max-problem including a constraint in $\Omega_* \times \Omega_*$. To summarize, it seems that a reduction of the dimension without increasing the complexity of the problem is impossible.

Since the objective $\mathcal{J}$ in the bilevel formulation of (OCP) is linear in π, Theorem 3.6 is applicable, which yields the following

Corollary 4.7 *Suppose that $\Omega_* \in \mathbb{R}^{d_*}$, $d_* \leq 3$, is such that $\overline{\operatorname{int}(\Omega_*)} = \Omega_*$ and $\operatorname{int}(\Omega_*)$ is a bounded Lipschitz domain. Let $\nu > 0$ and $\beta > 1$ be given and let $q > d_*$ be the exponent from Lemma 4.1. Moreover, assume that $\mu_2^d \in L^2(\Omega_*)$. Then, for a given a sequence $\{(\gamma_n, \delta_n)\}_{n\in\mathbb{N}}$ tending to zero and fulfilling (12), there exists a subsequence of solutions ($\mathrm{OCP}_\gamma^\delta$) denoted by $(\bar{\mu}_1^n, \bar{\alpha}_1^n, \bar{\alpha}_2^n, \bar{y}_n)$ such that*

$$\bar{\mu}_1^n \rightharpoonup^* \bar{\mu}_1 \ \ \text{in } \mathfrak{M}(\Omega_*), \quad \tfrac{1}{\gamma}(\bar{\alpha}_1^n \oplus \bar{\alpha}_2^n - c_d)_+ \rightharpoonup^* \bar{\pi} \ \ \text{in } \mathfrak{M}(\Omega), \quad \bar{y}_n \to \bar{y} \ \ \text{in } W_0^{1,q}(\Omega_*)$$

and the limit $(\bar{\mu}_1, \bar{y})$ is a solution of (OCP).

Proof We verify the assumptions of Theorem 3.6 for the reduced form of (OCP). First, $c_d(x_1, x_2) = h(x_1 - x_2) := |x_1 - x_2|^\beta$ is clearly symmetric and strictly convex (since $\beta > 1$ by assumption) such that condition 3 is met. Moreover, the upper-level objective is even weak-$*$ continuous due to the linearity of $\pi \mapsto \int_\Omega c_{\mathrm{d}} \mathrm{d}\pi$ and the compactness of G. Moreover, it is clearly bounded on bounded sets and therefore, $\mathcal{J}$ fulfills our standing assumptions as well as condition 4. Thus, since $\mu_2^{\mathrm{d}} \ll \lambda_*$ by assumption, all hypotheses of Theorem 3.6 are fulfilled and the assertion follows by Remark 3.7. □

Remark 4.8 Due to the compactness properties of the solution operator of Poisson's equation, see Lemma 4.1, and the weak-$*$ convergence of $\mathcal{T}_1^n(\bar{\mu}_1^n)$ to $\bar{\mu}_1$ by Lemma 2.7, the convergence result from Corollary 4.7 will also hold, if one replaces μ_1 by $\mathcal{T}_1^\delta(\mu_1)$ in the state equation in (OCP$_\gamma^\delta$), which results in $-\triangle y = \mathcal{T}_1^\delta(\mu_1)$. This has the advantage that μ_1 does no longer directly appear in (OCP$_\gamma^\delta$) but only $\mathcal{T}_1^\delta(\mu_1)$, which is however uniquely determined by α_1 and α_2 through the equality constraints in (OCP$_\gamma^\delta$). One can therefore drop μ_1 as an optimization variable to obtain a problem in α_1, α_2, and y only.

Similarly, one can replace the objective in (27) by

$$\|\nabla \varphi\|_{L^{p'}(D;\mathbb{R}^d)}^{p'} + \nu \, \|\mathcal{T}_1^\delta(\mu_1) - \mu_1^{\mathrm{d}}\|_{\mathfrak{M}(\Omega_*)}$$

without loosing the convergence result of Corollary 4.4. To see this, consider a sequence of solutions associated with the regularized problems with this new objective, again denoted by $(\bar{\mu}_1^n, \bar{\alpha}_1^n, \bar{\alpha}_2^n, \bar{\varphi}_n)$, where $\bar{\mu}_1^n$ converges weakly-$*$ to $\bar{\mu}_1$. Then Lemma 2.7 along with the weak-$*$ lower semicontinuity of the Radon-norm implies that,

