# A MIXED FORMULATION OF THE MONGE-KANTOROVICH EQUATIONS 

John W. Barrett ${ }^{1}$ and Leonid Prigozhin ${ }^{2}$


#### Abstract

We introduce and analyse a mixed formulation of the Monge-Kantorovich equations, which express optimality conditions for the mass transportation problem with cost proportional to distance. Furthermore, we introduce and analyse the finite element approximation of this formulation using the lowest order Raviart-Thomas element. Finally, we present some numerical experiments, where both the optimal transport density and the associated Kantorovich potential are computed for a coupling problem and problems involving obstacles and regions of cheap transportation.


Mathematics Subject Classification. 35D05, 35J85, 49J40, 65N12, 65N30, 82B27.
Received November 15, 2006. Revised May 21, 2007.

## 1. Introduction

The Monge-Kantorovich (MK) problem, first considered by Monge in 1781 and reformulated by Kantorovich in 1942, has a great variety of applications and was generalized in many different ways. Recently, the problem has been investigated with renewed interest, see $[1,2,15,16,28]$ and the references therein. Let $f^{+}, f^{-} \in \mathcal{M}^{+}\left(\mathbb{R}^{n}\right)$ be two given measures having the same finite total mass, $\int_{\mathbb{R}^{n}} f^{+}=\int_{\mathbb{R}^{n}} f^{-}$. Here $\mathcal{M}\left(\mathbb{R}^{n}\right)$ is the Banach space of Radon measures, with $\mathcal{M}^{+}\left(\mathbb{R}^{n}\right)$ being the subset of nonnegative measures. In the relaxed variational formulation by Kantorovich, the MK problem consists in determining the optimal transport plan; i.e., the measure $\gamma \in \mathcal{M}^{+}\left(\mathbb{R}^{n} \times \mathbb{R}^{n}\right)$ having projections $f^{+}$and $f^{-}$, such that $\int_{U \times \mathbb{R}^{n}} \gamma=\int_{U} f^{+}$and $\int_{\mathbb{R}^{n} \times U} \gamma=\int_{U} f^{-}$ for all Borel sets $U \subset \mathbb{R}^{n}$, and minimizing the cost of transportation

$$
\begin{equation*}
\int_{\mathbb{R}^{n} \times \mathbb{R}^{n}} c(\underline{x}, \underline{y}) \gamma(\underline{x}, \underline{y}) \tag{1.1}
\end{equation*}
$$

of one distribution of mass into another. Here $c(\underline{x}, \underline{y})$ is the cost of transportation of one unit of mass from $\underline{x}$ to $\underline{y}$. Under some additional assumptions on $f^{ \pm}$and $c$, see $[2,16,28]$, an optimal plan $\gamma$ exists and can be chosen as a measure concentrated in $\mathbb{R}^{n} \times \mathbb{R}^{n}$ on the graph of a map $T: \mathbb{R}^{n} \rightarrow \mathbb{R}^{n}$, the optimal transport map. This is the map originally sought by Monge, who did not allow mass splitting and assumed that cost is equal to the distance, i.e. $c(\underline{x}, \underline{y})=|\underline{x}-\underline{y}|$. Although MK problems have since been studied theoretically for a wide class of cost functions, $c(\underline{x}, \underline{y})=|\underline{x}-\underline{y}|^{p}$ with $p=1$ and $p=2$ are the two most often considered.

[^0](C) EDP Sciences, SMAI 2007

Several numerical algorithms, see $[4,9]$ and the references therein, have been derived for the quadratic case ( $p=2$ ), where the cost function is smooth and strictly convex. In this case the unique optimal transport plan is a map; and moreover, this map is the gradient of a convex function. Numerical methods for the classical linear case $(p=1)$, which is the case considered in this paper, are much less developed. This cost-equal-to-distance MK problem differs from those with $p>1$ mainly because the cost function is not strictly convex. In this case not every optimal plan is a map. With minor restrictions on $f^{ \pm}$, an optimal map exists; but it is typically not unique. However, its non-uniqueness is a one-dimensional phenomenon that can be eliminated by an additional condition of map monotonicity along transport rays, see [ 15,16$]$.

