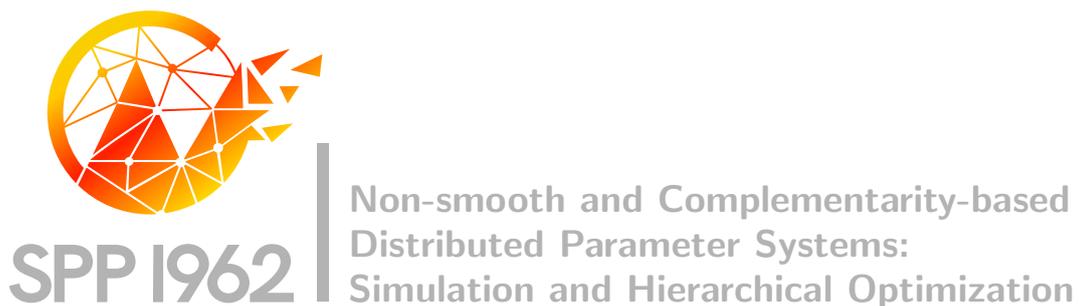


DFG Deutsche
Forschungsgemeinschaft
Priority Programme 1962

*Bilevel Optimization of the Kantorovich Problem
and its Quadratic Regularization Part II:
Convergence Analysis*

Sebastian Hillbrecht, Paul Manns, Christian Meyer



Preprint Number SPP1962-200

received on November 14, 2022

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

BILEVEL OPTIMIZATION OF THE KANTOROVICH PROBLEM AND ITS QUADRATIC REGULARIZATION PART II: CONVERGENCE ANALYSIS

SEBASTIAN HILLBRECHT, PAUL MANN, AND CHRISTIAN MEYER

ABSTRACT. This paper is concerned with an optimization problem that is constrained by the Kantorovich optimal transportation problem. This bilevel optimization problem can be reformulated as a mathematical problem with complementarity constraints in the space of regular Borel measures. Because of the non-smoothness induced by the complementarity relations, problems of this type are frequently regularized. Here we apply a quadratic regularization of the Kantorovich problem. 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.

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 $\mu_2 \in \mathfrak{M}(\Omega_2)$ on compact domains Ω_1 and Ω_2 with the same total mass and a continuous cost function $c \in C(\Omega_1 \times \Omega_2)$, the Kantorovich problem reads as follows:

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

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

$$(1.1) \quad \Pi(\mu_1, \mu_2) := \{\pi \in \mathfrak{M}(\Omega_1 \times \Omega_2) : P_{i\#}\pi = \mu_i, i = 1, 2\},$$

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 [30, 31, 2, 26] for more details on its mathematical background.

For the definition of the bilevel problem, the cost function is set to $c = c_d \in C(\Omega_1 \times \Omega_2)$ and we fix the second marginal $\mu_2 = \mu_2^d$, i.e., both c_d and μ_2^d are given

Date: November 14, 2022.

2010 Mathematics Subject Classification. 49Q20, 90C08, 49J45.

Key words and phrases. Optimal transport, Kantorovich problem, bilevel optimization, quadratic regularization.

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).

data. The upper-level optimization variables are π and μ_1 such that the overall bilevel problem is given by

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

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 $\mathcal{J}(\pi, \mu_1) = \mathcal{J}(\pi) := \|\pi - \pi^d\|_{\mathfrak{M}(D)}$, where $D \subset \Omega_1 \times \Omega_2$ is a given observation domain and $\pi^d \in \mathfrak{M}(D)$ denotes the measured transport plan in D . Another example in form of an optimal control problem that fits into the framework of (BK) will be given at the end of the paper.

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. [11] for a convergence analysis in function space), which leads to the well-known Sinkhorn algorithm, cf. [12, 13]. An alternative regularization approach was proposed in [23], 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 [23], 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 [22]. 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 [23] 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 [23] 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) and its regularized counterpart admit (globally optimal) solutions?
- Do solutions of the regularized bilevel problems (or subsequences thereof) converge to solutions of (BK) 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 [17], 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) 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 [18]. 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 [23, 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 [19, 16, 9].

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 [17] in Section 2. The remaining part of the paper is then devoted to the convergence analysis for vanishing regularization parameter $\gamma \searrow 0$. First, in Section 3.1, we show that weak-* accumulation points of solutions of the regularized bilevel problems are feasible for (BK), 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 Section 3.2, we establish the optimality of the weak-* limit under additional assumptions. The paper ends with two application-driven examples, 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 [14, 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. Let $(X, \mathfrak{B}(X))$ be a measurable space. We denote by $\mathfrak{M}(X)$ the space of (signed) regular Borel measures on X equipped with the total variation norm, i.e., $\|\mu\|_{\mathfrak{M}(X)} := |\mu|(X)$. 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 on 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

$$(1.2) \quad \Pi(\mu_1, \mu_2) := \{\pi \in \mathfrak{M}(\Omega): P_{1\#}\pi = \mu_1 \text{ and } P_{2\#}\pi = \mu_2\},$$

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

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

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 $\mu_1(\Omega_1) \neq \mu_2(\Omega_2)$. Throughout the paper, $\mu_2^d \in \mathfrak{M}(\Omega_2)$ is a fixed marginal satisfying $\mu_2^d \geq 0$ and, in order to ease notation, $|\mu_2^d|(\Omega_2) = 1$. 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}, μ) , 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 $c_d \in W^{1,p}(\Omega)$, $p > d$, where, with a slight abuse of notation, $W^{1,p}(\Omega)$ denotes the Sobolev space on $\text{int}(\Omega)$. Note that, due to the regularity of $\partial\Omega$, $W^{1,p}(\Omega)$ is compactly embedded in $C(\Omega)$, cf. e.g. [1, Theorem 6.3]. Thus, there exists a continuous representative of c_d , which we denote by the same symbol.