$$\|\bar{\mu}_1 - \mu_1^{\mathrm{d}}\|_{\mathfrak{M}(\Omega_*)} \leq \liminf_{n\to\infty} \|\mathcal{T}_1^n(\bar{\mu}_1^n) - \mu_1^{\mathrm{d}}\|_{\mathfrak{M}(\Omega_*)},$$

which gives the liminf-inequality from the proof of Corollary 3.4 for the new objective. To obtain the limsup-inequality for the recovery sequence from the proof of Theorem 3.6, we again choose $\mu_{1,n}^* := \mu_1^*$ as recovery sequence, cf. (20), and observe that the strict convergence of the convolution, see e.g. [33, Theorem 4.36], implies

$$\|\mathcal{T}_1^n(\mu_1^*) - \mu_1^{\mathrm{d}}\|_{\mathfrak{M}(\Omega_*)} \to \|\mu_1^* - \mu_1^{\mathrm{d}}\|_{\mathfrak{M}(\Omega_*)},$$

which yields the desired limsup-inequality. In summary, this implies that μ_1 can be replaced by $\mathcal{T}_1^\delta(\mu_1)$ in the optimization problem in (27), too. Again, in view of the equality constraints in (27), $\mathcal{T}_1^\delta(\mu_1)$ is uniquely determined by α_1 and α_2 so that μ_1 can again be dropped as optimization variable.

In particular, due to the non-smooth max-operator in its constraints, (OCP$_\gamma^\delta$) itself is a challenging problem. However, for instance by employing a further smoothing of max, it might be possible to approximate (OCP$_\gamma^\delta$) by smooth infinite-dimensional optimization problems, similarly to optimal control problems governed by VIs, see Remark 4.3. In light of Corollary 4.7, this might open a way for a numerical solution of (OCP) without "falling victim to the curse of dimensionality". An efficient solution of (OCP$_\gamma^\delta$) would however go beyond the scope of this paper and is subject to future research.

Acknowledgements The authors are very grateful to the anonymous referees for their valuable comments that substantially improved the paper.

Author Contributions All authors contributed to the paper. The first draft of the manuscript was written by Christian Meyer and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. This research was supported by the German Research Foundation (DFG) under grant number LO 1436/9-1 within the priority program Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization (SPP 1962).

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

1. Villani, C.: Topics in Optimal Transportation. American Mathematical Society, Providence, RI (2003)
2. Villani, C.: Optimal Transport: Old and New. Springer, Heidelberg (2009)
3. Ambrosio, L., Bressan, A., Helbing, D., Klar, A., Zuazua, E., Ambrosio, L., Gigli, N.: A user's guide to optimal transport. In: B. Piccoli, M. Rascle (Eds.), Modelling and Optimisation of Flows on Networks: Cetraro, Italy 2009, pp. 1–155 (2013)
4. Santambrogio, F.: Optimal Transport for Applied Mathematicians. Birkhäuser/Springer, Cham (2015). https://doi.org/10.1007/978-3-319-20828-2
5. Clason, C., Lorenz, D.A., Mahler, H., Wirth, B.: Entropic regularization of continuous optimal transport problems. J. Math. Anal. Appl. **494**(1), 124432–22 (2021). https://doi.org/10.1016/j.jmaa.2020.124432
6. Cuturi, M.: Sinkhorn distances: lightspeed computation of optimal transport. In: Burges, C., Bottou, L., Welling, M., Ghahramani, Z., Weinberger, K. (eds.) Advances in Neural Information Processing Systems, pp. 2292–2300. Springer, New York (2013)
7. Cuturi, M., Peyré, G.: Semidual regularized optimal transport. SIAM Rev. **60**(4), 941–965 (2018)
8. Lorenz, D.A., Manns, P., Meyer, C.: Quadratically regularized optimal transport. Appl. Math. Optim. **83**(3), 1919–1949 (2021). https://doi.org/10.1007/s00245-019-09614-w
9. Lorenz, D., Mahler, H.: Orlicz space regularization of continuous optimal transport problems. arXiv:2004.11574 (2021)
10. Hillbrecht, S., Meyer, C.: Bilevel Optimization of the Kantorovich Problem and its Quadratic Regularization, Part I: Existence Results. arXiv:2206.13220 (2022)
11. Hillbrecht, S., Meyer, C.: Bilevel Optimization of the Kantorovich Problem and Its Quadratic Regularization. Reverse Approximation in Finite Dimensions. In preparation, Part III (2024)
12. Hintermüller, M., Surowiec, T.: A bundle-free implicit programming approach for a class of elliptic MPECs in function space. Math. Program. **160**(1–2), 271–305 (2016)
13. Hertlein, L., Ulbrich, M.: An inexact bundle algorithm for nonconvex nonsmooth minimization in Hilbert space. SIAM J. Control Optim. **57**(5), 3137–3165 (2019). https://doi.org/10.1137/18M1221849
14. Christof, C., De los Reyes, J.C., Meyer, C.: A nonsmooth trust-region method for locally Lipschitz functions with application to optimization problems constrained by variational inequalities. SIAM J. Optim. **30**(3), 2163–2196 (2020). https://doi.org/10.1137/18M1164925