Optimal plans are solutions of an infinite-dimensional linear programming problem and Kantorovich formulated the dual problem; for $c(\underline{x}, \underline{y})=|\underline{x}-\underline{y}|$, it reads

$$
\begin{equation*}
\max \left\{\int_{\mathbb{R}^{n}} u f: u \in \operatorname{Lip}_{1}\left(\mathbb{R}^{n}\right)\right\} \tag{1.2}
\end{equation*}
$$

where $f=f^{+}-f^{-}$and $\operatorname{Lip}_{1}\left(\mathbb{R}^{n}\right):=\left\{u: \mathbb{R}^{n} \rightarrow \mathbb{R}:|u(\underline{x})-u(\underline{y})| \leq|\underline{x}-\underline{y}|, \forall \underline{x}, \underline{y} \in \mathbb{R}^{n}\right\}$. Under some additional regularity conditions on $f^{ \pm}$, Evans and Gangbo [16] have showed that the maximum in (1.2) is attained by a function $u$ satisfying, jointly with a suitable function $a$, the following Monge-Kantorovich equations

$$
\begin{equation*}
-\underline{\nabla} \cdot(a \underline{\nabla} u)=f, \quad \text { with } \quad|\underline{\nabla} u| \leq 1, \quad a \geq 0, \quad \text { and } \quad|\underline{\nabla} u|<1 \Rightarrow a=0 . \tag{1.3}
\end{equation*}
$$

Here $u$ is the Kantorovich potential, $a$ is the transport density, and the relations (1.3) were named the MK equations by Bouchitté et al. [10], who proved that for every measure $f$ with finite total variation and zero average there exists a weak solution to the MK equations, $(u, a) \in \operatorname{Lip}_{1}\left(\mathbb{R}^{n}\right) \times \mathcal{M}^{+}\left(\mathbb{R}^{n}\right)$.

The transport density plays a very important role, as the total mass of the measure $a$ is equal to the total transportation cost, (1.1) with $c(\underline{x}, \underline{y})=|\underline{x}-\underline{y}|$, for some optimal plan $\gamma$; and, for any Borel set $U \subset \mathbb{R}^{n}, a(U)$ is the cost of transportation inside $\bar{U}$. It has been shown that the transport density, unlike the optimal plan, is unique if either $f^{+}$or $f^{-}$is absolutely continuous, see [2]. In addition, the transport density is supported on transport rays, which are straight line segments starting in $X:=\operatorname{supp}\left(f^{+}\right)$and ending in $Y:=\operatorname{supp}\left(f^{-}\right)$.

The direction of optimal transportation coincides with the direction of $-\underline{\nabla} u$. Since $|\underline{\nabla} u|=1$ a.e. in $\operatorname{supp}(a)$, the vector measure $\underline{q}=-a \underline{\nabla} u$, the transport flux, contains all information on the direction and density of optimal transportation. Clearly, this flux calculation is of interest in various applications.

As shown in [16], under suitable assumptions on $f^{ \pm}$, a solution to the MK equations is the $p \rightarrow \infty$ limit of solutions to the $p$-Laplacian equation $-\nabla \cdot\left(\left|\nabla u_{p}\right|^{p-2} \nabla u_{p}\right)=f$. More precisely, $u_{p} \rightarrow u$ uniformly locally and $-\left|\nabla u_{p}\right|^{p-2} \nabla u_{p} \rightarrow q$ weak- in $L^{\infty}$ (for a subsequence) as $p \rightarrow \infty$. In practical calculations, however, such an approximation allows one to approximate only the Kantorovich potential $u$. Accurate computation of the transport flux $\underline{q}$ is difficult, because of the numerical instabilities for large values of $p$.

One possible approach to computing the optimal flux is based on the similarity, noted by Evans [5, 15], between the MK problem with cost equal to distance and the evolutionary model of sandpile growth [5, 24]. Indeed, the latter model can be regarded as a non-stationary version of the MK equations if one assumes that sand is poured out of a distributed source with the constant rate $f^{+}$and is, simultaneously, reclaimed from the pile surface with the rate $f^{-}$. Recently, a dual variational formulation of the sandpile model in terms of transport flux has been proposed and used in numerical simulations; see [7]. Making use of exactly the same method, it is possible (see [25]) to compute the optimal transport flux in this MK problem as the $t \rightarrow \infty$ limit of the flux in the sandpile model; and the Kantorovich potential can be found as $t \rightarrow \infty$ limit of the evolving pile surface. It may be noted that, as an auxiliary variable, time has been introduced also into numerical schemes $[4,9]$ for the MK problem with quadratic cost.

Here, on once again setting $\underline{q}=-a \underline{\nabla} u$, we derive a different (mixed and "steady-state") approximation for the linear cost problem based on the following equivalent reformulation of the MK equations (1.3):

$$
\begin{equation*}
\underline{\nabla} \cdot \underline{q}=f, \quad-\underline{\nabla} u \cdot(\underline{v}-\underline{q}) \leq|\underline{v}|-|\underline{q}| \quad \forall \underline{v} . \tag{1.4}
\end{equation*}
$$

On integrating the inequality, and noting the equation, in (1.4), one obtains the following alternative dual formulation of the MK problem:

$$
\begin{equation*}
\min \left\{\int_{\mathbb{R}^{n}}|\underline{q}|: \underline{q} \in\left[\mathcal{M}\left(\mathbb{R}^{n}\right)\right]^{n} \quad \text { and } \quad \underline{\nabla} \cdot \underline{q}=f\right\} . \tag{1.5}
\end{equation*}
$$

The formulation (1.5) appeared in [10] and also, in a different context, in some previous works (see, e.g. [27]). However, as far as we know, it has not been exploited in the numerical approximation of the cost-equal-todistance MK problem.