Bilevel Objective. The functional $\mathcal{J}: \mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1) \rightarrow \mathbb{R}$ 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

$$(1.3) \quad \sup\{|\mathcal{J}(\pi, \mu_1)| : \|(\pi, \mu_1)\|_{\mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1)} \leq M\} < \infty.$$

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 [17], we consider a slightly different problem involving an additional constraint on the distance of $\text{supp}(\mu_1)$ to the boundary $\partial\Omega_1$. This constraint is needed to ensure the weak-* convergence of the mollified marginal, but can be avoided by passing on to an equivalent problem on an enlarged domain, see Lemma 2.4 below. Our bilevel problem including the additional constraint with distance parameter $\rho > 0$ reads as follows:

$$(BK) \quad \left\{ \begin{array}{l} \inf_{\pi, \mu_1} \mathcal{J}(\pi, \mu_1) \\ \text{s.t. } \mu_1 \in \mathfrak{P}(\Omega_1), \quad \text{dist}(\text{supp}(\mu_1), \partial\Omega_1) \geq \rho, \\ \pi \in \arg \min \left\{ \int_{\Omega} c_d \, d\varphi : \varphi \in \Pi(\mu_1, \mu_2^d), \varphi \geq 0 \right\}. \end{array} \right.$$

In the following, we tacitly assume that $\rho > 0$ such that the feasible set of (BK) 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*

$$(2.1) \quad \mathcal{M} := \{\mu_1 \in \mathfrak{M}(\Omega_1) : \mu_1 \geq 0, \text{dist}(\text{supp}(\mu_1), \partial\Omega_1) \geq \rho\}$$

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 $x \in \text{supp}(\mu_1)$ with

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

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 $x \in \text{supp}(\mu_1)$, this yields the desired contradiction:

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

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

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

Proof. In [17, Theorem 3.2], the assertion is shown by means of the stability of transport plans according to [31, Theorem 5.20] for an analogous bilevel problem without the additional constraint $\text{dist}(\text{supp}(\mu_1), \partial\Omega_1) \geq \rho$. Since the set \mathcal{M} from (2.1) is weakly-* closed as seen in Lemma 2.1, the proof of [17, Theorem 3.2] readily carries over to (BK). \square

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).

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

As indicated above, due to the following elementary result, Assumption 2.3 as well as the additional constraint in (BK) can be avoided.

Lemma 2.4. *Consider a bilevel Kantorovich problem of the form*

$$(2.2) \quad \left\{ \begin{array}{l} \inf_{\pi, \mu_1} \mathcal{J}(\pi, \mu_1) \\ \text{s.t.} \quad \mu_1 \in \mathfrak{P}(\Omega_1), \quad \pi \in \arg \min \left\{ \int_{\Omega} c_d d\varphi : \varphi \in \Pi(\mu_1, \mu_2^d), \varphi \geq 0 \right\}, \end{array} \right.$$

where $\mu_2^d \in \mathfrak{P}(\Omega_2)$ need not necessarily satisfy Assumption 2.3. If we define $\Omega_i^\rho := \Omega_i + \overline{B(0, \rho)}$, $i = 1, 2$, and $\Omega^\rho := \Omega_1^\rho \times \Omega_2^\rho$, then (2.2) is equivalent to

$$(2.3) \quad \left\{ \begin{array}{l} \inf_{\pi, \mu_1} \mathcal{J}(\pi|_{\Omega}, \mu_1|_{\Omega_1}) \\ \text{s.t.} \quad \mu_1 \in \mathfrak{P}(\Omega_1^\rho), \quad \text{dist}(\text{supp}(\mu_1), \partial\Omega_1^\rho) \geq \rho, \\ \pi \in \arg \min \left\{ \int_{\Omega} c_d d\varphi : \varphi \in \Pi_\rho(\mu_1, \mu_2^\rho), \varphi \geq 0 \right\}, \end{array} \right.$$

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

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

The above equivalence means that, if μ_1 and π solve (2.2), their extensions by zero, denoted by μ_1^ρ and π^ρ , also solve (2.3), whereas, if μ_1 and π solve (2.3), their

restrictions $\mu_1|_{\Omega_1}$ and $\pi|_{\Omega}$ are solutions of (2.2), each time with the same optimal value.

Proof. If μ_1 and π are feasible for (2.2), then μ_1^ρ and π^ρ satisfy $\text{dist}(\text{supp}(\mu_1^\rho), \partial\Omega_1^\rho) \geq \rho$ and $\pi^\rho \in \Pi_\rho(\mu_1^\rho, \mu_2^\rho)$ by construction. Vice versa, if μ_1 and π are feasible for (2.3), then $\mu_1|_{\Omega_1}$ and $\pi|_{\Omega}$ satisfy $\pi|_{\Omega} \in \Pi(\mu_1|_{\Omega_1}, \mu_2^d)$, since $\pi \in \Pi_\rho(\mu_1, \mu_2^\rho)$ implies $\text{supp}(\pi) \subset \text{supp}(\mu_1) \times \text{supp}(\mu_2^\rho) = \text{supp}(\mu_1|_{\Omega_1}) \times \text{supp}(\mu_2^d)$. Therefore, because the objectives of the lower-level problems in (2.2) and (2.3) are the same, the same holds for the feasible sets of (2.2) and (2.3) (after extension by zero and restriction, respectively). Since the upper-level objectives are also identical, this gives the assertion. \square

Lemma 2.4 shows that one can equivalently solve (2.3) instead of (2.2) and the former problem satisfies an assumption analogous to Assumption 2.3 and guarantees the additional constraint on the distance of $\text{supp}(\mu_1)$ to the boundary. Without loss of generality, we can therefore assume Assumption 2.3 to hold and consider the modified problem (BK) involving the additional constraint on $\text{supp}(\mu_1)$. Consequently, we will tacitly take Assumption 2.3 for granted in the rest of the paper.

2.1. Quadratic Regularization. We now turn to the quadratic regularization of (BK). 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):

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

Lemma 2.5 ([23, Lemma 2.1]). *Problem (KP_γ) 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_γ) :

$$(2.4) \quad \mathcal{S}_\gamma : L^2(\Omega) \times \mathcal{M}_0(\Omega) \ni (c, \mu_1, \mu_2) \mapsto \pi_\gamma \in L^2(\Omega),$$

$$(2.5) \quad \mathcal{M}_0(\Omega) := \left\{ (\mu_1, \mu_2) \in L^2(\Omega_1) \times L^2(\Omega_2) : \|\mu_1\|_{L^1(\Omega_1)} = \|\mu_2\|_{L^1(\Omega_2)}, \right. \\ \left. \mu_i \geq 0 \text{ } \lambda_i\text{-a.e. in } \Omega_i, i = 1, 2 \right\}.$$

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_γ) admits a solution, too, see [23, Theorem 2.11] and Lemma 3.8 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_γ) 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

$$(2.6) \quad \mathcal{T}_i^\delta : \mathfrak{M}(\Omega_i) \ni \mu_i \mapsto \varphi_i^\delta * \mu_i + \frac{\delta}{|\Omega_i|} \in L^2(\Omega_i), \quad i = 1, 2.$$

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. [3, 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).