15. Grisvard, P.: Elliptic Problems in Nonsmooth Domains. Classics in Applied Mathematics, SIAM, Philadelphia (1985)
16. Adams, R.A., Fournier, J.J.F.: Sobolev Spaces. Pure and Applied Mathematics, vol. 140, 2nd edn. Elsevier/Academic Press, Amsterdam (2003)
17. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order. Classics in Mathematics, vol. 224. Springer, Berlin (1983)
18. Attouch, H., Buttazzo, G., Michaille, G.: Variational Analysis in Sobolev and BV Spaces. MPS/SIAM Series on Optimization, vol. 6, p. 634. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2006). Applications to PDEs and optimization
19. Meyer, C., Walther, S.: Optimal control of perfect plasticity Part II: displacement tracking. SIAM J. Control Optim. **59**(4), 2498–2523 (2021). https://doi.org/10.1137/20M1327331
20. Bogachev, V., Popova, S.: Optimal transportation of measures with a parameter. arXiv:2111.13014 (2021)
21. Simader, C.G.: On Dirichlet's Boundary Value Problem. Lecture Notes in Mathematics, vol. 268. Springer, New York (1972)
22. Gröger, K.: A $W^{1,p}$-estimate for solutions to mixed boundary value problems for second order elliptic differential equations. Math. Ann. **283**(4), 679–687 (1989). https://doi.org/10.1007/bf01442860
23. Jerison, D., Kenig, C.: The inhomogeneous Dirichlet problem in Lipschitz domains. J. Funct. Anal. **130**(1), 161–219 (1995)
24. Hoheisel, T., Kanzow, C., Schwartz, A.: Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints. Math. Program. **137**(1–2, Ser. A), 257–288 (2013). https://doi.org/10.1007/s10107-011-0488-5
25. Schiela, A., Wachsmuth, D.: Convergence analysis of smoothing methods for optimal control of stationary variational inequalities with control constraints. ESAIM: M2AN **47**, 771–787 (2013). https://doi.org/10.1051/m2an/2012049
26. Christof, C., Meyer, C., Walther, S., Clason, C.: Optimal control of a non-smooth semilinear elliptic equation. Math. Control Relat. Fields **8**(1), 247–276 (2018)
27. Casas, E., Clason, C., Kunisch, K.: Approximation of elliptic control problems in measure spaces with sparse solutions. SIAM J. Control Optim. **50**(4), 1735–1752 (2012)
28. Casas, E., Kunisch, K.: Optimal control of semilinear elliptic equations in measure spaces. SIAM J. Control Optim. **52**(1), 339–364 (2014)
29. Casas, E., Kunisch, K.: Optimal control of the two-dimensional stationary Navier-Stokes equations with measure valued controls. SIAM J. Control Optim. **57**(2), 1328–1354 (2019)
30. Bogachev, V.I.: Weak Convergence of Measures. American Mathematical Society, Providence, RI (2018)
31. Panaretos, V.M., Zemel, Y.: Statistical aspects of Wasserstein distances. Annu. Rev. Stat. Appl. **6**, 405–431 (2019)
32. Tartavel, G., Peyré, G., Gousseau, Y.: Wasserstein loss for image synthesis and restoration. SIAM J. Imaging Sci. **9**(4), 1726–1755 (2016)
33. Maggi, F.: Sets of Finite Perimeter and Geometric Variational Problems: An Introduction to Geometric Measure Theory. Cambridge Studies in Advanced Mathematics. Cambridge University Press. https://doi.org/10.1017/CBO9781139108133