In this work, we consider a slightly more general problem. Instead of working with the whole space, $\mathbb{R}^{n}$, we define the region of possible transportation as the closure of a connected bounded open set $\Omega \subset \mathbb{R}^{n}$ with a Lipschitz boundary $\partial \Omega$, by assuming $X, Y \subset \bar{\Omega}$ and setting the no-flux boundary condition $\left.\underline{q} \cdot \underline{\nu}\right|_{\partial \Omega}=0$. Furthermore, we assume the cost of transporting one unit of mass an infinitesimal distance $\mathrm{d} \underline{x}$ near the point $\underline{x} \in \bar{\Omega}$ is $k(\underline{x}) \mathrm{d} \underline{x}$, where $k: \bar{\Omega} \rightarrow \mathbb{R}_{>0}$ is a given continuous function. In this set up, the transportation cost is defined as

$$
\begin{equation*}
c_{k, \Omega}(\underline{x}, \underline{y})=\inf \left\{\int_{0}^{1} k(s(t))\left|s^{\prime}(t)\right| \mathrm{d} t: s \in C^{0,1}([0,1] ; \bar{\Omega}), s(0)=\underline{x}, s(1)=\underline{y}\right\}, \tag{1.6}
\end{equation*}
$$

and the possible transport rays are the continuous paths minimizing this cost. Clearly, if $k \equiv 1$ and $\bar{\Omega}$ contains all straight line segments $[\underline{x}, \underline{y}]$, with $\underline{x} \in X$ and $\underline{y} \in Y$, the problem becomes equivalent to the classical MK problem in $\mathbb{R}^{n}$ because all optimal transportation occurs along such segments in this case. Otherwise for the problem that we consider in this paper, one should replace the convex set $\operatorname{Lip}_{1}\left(\mathbb{R}^{n}\right)$ in the dual formulation (1.2) by

$$
\begin{equation*}
\operatorname{Lip}_{k}(\bar{\Omega}):=\left\{u: \bar{\Omega} \rightarrow \mathbb{R}:|u(\underline{x})-u(\underline{y})| \leq c_{k, \Omega}(\underline{x}, \underline{y}), \forall \underline{x}, \underline{y} \in \bar{\Omega}\right\} \tag{1.7}
\end{equation*}
$$

and to minimize the integral $\int_{\bar{\Omega}} k|\underline{q}|$ in the dual formulation (1.5). For the benefit of the reader, we present in the Appendix a proof, based on standard convex analysis theory and a density result, of this duality between the potential and flux formulations. The weak form of the mixed formulation (1.4), for these generalized cost-equal-to-distance MK equations, can be correspondingly written for given $f \in \mathcal{M}_{I}(\bar{\Omega}):=\left\{\eta \in \mathcal{M}(\bar{\Omega}): \int_{\bar{\Omega}} \eta=0\right\}$ and positive $k \in C(\bar{\Omega})$ as:
$(\widetilde{\mathbf{Q}})$ Find $\underline{q} \in \underline{\mathcal{V}}_{0}^{\mathcal{M}}(\Omega)$ and $u \in C(\bar{\Omega})$ such that

$$
\begin{array}{rlrl}
\langle\underline{\nabla} \cdot \underline{q}, \eta\rangle_{C(\bar{\Omega})} & =\langle f, \eta\rangle_{C(\bar{\Omega})} & & \forall \eta \in C(\bar{\Omega}), \\
\langle | \underline{v}|, k\rangle_{C(\bar{\Omega})}-\langle | \underline{q}|, k\rangle_{C(\bar{\Omega})} \geq\langle\underline{\nabla} \cdot(\underline{v}-\underline{q}), u\rangle_{C(\bar{\Omega})} & & \forall \underline{v} \in \underline{\mathcal{V}}_{0}^{\mathcal{M}}(\Omega) ; \tag{1.8b}
\end{array}
$$

where

$$
\begin{align*}
& \underline{\mathcal{V}}_{0}^{\mathcal{M}}(\Omega):=\left\{\underline{v} \in \underline{\mathcal{V}}^{\mathcal{M}}(\Omega):\langle\underline{\nabla} \cdot \underline{v}, \eta\rangle_{C(\bar{\Omega})}=-\langle\underline{v}, \underline{\nabla} \eta\rangle_{C(\bar{\Omega})} \quad \forall \eta \in C^{1}(\bar{\Omega})\right\},  \tag{1.9a}\\
& \underline{\mathcal{V}}^{\mathcal{M}}(\Omega):=\left\{\underline{v} \in[\mathcal{M}(\bar{\Omega})]^{n}: \underline{\nabla} \cdot \underline{v} \in \mathcal{M}(\bar{\Omega})\right\} \tag{1.9b}
\end{align*}
$$