Lemma 2.6. *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*

$$(2.7) \quad \text{dist}(\text{supp}(\mu_n), \partial\Lambda) \geq \rho > 0 \quad \forall n \in \mathbb{N}.$$

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

Proof. Let $v \in C(\Lambda)$ be arbitrary and denote $\Lambda_\rho := \{x \in \Lambda : \text{dist}(x, \partial\Lambda) \geq \rho\}$. Then Fubini's theorem along with (2.7) yields

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

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

We note that the restrictions on Ω , non-empty interior and Lipschitz boundary, are not required for Lemma 2.6 because $A_\rho \subset \text{int}(\Lambda)$ is ensured by $\text{dist}(\text{supp} \mu_n, \partial\Lambda) \geq \rho$.

According to Lemma 2.5, (KP_γ) only admits a solution if the total mass of the marginals is the same. In context of the bilevel problem $(\text{BK}_\gamma^\delta)$ below, this is ensured for the smoothed marginals by Assumption 2.3 and the additional constraint on $\text{supp}(\mu_1)$, provided that $\delta \leq \rho$. Since these imply $\text{supp}(\varphi_2^\delta * \mu_2^d) \subset \Omega_2$ and $\text{supp}(\varphi_1^\delta * \mu_1) \subset \Omega_1$, respectively, we obtain for every $\mu_1 \in \mathfrak{P}(\Omega_1)$ with $\text{dist}(\text{supp}(\mu_1), \partial\Omega_1) \geq \rho$ that

$$(2.8) \quad \begin{aligned} \|\mathcal{T}_1^\delta(\mu_1)\|_{L^1(\Omega_1)} &= \int_{\Omega_1} \varphi_1^\delta * \mu_1 d\lambda_1 + \delta \\ &= \|\mu_1\|_{\mathfrak{M}(\Omega_1)} + \delta = \|\mu_2^d\|_{\mathfrak{M}(\Omega_2)} + \delta = \|\mathcal{T}_2^\delta(\mu_2^d)\|_{L^1(\Omega_2)} \end{aligned}$$

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

$$(\text{BK}_\gamma^\delta) \quad \begin{cases} \inf_{\pi_\gamma, \mu_1, c} & \mathcal{J}_\gamma(\pi_\gamma, \mu_1, c) := \mathcal{J}(\pi_\gamma, \mu_1) + \frac{1}{p\gamma} \|c - c_d\|_{W^{1,p}(\Omega)}^p \\ \text{s.t.} & c \in W^{1,p}(\Omega), \quad \mu_1 \in \mathfrak{P}(\Omega_1), \quad \text{dist}(\text{supp}(\mu_1), \partial\Omega_1) \geq \rho, \\ & \pi_\gamma = \mathcal{S}_\gamma(c, \mathcal{T}_1^\delta(\mu_1), \mathcal{T}_2^\delta(\mu_2^d)). \end{cases}$$

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), 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 [24]. For this

reason, c is treated as an additional optimization variable to have more flexibility at this point. In the companion paper [18], 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}_γ will ensure that, in the limit, c equals the given costs c_d , see (3.3) below.

Proposition 2.7. *Let $\gamma > 0$ and $\delta \in (0, \rho]$ be fixed. There exists at least one globally optimal solution to the regularized bilevel Kantorovich problem $(\text{BK}_\gamma^\delta)$.*

Proof. The existence of solutions for a slightly different problem has been shown in [17, Theorem 4.7], which differs from $(\text{BK}_\gamma^\delta)$ as follows: First, the bilevel problem in [17] does not contain the additional constraint on $\text{supp}(\mu_1)$, 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 $\mathcal{T}_2^\delta(\mu_2^d)$ have the same total mass. In our case, however, this is guaranteed by Assumption 2.3 and the constraint on $\text{supp}(\mu_1)$ together with $\delta \leq \rho$, see (2.8). 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 [17, Theorem 4.7]. \square

Remark 2.8. 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 ([17, Theorem 4.7]). If one can drop the term $\frac{1}{p\gamma} \|c - c_d\|_{W^{1,p}(\Omega)}^p$, 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 $(\text{BK}_\gamma^\delta)$ 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

$$(3.1) \quad 0 < \delta_n \leq \rho \quad \text{for all } n \in \mathbb{N} \quad \text{and} \quad \frac{\gamma_n}{\delta_n^d} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

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 $(\text{BK}_{\gamma_n}^{\delta_n})$ 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 2.7 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

$$(3.2) \quad \begin{aligned} \|\bar{\pi}_n\|_{\mathfrak{M}(\Omega)} &= \int_{\Omega_1} \int_{\Omega_2} \bar{\pi}_n \, d\lambda_2 \, d\lambda_1 \\ &= \int_{\Omega_1} \mathcal{T}_1^n(\bar{\mu}_1^n) \, d\lambda_1 = \|\varphi_1^n\|_{L^1(\mathbb{R}^{d_1})} \|\bar{\mu}_1^n\|_{\mathfrak{M}(\Omega_1)} + \delta_n \leq 1 + \rho \end{aligned}$$

for all $n \in \mathbb{N}$, where we make use of $\bar{\pi}_n \geq 0$, (2.8), and (3.1). Hence, there is a subsequence (denoted by the same symbol to ease notation) such that

$$(\bar{\pi}_n, \bar{\mu}_1^n) \rightharpoonup^* (\bar{\pi}, \bar{\mu}_1) \quad \text{in } \mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1) \quad \text{as } n \rightarrow \infty.$$