and $\langle\cdot, \cdot\rangle_{C(\bar{\Omega})}$ is the duality pairing on $[C(\bar{\Omega})]^{*} \times C(\bar{\Omega})$ with $\mathcal{M}(\bar{\Omega}) \equiv[C(\bar{\Omega})]^{*}$ being the dual of $C(\bar{\Omega})$, which is naturally extended to the case of vectors. The constraint on $\underline{\mathcal{V}}^{\mathcal{M}}(\Omega)$, in defining $\underline{\mathcal{V}}_{0}^{\mathcal{M}}(\Omega)$ in (1.9a), is a weak formulation of the no-flux boundary condition $\underline{v} \cdot \underline{\nu}=0$ on $\partial \Omega$, where $\underline{\nu}$ is normal to $\partial \Omega$.

In order to consider inhomogeneous transportation domains, it is convenient to further generalize the problem $(\widetilde{\mathrm{Q}})$ above by allowing $k$ to be a piecewise continuous positive function. Obviously, this requires $\underline{\mathcal{V}}_{0}^{\mathcal{M}}(\Omega)$ to be modified so that the terms on the left-hand side of (1.8b) are well-defined. As this requires some technical
details, we postpone the exact formulation of this generalization of ( $\widetilde{Q}$ ), which we denote as problem (Q), until later in this section after we have introduced our notation.

Another extension of the optimal transportation problem is when only a specified fraction, $m \leq \min \left\{\int_{\bar{\Omega}} f^{+}\right.$, $\left.\int_{\bar{\Omega}} f^{-}\right\}$, is allowed to be transported and the problem may be unbalanced, $\int_{\bar{\Omega}} f \neq 0$. For theoretical work on this fractional MK problem with a quadratic cost function, $p=2$, and with $f^{ \pm} \in L^{1}(\Omega)$ see the recent preprint [11]. In [8], we extend our approach in this paper to fractional MK problems with cost-equal-to-distance, $p=1$.

The outline of this paper is as follows. In the next section we introduce a regularized version, $\left(\mathrm{Q}_{r}\right), r>1$, of (Q) by replacing the non-differentiable function $|\underline{v}|$ by $\frac{1}{r}|\underline{v}|^{r}$. First, we prove existence of a unique solution to $\left(\mathrm{Q}_{r}\right)$. Then, on passing to the limit $r \rightarrow 1$, we prove existence of a solution to (Q). In Section 3 we introduce a fully practical finite element approximation, $\left(\mathrm{Q}_{r}^{h}\right)$, of $\left(\mathrm{Q}_{r}\right)$ based on the lowest order Raviart-Thomas element with vertex sampling on the nonlinear term. We establish well-posedness of this approximation, and establish, for any fixed $r \in\left(1, \frac{4}{3}\right)$, convergence of $\left(\mathrm{Q}_{r}^{h}\right)$ to $\left(\mathrm{Q}_{r}\right)$ as $h \rightarrow 0$. In Section 4 we describe our numerical algorithm for solving the nonlinear algebraic system arising from our approximation $\left(\mathrm{Q}_{r}^{h}\right)$. In addition, we describe our adaptive mesh refinement strategy, which is used to increase the accuracy of our approximation to solutions with singularities. Finally in Section 5, we present various numerical experiments based on the discretization $\left(\mathrm{Q}_{r}^{h}\right)$, where we compute approximations to both the transport density and the Kantorovich potential for some typical problems.

We end this section with a few remarks about the notation employed in this paper, and our precise formulation of (Q). Above and throughout we adopt the standard notation for Sobolev spaces on a bounded Lipschitz domain $D$, denoting the norm of $W^{\ell, p}(D)(\ell \in \mathbb{N}, p \in[1, \infty])$ by $\|\cdot\|_{\ell, p, D}$ and the semi-norm by $|\cdot|_{\ell, p, D}$. Of course, we have that $|\cdot|_{0, p, D} \equiv\|\cdot\|_{0, p, D}$. We extend these norms and semi-norms in the natural way to the corresponding spaces of vector valued functions. For $p=2, W^{\ell, 2}(D)$ will be denoted by $H^{\ell}(D)$ with the associated norm and semi-norm written as, respectively, $\|\cdot\|_{\ell, D}$ and $|\cdot|_{\ell, D}$. We introduce $L_{I}^{p}(D):=\left\{\eta \in L^{p}(D): \int_{D} \eta \mathrm{~d} \underline{x}=0\right\}$, and recall the Poincaré inequality for any $p \in[1, \infty]$ :

$$
\begin{equation*}
|\eta|_{0, p, D} \leq C_{D}|\underline{\nabla} \eta|_{0, p, D} \quad \forall \eta \in W_{I}^{1, p}(D):=W^{1, p}(D) \cap L_{I}^{p}(D), \tag{1.10}
\end{equation*}
$$

where $C_{D}$ depends on $D$, but is independent of $p$. The measure of $D$ will be denoted by $|D|$.
For our regularized problem, $\left(\mathrm{Q}_{r}\right)$, in the next section we require, for $r \in(1, \infty)$, the reflexive Banach space

$$
\begin{align*}
\underline{V}_{0}^{r}(\Omega) & :=\left\{\underline{v} \in \underline{V}^{r}(\Omega): \underline{v} \cdot \underline{\nu}=0 \text { on } \partial \Omega\right\},  \tag{1.11a}\\
\text { where } \quad \underline{V}^{r}(\Omega) & :=\left\{\underline{v} \in\left[L^{r}(\Omega)\right]^{n}: \underline{\nabla} \cdot \underline{v} \in L^{r}(\Omega)\right\},  \tag{1.11b}\\
\text { with norm } \quad\|\underline{v}\|_{\underline{V}^{r}(\Omega)} & :=\left[|\underline{v}|_{0, r, \Omega}^{r}+|\underline{\nabla} \cdot \underline{v}|_{0, r, \Omega}^{r}\right]^{\frac{1}{r}} . \tag{1.11c}
\end{align*}
$$

We note that $\underline{V}_{0}^{r}(\Omega)$ is the strong closure of $\left[C_{0}^{\infty}(\Omega)\right]^{n}$ in the norm $\|\cdot\|_{V^{r}(\Omega)}$; e.g. this can be shown by a simple extension of the argument for the case $r=2$ in [21], pp. 26-29.

Let $C(\bar{D})$ denote the space of continuous functions on $\bar{D}$, and $C_{I}(\bar{D}):=\left\{\eta \in C(\bar{D}): \int_{D} \eta \mathrm{~d} \underline{x}=0\right\}$. As one can identify $L^{1}(D)$ as a closed subspace of $\mathcal{M}(\bar{D})$, it is convenient to adopt the notation

$$
\begin{equation*}
\int_{\bar{D}}|\mu| \equiv\|\mu\|_{\mathcal{M}(\bar{D})}:=\sup _{\eta \in C(\bar{D})} \frac{\langle\mu, \eta\rangle_{C(\bar{D})}}{|\eta|_{0, \infty, D}}<\infty . \tag{1.12}
\end{equation*}
$$

We note that if $\left\{\mu_{j}\right\}_{j \geq 0}$ is a bounded sequence in $\mathcal{M}(\bar{D})$, then there exist a subsequence $\left\{\mu_{j_{\ell}}\right\}_{j_{\ell} \geq 0}$ and a $\mu \in \mathcal{M}(\bar{D})$ such that as $j_{\ell} \rightarrow \infty$

$$
\begin{equation*}
\mu_{j_{\ell}} \rightarrow \mu \quad \text { vaguely in } \mathcal{M}(\bar{D}) ; \quad \text { i.e. } \quad\left\langle\mu_{j_{\ell}}-\mu, \eta\right\rangle_{C(\bar{D})} \rightarrow 0 \quad \forall \eta \in C(\bar{D}) . \tag{1.13}
\end{equation*}
$$

In addition, we have that

$$
\begin{equation*}
\liminf _{j_{\ell} \rightarrow \infty} \int_{\bar{D}}\left|\mu_{j_{\ell}}\right| \geq \int_{\bar{D}}|\mu| ; \tag{1.14}
\end{equation*}
$$

see e.g. [14], p. 5 and [20], p. 223.
We assume that our generalized cost function is determined by

$$
\begin{equation*}
k(\underline{x})=k^{(i)}(\underline{x}) \geq k_{\min }>0 \quad \forall \underline{x} \in \overline{\Omega^{(i)}}, \quad i=1 \rightarrow N, \tag{1.15}
\end{equation*}
$$