Furthermore, take an arbitrary, but fixed $\mu_1 \in \mathfrak{P}(\Omega_1)$ and consider the regularized optimal transport plans $\pi_n = \mathcal{S}_n(c_d, \mathcal{T}_1^n(\mu_1), \mathcal{T}_2^n(\mu_2^d))$ for $n \in \mathbb{N}$. Then, the triple (π_n, μ_1, c_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. (3.2). The optimality of $(\bar{\pi}_n, \bar{\mu}_1^n, \bar{c}_n)$ for (BK_n) thus yields

$$\|\bar{c}_n - c_d\|_{W^{1,p}(\Omega)}^p \leq p \gamma_n (\mathcal{J}(\pi_n, \mu_1) - \mathcal{J}(\bar{\pi}_n, \bar{\mu}_1^n)) \leq \gamma_n C$$

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

$$(3.3) \quad \bar{c}_n \rightarrow c_d \quad \text{in } W^{1,p}(\Omega) \text{ as } n \rightarrow \infty.$$

Now that we have found an accumulation point $(\bar{\pi}, \bar{\mu}_1, c_d)$ 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) . We start with the feasibility of $(\bar{\pi}, \bar{\mu}_1)$ for (BK) in the next subsection.

3.1. Feasibility of the Limit Plan.

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., $\pi \in \Pi(\mu_1, \mu_2^d)$.*

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_γ) imply

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

where we use Lemma 2.6 and the additional constraint on $\text{supp}(\mu_1)$ in (KP_γ) 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 $P_{2\#}\pi = \mu_2^d$. \square

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 $\pi \in \Pi(\mu_1, \mu_2^d)$ be a non-negative coupling between μ_1 and μ_2^d . Then, there exists a sequence of non-negative couplings $\pi_n \in \Pi(\mu_1^n, \mu_2^d)$ 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 [31, 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., [26, 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,

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

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 $B_1 \in \mathfrak{B}(\Omega_1)$

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

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

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

so that $\pi_n \in \Pi(\mu_1^n, \mu_2^d)$ 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 [4, Theorem 3.1]. For this purpose, define the mapping

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

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

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

as well as

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

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 :

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

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 [31, Theorem 6.9]. Using this equivalence once more finally yields $\pi_n \rightharpoonup^* \pi$ in $\mathfrak{M}(\Omega)$ as $n \rightarrow \infty$. \square

Proposition 3.3. *Assume that the vanishing sequences $\{\gamma_n\}_{n \in \mathbb{N}}$ and $\{\delta_n\}_{n \in \mathbb{N}}$ satisfy (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 $(\pi, \mu_1, c_d) \in \mathfrak{M}(\Omega) \times \mathfrak{M}(\Omega_1) \times W^{1,p}(\Omega)$ 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) , i.e., $\mu_1 \in \mathfrak{P}(\Omega_1)$, $\text{dist}(\text{supp}(\mu_1), \partial\Omega_1) \geq \rho$, 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 (2.1) 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 :

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

By [31, 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 limit cost function c_d . We consider an arbitrary of these solutions and denote it by $\pi^* \in \mathfrak{M}(\Omega)$. 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

$$\varphi_n(x_1, x_2) := \varphi_1^n(x_1) \varphi_2^n(x_2), \quad (x_1, x_2) \in \Omega,$$

$$\text{and } \vartheta_n^* := \varphi_n * \pi_n^* + \frac{\delta_n}{|\Omega_1| |\Omega_2|} = \int_{\Omega} \varphi_n(\xi - \cdot) \, d\pi_n^*(\xi) + \frac{\delta_n}{|\Omega_1| |\Omega_2|} \in L^2(\Omega).$$

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

$$\int_{\Omega_2} \vartheta_n^* \, d\lambda_2 = \varphi_1^n * \mu_1^n + \frac{\delta_n}{|\Omega_1|} = \mathcal{T}_1^n(\mu_1^n), \quad \int_{\Omega_1} \vartheta_n^* \, d\lambda_1 = \varphi_2^n * \mu_2^d + \frac{\delta_n}{|\Omega_2|} = \mathcal{T}_2^n(\mu_2^n)$$

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

$$(3.4) \quad \begin{aligned} \langle c_d, \pi \rangle &= \lim_{n \rightarrow \infty} \langle c_n, \pi_n \rangle \leq \liminf_{n \rightarrow \infty} \int_{\Omega} c_n \pi_n \, d\lambda + \frac{\gamma_n}{2} \|\pi_n\|_{L^2(\Omega)}^2 \\ &\leq \limsup_{n \rightarrow \infty} \int_{\Omega} c_n \vartheta_n^* \, d\lambda + \frac{\gamma_n}{2} \|\vartheta_n^*\|_{L^2(\Omega)}^2. \end{aligned}$$

Let us investigate the two addends on the right hand side of this inequality separately. Since $c_n \rightarrow c_d$ 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 $\Omega_\rho := \{x \in \Omega :$

$\text{dist}(x, \partial\Omega) \geq \rho\}$. Then, the uniform convergence of both c_n and the convolution in compact subsets of $\text{int}(\Omega)$ yields

$$\max_{x \in \Omega_\rho} \left| \int_{\Omega} c_n(\xi) \varphi(x - \xi) d\lambda(\xi) - c_d(x) \right| = 0.$$

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

$$(3.5) \quad \int_{\Omega} c_n \vartheta_n^* d\lambda = \int_{\Omega_\rho} \int_{\Omega} c_n(x) \varphi(x - \xi) d\lambda(x) d\pi_n^*(\xi) + \frac{\delta_n}{|\Omega_1| |\Omega_2|} \int_{\Omega} c_n(x) d\lambda(x) \rightarrow \langle c_d, \pi^* \rangle.$$

For the second addend on the right hand side of (3.4), we obtain

$$\|\vartheta_n^*\|_{L^2(\Omega)}^2 \leq 2 \int_{\Omega} \int_{\Omega} \varphi(x - \xi) d\pi_n^*(\xi)^2 d\lambda(x) + 2\delta_n^2 \leq 2\|\varphi_n\|_{L^2(B(0, \delta_n))}^2 \|\pi_n^*\|_{\mathfrak{M}(\Omega)}^2 + 2\delta_n^2,$$

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

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

with a constant $C > 0$. In view of the coupling of γ_n and δ_n in (3.1) and $\|\pi_n^*\|_{\mathfrak{M}(\Omega)} = 1$ for all $n \in \mathbb{N}$, we thus arrive at

$$\frac{\gamma_n}{2} \|\vartheta_n^*\|_{L^2(\Omega)}^2 \leq C \left(\frac{\gamma_n}{\delta_n^d} + \gamma_n \delta_n^2 \right) \rightarrow 0.$$

Inserting this together with (3.5) in (3.4) implies $\langle c_d, \pi \rangle \leq \langle c_d, \pi^* \rangle$ 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) as claimed. \square

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). 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 (3.1). Suppose moreover that (π^*, μ_1^*) is a solution to the bilevel problem (BK) 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) $\limsup_{n \rightarrow \infty} \mathcal{J}_n(\pi_n^*, \mu_{1,n}^*, c_n^*) \leq \mathcal{J}(\pi^*, \mu_1^*)$.

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).*

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 .

$$\begin{aligned} \mathcal{J}(\bar{\pi}, \bar{\mu}_1) &\leq \liminf_{n \rightarrow \infty} \mathcal{J}(\bar{\pi}_n, \bar{\mu}_1^n) \\ &\leq \liminf_{n \rightarrow \infty} \mathcal{J}(\bar{\pi}_n, \bar{\mu}_1^n) + \frac{1}{p\gamma_n} \|\bar{c}_n - c_d\|_{W^{1,p}(\Omega)}^p \\ &= \liminf_{n \rightarrow \infty} \mathcal{J}_n(\bar{\pi}_n, \bar{\mu}_1^n, \bar{c}_n) \\ &\leq \limsup_{n \rightarrow \infty} \mathcal{J}_n(\pi_n^*, \mu_{1,n}^*, c_n^*) \leq \mathcal{J}(\pi^*, \mu_1^*). \end{aligned}$$

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

The crucial task is now of course to establish the existence of a recovery sequence satisfying (i) and (ii). So far, unfortunately, we are not able to guarantee the existence of such a sequence in the general setting without further assumptions. If, however, $\mu_2^d \ll \lambda^{d_2}$, $c(x_1, x_2) = h(x_1 - x_2)$ 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.

Theorem 3.5. *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) $\mu_2^d \ll \lambda_*$ with $\lambda_* := \lambda_1 = \lambda_2$,
- (3) $c_d(x_1, x_2) = h(x_1 - x_2)$ with a function $h : \mathbb{R}^{d_*} \rightarrow \mathbb{R}$ that is strictly convex and even symmetric, i.e., $h(-\xi) = h(\xi)$ for all $\xi \in \mathbb{R}^{d_*}$,
- (4) \mathcal{J} is 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

$$\limsup_{n \rightarrow \infty} \mathcal{J}(\pi_n, \mu_1) \leq \mathcal{J}(\pi, \mu_1),$$

- (5) The null sequences $\{\gamma_n\}_{n \in \mathbb{N}}$ and $\{\delta_n\}_{n \in \mathbb{N}}$ satisfy (3.1).

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).

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). According to the generalized version of Brenier's theorem in [31, 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 (μ_2^d, μ_1^*) and cost $c_d(x_1, x_2)$, i.e.,

$$\min \left\{ \int_{\Omega} c_d(x_2, x_1) d\varphi(x_1, x_2) : \varphi \in \Pi(\mu_2^d, \mu_1^*), \varphi \geq 0 \right\}.$$

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

$$(KP^*) \quad \min \left\{ \int_{\Omega} c_d(x_1, x_2) d\varphi(x_1, x_2) : \varphi \in \Pi(\mu_1^*, \mu_2^d), \varphi \geq 0 \right\}$$

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

Let us define the sequence

$$(3.6) \quad \mu_{1,n}^* := \mu_1^*, \quad c_n^* := c_d, \quad \pi_n^* := \mathcal{S}_n(c_d, \mathcal{T}_1^n(\mu_1^*), \mathcal{T}_2^n(\mu_2^d)), \quad n \in \mathbb{N}.$$

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) , 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 (π^*, μ_1^*, c_d) . The presupposed weak-* continuity of \mathcal{J} finally yields that

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

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. \square

Remark 3.6. An inspection of the proof of Theorem 3.5 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 (3.6). Accordingly, the assertion of Theorem 3.5 remains true, if one considers

$$(BK_{\gamma, c_d}^{\delta}) \quad \begin{cases} \inf_{\pi_{\gamma}, \mu_1} & \mathcal{J}(\pi_{\gamma}, \mu_1) \\ \text{s.t.} & \mu_1 \in \mathfrak{P}(\Omega_1), \quad \text{dist}(\text{supp}(\mu_1), \partial\Omega_1) \geq \rho, \\ & \pi_{\gamma} = \mathcal{S}_{\gamma}(c_d, \mathcal{T}_1^{\delta}(\mu_1), \mathcal{T}_2^{\delta}(\mu_2^d)). \end{cases}$$

instead of (BK_{γ}^{δ}) . In the finite-dimensional setting, however, it is exactly the additional optimization variable c , which enables the construction of a recovery sequence, see [18].

Consequently, one may drop the additional optimization over c and the term $\frac{1}{\gamma^p} \|c - c_d\|_{W^{1,p}(\Omega)}^p$ in the setting of Theorem 3.5. 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 2.8.

The additional assumptions in Theorem 3.5 are certainly rather restrictive, in particular condition (4). Let us therefore end our considerations by giving two examples fulfilling these assumptions.