where $\Omega^{(i)}$ are disjoint open connected subsets of $\Omega$ with Lipschitz boundaries and $\bar{\Omega} \equiv \cup_{i=1}^{N} \overline{\Omega^{(i)}}$, and $k^{(i)} \in$ $C\left(\overline{\Omega^{(i)}}\right), i=1 \rightarrow N$. In order to extend $(\widetilde{\mathrm{Q}})$ to such piecewise continuous $k$, we introduce $\underline{\Phi}$ to be the Banach space of vector functions $\underline{\phi}: \cup_{i=1}^{N} \Omega^{(i)} \rightarrow \mathbb{R}^{n}$ such that for each $i$ the restriction $\left.\underline{\phi}\right|_{\Omega^{(i)}}$ is continuous and can be extended to a function from $\left[C\left(\overline{\Omega^{(i)}}\right)\right]^{n}$. The elements of the dual space $\underline{\mathcal{M}}:=[\underline{\Phi}]^{*}$ can be represented as $\underline{v}=$ $\left(\underline{v}^{(1)}, \ldots, \underline{v}^{(N)}\right)$, where $\underline{v}^{(i)} \in \underline{\mathcal{M}}\left(\overline{\Omega^{(i)}}\right):=\left[\left[C\left(\overline{\Omega^{(i)}}\right)\right]^{n}\right]^{*}$ is a vector Radon measure, $\|\underline{v}\|_{\underline{\mathcal{M}}}:=\sum_{i=1}^{N}\left\|\underline{v}^{(i)}\right\|_{\underline{\mathcal{M}}}\left(\overline{\Omega^{(i)}}\right)$, and $\langle\underline{v}, \underline{\phi}\rangle_{\underline{\Phi}}:=\sum_{i=1}^{N}\left\langle\underline{v}^{(i)}, \underline{\phi}{\left.\overline{\Omega^{(i)}}\right\rangle}_{\left[C\left(\overline{\Omega^{(i)}}\right)\right]^{n}}\right.$. To simplify our notation, we write

$$
\begin{equation*}
\langle | \underline{v}|, k\rangle:=\sum_{i=1}^{N}\langle | \underline{v}^{(i)}\left|, k^{(i)}\right\rangle_{C\left(\overline{\Omega^{(i)}}\right)} \quad \forall \underline{v} \in \underline{\mathcal{M}} . \tag{1.16}
\end{equation*}
$$

The weak form of the mixed formulation (1.4), that is the basis of our numerical approximation of these generalized cost-equal-to-distance MK equations for given $f \in \mathcal{M}_{I}(\bar{\Omega})$ and $k=\left(k^{(1)}, \ldots, k^{(N)}\right)$ satisfying (1.15), is:
(Q) Find $\underline{q} \in \underline{V}_{0}^{\mathcal{M}}(\Omega)$ and $u \in C(\bar{\Omega})$ such that

$$
\begin{array}{rlrl}
\langle\underline{\nabla} \cdot \underline{q}, \eta\rangle_{C(\bar{\Omega})} & =\langle f, \eta\rangle_{C(\bar{\Omega})} & \forall \eta \in C(\bar{\Omega}), \\
\langle | \underline{v}|, k\rangle-\langle | \underline{q}|, k\rangle \geq\langle\underline{\nabla} \cdot(\underline{v}-\underline{q}), u\rangle_{C(\bar{\Omega})} & \forall \underline{v} \in \widehat{\underline{V}}_{0}^{\mathcal{M}}(\Omega) ; \tag{1.17b}
\end{array}
$$

where

$$
\begin{align*}
\underline{V}_{0}^{\mathcal{M}}(\Omega):= & \left\{\underline{v} \in \underline{V}^{\mathcal{M}}(\Omega):\langle\underline{\nabla} \cdot \underline{v}, \eta\rangle_{C(\bar{\Omega})}=-\langle\underline{v}, \underline{\nabla} \eta\rangle_{C(\bar{\Omega})} \quad \forall \eta \in C^{1}(\bar{\Omega})\right\},  \tag{1.18a}\\
\underline{V}^{\mathcal{M}}(\Omega):= & \{\underline{v} \in \underline{\mathcal{M}}: \underline{\nabla} \cdot \underline{v} \in \mathcal{M}(\bar{\Omega})\},  \tag{1.18b}\\
\widehat{\widehat{V}}_{0}^{\mathcal{M}}(\Omega):= & \left\{\underline{v} \in \underline{V}_{0}^{\mathcal{M}}(\Omega): \exists \underline{v}_{j}=\left(\underline{v}_{j}^{(1)}, \ldots, \underline{v}_{j}^{(N)}\right) \in\left[C_{0}^{\infty}(\Omega)\right]^{n} \text { such that as } j \rightarrow \infty\right. \\
& \left.\underline{v}_{j}^{(i)} \equiv \underline{v}_{j}\right|_{\bar{\Omega}^{(i)}} \in\left[C^{\infty}\left(\overline{\Omega^{(i)}}\right)\right]^{n} \rightarrow \underline{v}^{(i)} \text { vaguely in }\left[\mathcal{M}\left(\overline{\Omega^{(i)}}\right)\right]^{n}, i=1 \rightarrow N, \\
& \left.\underline{\nabla} \cdot \underline{v}_{j} \rightarrow \underline{\nabla} \cdot \underline{v} \text { vaguely in } \mathcal{M}(\bar{\Omega}), \text { and } \limsup _{j \rightarrow \infty}\langle | \underline{v}_{j}|, k\rangle \leq\langle | \underline{v}|, k\rangle\right\} . \tag{1.18c}
\end{align*}
$$