3.2.1. 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 Section 1. The most natural choice for the upper-level objective probably reads

$$(3.7) \quad \mathcal{J}(\pi, \mu_1) := \|\pi - \pi^d\|_{\mathfrak{M}(D)} + \nu \|\mu_1 - \mu_1^d\|_{\mathfrak{M}(\Omega_*)},$$

where $\pi^d \in \mathfrak{M}(D)$ denotes the measurement of the transport plan, while $\mu_1^d \in \mathfrak{P}(\Omega_*)$ 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 3.5. 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 $W_0^{1,p}(D) \hookrightarrow C(\overline{D})$ is compact and so, by Schauder's theorem, $\mathfrak{M}(\overline{D})$ embeds compactly in $W^{-1,p'}(D) := W_0^{1,p}(D)^*$, where, as usual, $p' = p/(p-1)$ denotes the conjugate exponent. Therefore, for a given $\pi^d \in \mathfrak{M}(\overline{D})$, an objective of the form

$$(3.8) \quad \tilde{\mathcal{J}}(\pi, \mu_1) := \|\pi - \pi^d\|_{W^{-1,p'}(D)}^{p'} + \nu \|\mu_1 - \mu_1^d\|_{\mathfrak{M}(\Omega_*)}$$

fulfills condition (4). 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

$$(3.9) \quad -\operatorname{div}(|\nabla\psi|^{p-2}\nabla\psi) = \pi - \pi^d \quad \text{in } W^{-1,p'}(D).$$

Then, it is easily seen that

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

and hence, the objective from (3.8) becomes $\|\nabla\psi\|_{L^p(D;\mathbb{R}^d)}^p + \nu \|\mu_1 - \mu_1^d\|$. 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 $\eta \in \mathfrak{M}(\overline{D})$ be given and consider

$$(3.10) \quad \varphi \in W_0^{1,p'}(D), \quad -\Delta\varphi = \eta \quad \text{in } W^{-1,p'}(D),$$

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

Lemma 3.7. *Let D be a bounded domain of class C^1 . Then there exists an exponent $p > d$ such that, for every $\eta \in \mathfrak{M}(\overline{D})$, there exists a unique solution $\varphi \in W_0^{1,p'}(D)$ of (3.10). The associated solution operator denoted by $G : \mathfrak{M}(\overline{D}) \rightarrow W_0^{1,p'}(D)$ is linear and compact.*

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

Proof. According to [28, Theorem 4.6], there exists a $p > d$ such that the Poisson equation 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 [15] for the case $d = 2$ and [21] for $d = 3$. By duality, there is thus a unique solution in $\varphi \in W_0^{1,p'}(D)$ to the state equation in (3.10) for every right hand side in $W^{-1,p'}(D)$. Due to the continuous embedding $W_0^{1,p}(D) \hookrightarrow C(\overline{D})$ already mentioned above, there holds $\mathfrak{M}(\overline{D}) \hookrightarrow W^{-1,p'}(D)$ and we obtain the existence and uniqueness of $\varphi \in W_0^{1,p'}(D)$ for every $\pi \in \mathfrak{M}(\overline{D})$. 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 $W^{1,p}(D) \hookrightarrow C(\overline{D})$. \square

¹On C^1 -domains, this even holds for every $p < \infty$, see [28, Theorem 4.6].

Unfortunately, the $W^{-1,p'}(D)$ -norm cannot be expressed by means of the solution operator G , but there holds $\|\eta\|_{W^{-1,p'}(D)} \leq \|\nabla\varphi\|_{L^{p'}(\Omega;\mathbb{R}^d)} \leq \|G\| \|\eta\|_{W^{-1,p'}(D)}$ and therefore,

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

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

$$(3.11) \quad \widehat{\mathcal{J}}(\pi, \mu_1) := \|\nabla G(\pi - \pi^d)\|_{L^{p'}(D;\mathbb{R}^d)}^{p'} + \nu \|\mu_1 - \mu_1^d\|_{\mathfrak{M}(\Omega_*)}$$

satisfies the condition (4) in Theorem 3.5. Thus, if we consider the bilevel Kantorovich problem (BK) with $\widetilde{\mathcal{J}}$ or $\widehat{\mathcal{J}}$, Theorem 3.5 applies. Let us shortly turn to the associated regularized problems. For this purpose, recall the following result from [23]:

Lemma 3.8 ([23, 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 exist functions $\alpha_1 \in L^2(\Omega_1)$ and $\alpha_2 \in L^2(\Omega_2)$ satisfying*

$$(3.12a) \quad \pi_\gamma - \frac{1}{\gamma}(\alpha_1 \oplus \alpha_2 - c)_+ = 0 \quad \lambda\text{-a.e. in } \Omega,$$

$$(3.12b) \quad \int_{\Omega_2} \pi_\gamma(x_1, x_2) d\lambda_2(x_2) = \mu_1(x_1) \quad \lambda_1\text{-a.e. in } \Omega_1,$$

$$(3.12c) \quad \int_{\Omega_1} \pi_\gamma(x_1, x_2) d\lambda_1(x_1) = \mu_2(x_2) \quad \lambda_2\text{-a.e. in } \Omega_2.$$

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 $(\text{BK}_{\gamma, c_d}^\delta)$ by using the necessary and sufficient conditions in (3.12). Let us exemplarily consider $(\text{BK}_{\gamma, c_d}^\delta)$ with the objective from (3.11). The corresponding regularized bilevel problem then reads as follows:

$$(3.13) \quad \left\{ \begin{array}{l} \min \quad \|\nabla\varphi\|_{L^{p'}(D;\mathbb{R}^d)}^{p'} + \nu \|\mu_1 - \mu_1^d\|_{\mathfrak{M}(\Omega_*)} \\ \text{s.t.} \quad \alpha_1, \alpha_2 \in L^2(\Omega_*), \quad \mu_1 \in \mathfrak{P}(\Omega_*), \quad \varphi \in W_0^{1,p'}(D), \\ \quad \text{dist}(\text{supp}(\mu_1), \partial\Omega_*) \geq \rho, \\ \quad -\Delta\varphi = \frac{1}{\gamma}(\alpha_1 \oplus \alpha_2 - c_d)_+ - \pi^d \quad \text{in } W^{-1,p'}(D), \\ \quad \int_{\Omega_*} (\alpha_1 \oplus \alpha_2 - c_d)_+(x_1, x_2) d\lambda_*(x_2) = \gamma \mathcal{T}_1^\delta(\mu_1)(x_1) \quad \text{a.e. in } \Omega_*, \\ \quad \int_{\Omega_*} (\alpha_1 \oplus \alpha_2 - c_d)_+(x_1, x_2) d\lambda_*(x_1) = \gamma \mathcal{T}_2^\delta(\mu_1^d)(x_2) \quad \text{a.e. in } \Omega_*. \end{array} \right.$$

Remark 3.9. If one replaces the lower-level Kantorovich problem as in (BK) 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 [17, 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 [20] and the references therein. In light of the above reformulation based on Lemma 3.8, 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 [23, 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 (3.13). Even though it complicates the (numerical) solution of (3.13), this is a desirable feature, as the max-function promotes the sparsity of the optimal transport plan, cf. the numerical experiments in [23, Section 4.3]. A similar observation is also made in 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. [27]) or tackled by a semi-smooth Newton method (cf. [10]).