It immediately follows from $(1.18 \mathrm{a}-\mathrm{c})$ and $(1.11 \mathrm{a}-\mathrm{c})$ that for any $r \in(1, \infty), \underline{V}_{0}^{r}(\Omega) \subset \underline{\widehat{V}}_{0}^{\mathcal{M}}(\Omega) \subset \underline{V}_{0}^{\mathcal{M}}(\Omega)$. If $N=1$, i.e. $k \in C(\bar{\Omega})$, one can show that $\widehat{V}_{0}^{\mathcal{M}}(\Omega) \equiv \underline{V}_{0}^{\mathcal{M}}(\Omega) \equiv \underline{\mathcal{V}}_{0}^{\mathcal{M}}(\Omega)$, see Lemma 2.4 below, and so (Q) collapses to ( $\widetilde{\mathrm{Q}})$. Unfortunately, for general $k$ satisfying (1.15) we are not able to show that $\underline{\widehat{V}}_{0}^{\mathcal{M}}(\Omega) \equiv \underline{V}_{0}^{\mathcal{M}}(\Omega)$.

Finally, throughout $C$ denotes a generic positive constant independent of the regularization parameter, $r \in$ $(1, \infty)$, and the mesh parameter $h$. Whereas, $C_{s}$ denotes a positive dependent on the parameter $s$.

## 2. EXISTENCE THEORY FOR (Q) VIA REGULARIZATION

Firstly, we gather together our assumptions on the data.
(A1) $\Omega \subset \mathbb{R}^{n}, n \geq 1$, is a connected bounded open set, with a Lipschitz boundary $\partial \Omega$ if $n \geq 2, f \in \mathcal{M}_{I}(\bar{\Omega})$, $k$ satisfies (1.15) and we set $k_{\max }:=\max _{i=1 \rightarrow N} \max _{\underline{x} \in \overline{\Omega^{(i)}}} k^{(i)}(\underline{x})$.

The aim of this paper is to prove existence of, and approximate, solutions to (Q), (1.17a,b).
For any $r>1$, we regularize the non-differentiable nonlinearity $|\cdot|$ by the strictly convex function $\frac{1}{r}|\cdot| r$. We note for all $\underline{b}, \underline{c} \in \mathbb{R}^{n}$, with $\underline{b} \neq \underline{c}$, that

$$
\begin{equation*}
\frac{1}{r} \frac{\partial|\underline{b}|^{r}}{\partial b_{i}}=|\underline{b}|^{r-2} b_{i} \quad \Rightarrow \quad|\underline{b}|^{r-2} \underline{b} \cdot(\underline{b}-\underline{c})>\frac{1}{r}\left[|\underline{b}|^{r}-|\underline{c}|^{r}\right] . \tag{2.1}
\end{equation*}
$$

For a given $r>1$, on setting $p=\frac{r}{r-1}$ such that $\frac{1}{r}+\frac{1}{p}=1$, we then consider the following regularization of (Q) for a given $f_{r} \in L_{I}^{r}(\Omega)$, a regularization of $f$ :
$\left(\mathbf{Q}_{r}\right)$ Find $\underline{q}_{r} \in \underline{V}_{0}^{r}(\Omega)$ and $u_{r} \in L_{I}^{p}(\Omega)$ such that

$$
\begin{align*}
\left(\underline{\nabla} \cdot \underline{q}_{r}, \eta\right) & =\left(f_{r}, \eta\right) & & \forall \eta \in L^{p}(\Omega),  \tag{2.2a}\\
\left(k\left|\underline{q}_{r}\right|^{r-2} \underline{q}_{r}, \underline{v}\right) & =\left(u_{r}, \underline{\nabla} \cdot \underline{v}\right) & & \forall \underline{v} \in \underline{V}_{0}^{r}(\Omega) ; \tag{2.2~b}
\end{align*}
$$

where $\left(\eta_{1}, \eta_{2}\right):=\int_{\Omega} \eta_{1} \eta_{2} \mathrm{~d} \underline{x}$, and is naturally extended to vector functions. We note that imposing the constraint $\left(u_{r}, 1\right)=0$ leads to uniqueness of $u_{r}$, see Theorem 2.2 below. Of course, if one imposed the constraint $(u, 1)=0$ in $(\mathrm{Q}),(1.17 \mathrm{a}, \mathrm{b})$, then this would still not lead to uniqueness of $u$; since one can only expect uniqueness for $u$ on the $\operatorname{supp}(f)$, recall (1.2).