Concerning the convergence of solutions to (3.13) for regularization parameters tending to zero, Theorem 3.5 and Remark 3.6 imply the following result:

Corollary 3.10. *Let $D \subset \Omega$ satisfy the assumptions from Lemma 3.7 and let $\pi^d \in \mathfrak{M}(\overline{D})$ be given. Moreover, assume that $\mu_2^d \in L^1(\Omega_*)$ and that the transportation costs c_d fulfill condition (3) of Theorem 3.5. Then, for every sequence $\{(\gamma_n, \delta_n)\}_{n \in \mathbb{N}}$ tending to zero and fulfilling (3.1), there exists a subsequence of solutions to (3.13) denoted by $(\bar{\mu}_1^n, \bar{\alpha}_1^n, \bar{\alpha}_2^n, \bar{\varphi}_n)$ such that*

$\bar{\mu}_1^n \rightharpoonup^* \bar{\mu}_1$ in $\mathfrak{M}(\Omega_*)$, $\frac{1}{\gamma}(\bar{\alpha}_1^n \oplus \bar{\alpha}_2^n - c_d)_+ \rightharpoonup^* \bar{\pi}$ in $\mathfrak{M}(\Omega)$, $\bar{\varphi}_n \rightarrow \bar{\varphi}$ 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 (3.11).

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 loses convexity.

3.2.2. Optimal Control in Wasserstein Spaces. For our second example for a problem fulfilling the assumptions of Theorem 3.5, we suppose that $\Omega_* := \Omega_1 = \Omega_2$ satisfies the same assumptions as $\Omega = \Omega_* \times \Omega_*$, i.e., it coincides with the closure of its interior and has a Lipschitz boundary. As in case of Ω , we write $W_0^{1,q}(\Omega_*) := W_0^{1,q}(\text{int}(\Omega_*))$ for the Sobolev space with vanishing trace. In order to have Lemma 3.7 at our disposal, we moreover assume that $d_* \leq 3$. 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*:

$$(OCP) \quad \begin{cases} \min_{y, \mu_1} & \frac{1}{2} \|y - y_d\|_{L^2(\Omega_*)}^2 + \nu W_\beta(\mu_1, \mu_2^d)^\beta \\ \text{s.t.} & y \in W_0^{1,q'}(\Omega_*), \quad \mu_1 \in \mathfrak{P}(\Omega_*), \\ & \text{dist}(\text{supp}(\mu_1), \partial\Omega_*) \geq \rho, \quad -\Delta y = \mu_1 \text{ in } W^{-1,q'}(\Omega_*), \end{cases}$$

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

$$(3.14) \quad W_\beta(\mu_1, \mu_2^d) := \min \left\{ \int_\Omega |x_1 - x_2|^\beta d\varphi(x_1, x_2) : \varphi \in \Pi(\mu_1, \mu_2^d), \varphi \geq 0 \right\}^{\frac{1}{\beta}}.$$

Moreover, $y_d \in L^2(\Omega_*)$ is a given desired state. Note that, again, due to $q > d_*$, there holds $\mathfrak{M}(\Omega_*) \hookrightarrow W^{-1, q'}(\Omega_*)$ with compact embedding such that the right hand side in Poisson equation in (OCP) is well defined.

Remark 3.11. 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 $|\mu_1 - \mu_2^d|(\Omega_*)$. Optimal control problems in measure spaces with the total variation as control costs have intensively been studied in literature, we only refer to [6, 7, 8] and the references therein. However, the total variation might be a too strong norm for several applications, see, e.g., [5, 25, 29] for beneficial properties of the Wasserstein distance.

Due to $d_* \leq 3$ and our regularity assumptions on Ω_* , Lemma 3.7 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 $G : \mathfrak{M}(\Omega_*) \rightarrow W_0^{1, q'}(\Omega_*)$. Given this solution operator, we can erase the state y from (OCP), which, together with (3.14), leads to the following reformulation:

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

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

While the first reformulation is a bilevel problem of the form (BK), 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 3.8, the quadratic regularization allows 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:

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

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 $(\text{OCP}_\gamma^\delta)$ is no longer a convex problem due to the equality constraints including the max-function.

Remark 3.12. We point out that alternative regularization procedures like the entropic regularization lead to similar non-convex equality constraints, see [11, 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 [31, 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 3.5 is applicable, which yields the following

Corollary 3.13. *Suppose that $\Omega_* \in \mathbb{R}^{d_*}$, $d_* \leq 3$, is such that $\overline{\text{int}(\Omega_*)} = \Omega_*$ and $\text{int}(\Omega_*)$ is a bounded Lipschitz domain. Let $\nu > 0$ and $\beta > 1$ be given and let $q > d_*$ be the exponent from Lemma 3.7. 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 (3.1), there exists a subsequence of solutions $(\text{OCP}_\gamma^\delta)$ denoted by $(\bar{\mu}_1^n, \bar{\alpha}_1^n, \bar{\alpha}_2^n, \bar{y}_n)$ such that*

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

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

Proof. We verify the assumptions of Theorem 3.5 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 $\pi \mapsto \int_\Omega c_d d\pi$ 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 $\mu_2^d \ll \lambda_*$ by assumption, all hypotheses of Theorem 3.5 are fulfilled and the assertion follows by Remark 3.6. \square

In particular, due to the non-smooth max-operator in its constraints, $(\text{OCP}_\gamma^\delta)$ itself is a challenging problem. However, for instance by employing a further smoothing of max, it might be possible to approximate $(\text{OCP}_\gamma^\delta)$ by smooth infinite-dimensional optimization problems, similarly to optimal control problems governed by VIs, see Remark 3.9. In light of Corollary 3.13, this might open a way for a numerical solution of (OCP) without “falling victim to the curse of dimensionality”. An efficient solution of $(\text{OCP}_\gamma^\delta)$ would however go beyond the scope of this paper and is subject to future research.

REFERENCES

- [1] Robert A. Adams and John J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003.
- [2] Luigi Ambrosio and Nicola Gigli. A user’s guide to optimal transport. In *Modelling and optimisation of flows on networks*, pages 1–155. Springer, 2013.

- [3] Hedy Attouch, Giuseppe Buttazzo, and Gérard Michaille. *Variational analysis in Sobolev and BV spaces*, volume 6 of *MPS/SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA, 2006. Applications to PDEs and optimization.
- [4] V. Bogachev and S. Popova. Optimal transportation of measures with a parameter. arXiv:2111.13014, 2021.
- [5] Vladimir Igorevich Bogachev. *Weak Convergence of Measures*. American Mathematical Society, Providence, 2018.
- [6] Eduardo Casas, Christian Clason, and Karl Kunisch. Approximation of elliptic control problems in measure spaces with sparse solutions. *SIAM J. Control Optim.*, 50(4):1735–1752, 2012.
- [7] Eduardo Casas and Karl Kunisch. Optimal control of semilinear elliptic equations in measure spaces. *SIAM J. Control Optim.*, 52(1):339–364, 2014.
- [8] Eduardo Casas and Karl Kunisch. Optimal control of the two-dimensional stationary Navier-Stokes equations with measure valued controls. *SIAM J. Control Optim.*, 57(2):1328–1354, 2019.
- [9] Constantin Christof, Juan Carlos De los Reyes, and Christian Meyer. 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.
- [10] Constantin Christof, Christian Meyer, Stephan Walther, and Christian Clason. Optimal control of a non-smooth semilinear elliptic equation. *Math. Control Relat. Fields*, 8(1):247–276, 2018.
- [11] Christian Clason, Dirk A. Lorenz, Hinrich Mahler, and Benedikt Wirth. Entropic regularization of continuous optimal transport problems. *J. Math. Anal. Appl.*, 494(1):Paper No. 124432, 22, 2021.
- [12] Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in neural information processing systems*, pages 2292–2300, 2013.
- [13] Marco Cuturi and Gabriel Peyré. Semidual regularized optimal transport. *SIAM Review*, 60(4):941–965, 2018.
- [14] Pierre Grisvard. *Elliptic Problems in Nonsmooth Domains*. Classics in Applied Mathematics. SIAM, Philadelphia, 1985.
- [15] Konrad Gröger. A $W^{1,p}$ -estimate for solutions to mixed boundary value problems for second order elliptic differential equations. *Mathematische Annalen*, 283(4):679–687, 1989.
- [16] Lukas Hertle and Michael Ulbrich. An inexact bundle algorithm for nonconvex nonsmooth minimization in Hilbert space. *SIAM J. Control Optim.*, 57(5):3137–3165, 2019.
- [17] Sebastian Hillbrecht and Christian Meyer. Bilevel optimization of the Kantorovich problem and its quadratic regularization, part I: Existence results. arXiv:2206.13220, 2022.
- [18] Sebastian Hillbrecht and Christian Meyer. Bilevel optimization of the Kantorovich problem and its quadratic regularization, part III: Reverse approximation in finite dimensions. In preparation, 2022.
- [19] Michael Hintermüller and Thomas Surowiec. A bundle-free implicit programming approach for a class of elliptic MPECs in function space. *Math. Program.*, 160(1–2):271–305, 2016.
- [20] Tim Hoheisel, Christian Kanzow, and Alexandra Schwartz. Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints. *Math. Program.*, 137(1–2, Ser. A):257–288, 2013.
- [21] D. Jerison and C. Kenig. The inhomogeneous Dirichlet problem in Lipschitz domains. *J. Funct. Anal.*, 130(1):161–219, 1995.
- [22] Dirk Lorenz and Hinrich Mahler. Orlicz space regularization of continuous optimal transport problems. arXiv:2004.11574, 2021.
- [23] Dirk A. Lorenz, Paul Manns, and Christian Meyer. Quadratically regularized optimal transport. *Appl. Math. Optim.*, 83(3):1919–1949, 2021.
- [24] Christian Meyer and Stephan Walther. Optimal control of perfect plasticity Part II: Displacement tracking. *SIAM J. Control Optim.*, 59(4):2498–2523, 2021.
- [25] Victor M Panaretos and Yoav Zemel. Statistical aspects of wasserstein distances. *Annual review of statistics and its application*, 6:405–431, 2019.
- [26] Filippo Santambrogio. *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.

- [27] A. Schiela and D. Wachsmuth. Convergence analysis of smoothing methods for optimal control of stationary variational inequalities with control constraints. *ESAIM: M2AN*, 47:771–787, 2013.
- [28] Christian G. Simader. *On Dirichlet's Boundary Value Problem*, volume 268 of *Lecture Notes in Mathematics*. Springer, New York, 1972.
- [29] Guillaume Tartavel, Gabriel Peyré, and Yann Gousseau. Wasserstein loss for image synthesis and restoration. *SIAM Journal on Imaging Sciences*, 9(4):1726–1755, 2016.
- [30] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
- [31] Cédric Villani. *Optimal transport. Old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009.

SEBASTIAN HILLBRECHT, TECHNISCHE UNIVERSITÄT DORTMUND, FAKULTÄT FÜR MATHEMATIK,
LEHRSTUHL X, VOGELPOTHSWEG 87, 44227 DORTMUND, GERMANY
Email address: `sebastian.hillbrecht@tu-dortmund.de`

PAUL MANNS, TECHNISCHE UNIVERSITÄT DORTMUND, FAKULTÄT FÜR MATHEMATIK, LEHRSTUHL
X, VOGELPOTHSWEG 87, 44227 DORTMUND, GERMANY
Email address: `paul.manns@tu-dortmund.de`

CHRISTIAN MEYER, TECHNISCHE UNIVERSITÄT DORTMUND, FAKULTÄT FÜR MATHEMATIK, LEHRSTUHL
X, VOGELPOTHSWEG 87, 44227 DORTMUND, GERMANY
Email address: `christian2.meyer@tu-dortmund.de`