It follows from (2.1) that a solution of $\left(\mathrm{Q}_{r}\right)$ is such that $\underline{q}_{r} \in \underline{X}\left(f_{r}\right):=\left\{\underline{v} \in \underline{V}_{0}^{r}(\Omega):(\underline{\nabla} \cdot \underline{v}, \eta)=\left(f_{r}, \eta\right) \forall \eta \in\right.$ $\left.L^{p}(\Omega)\right\}$ and

$$
\begin{equation*}
E\left(\underline{q}_{r}\right) \leq E(\underline{v}):=\frac{1}{r} \int_{\Omega} k|\underline{v}|^{r} \mathrm{~d} \underline{x} \quad \forall \underline{v} \in \underline{X}\left(f_{r}\right) . \tag{2.3}
\end{equation*}
$$

Lemma 2.1. Let the Assumptions (A1) hold. Then for all $r>1$ with $p=\frac{r}{(r-1)}$, given $\rho \in L_{I}^{r}(\Omega)$ there exists $\underline{q}_{r}^{\rho} \in \underline{X}(\rho)$ and

$$
\begin{equation*}
\left\|\underline{q}_{r}^{\rho}\right\|_{\underline{V}^{r}(\Omega)} \leq\left(C_{\Omega}+1\right)|\rho|_{0, r, \Omega} \tag{2.4}
\end{equation*}
$$

Hence it follows that

$$
\begin{equation*}
\inf _{\eta \in L_{I}^{p}(\Omega)} \sup _{\underline{v} \in \underline{V}_{0}^{r}(\Omega)} \frac{(\underline{\nabla} \cdot \underline{v}, \eta)}{\|\underline{v}\|_{\underline{v}^{r}(\Omega)}|\eta|_{0, p, \Omega}} \geq \beta^{-1} \tag{2.5}
\end{equation*}
$$

where $\beta=2\left(C_{\Omega}+1\right)$ and so is independent of $r$ and $p$.
Proof. We adapt the proof of Proposition 2.1 in [18]; where (2.5) is proved with $L_{I}^{p}(\Omega)$ and $\underline{V}_{0}^{r}(\Omega)$ replaced by $L^{p}(\Omega)$ and $\underline{V}^{r}(\Omega)$, respectively, for some constant $\beta>0$. We note that the uniformity of $\beta$, shown below, is crucial for later developments in this paper.

Given $\rho \in L_{I}^{r}(\Omega)$, let $w \in W_{I}^{1, p}(\Omega)$ be the unique solution to

$$
\begin{equation*}
\left(|\underline{\nabla} w|^{p-2} \underline{\nabla} w, \underline{\nabla} z\right)=(\rho, z) \quad \forall z \in W^{1, p}(\Omega) \tag{2.6}
\end{equation*}
$$

It follows from (2.6) and (1.10) that

$$
\begin{equation*}
|\underline{\nabla} w|_{0, p, \Omega}^{p}=(\rho, w) \leq|\rho|_{0, r, \Omega}|w|_{0, p, \Omega} \leq C_{\Omega}^{r}|\rho|_{0, r, \Omega}^{r} \tag{2.7}
\end{equation*}
$$

Let $\underline{q}_{r}^{\rho}=-|\underline{\nabla} w|^{p-2} \underline{\nabla} w$, and so $\left|\underline{q}_{r}^{\rho}\right|_{0, r, \Omega}^{r}=|\underline{\nabla} w|_{0, p, \Omega}^{p}$. It then follows from (2.6), (2.7) and (1.11c) that $\underline{q}_{r}^{\rho} \in \underline{X}(\rho)$ and the bound (2.4) holds.

Given $\eta \in L_{I}^{p}(\Omega)$, let $\rho:=(I-f)\left[|\eta|^{p-2} \eta\right]$, where $f \eta=\frac{1}{|\Omega|} \int_{\Omega} \eta \mathrm{d} \underline{x}$. It follows that $\rho \in L_{I}^{r}(\Omega)$ and, as $(b+c)^{r} \leq 2^{r-1}\left(b^{r}+c^{r}\right)$ for all $b, c \in \mathbb{R}_{\geq 0}$, that

$$
\begin{equation*}
|\rho|_{0, r, \Omega}^{r} \leq\left.\left.\int_{\Omega}|(I+f)| \eta\right|^{p-1}\right|^{r} \mathrm{~d} \underline{x} \leq 2^{r-1} \int_{\Omega}\left[|\eta|^{p}+\left(f|\eta|^{p-1}\right)^{r}\right] \mathrm{d} \underline{x} \leq 2^{r}|\eta|_{0, p, \Omega}^{p} \tag{2.8}
\end{equation*}
$$

Hence choosing $\underline{v}=\underline{q}_{r}^{\rho} \in \underline{X}(\rho)$ in (2.5), and noting (2.4), (2.8) and that $(\underline{\nabla} \cdot \underline{v}, \eta)=|\eta|_{0, p, \Omega}^{p}$, yields the desired result (2.5).

Theorem 2.2. Let the Assumptions (A1) hold. Then for any given $r>1$ there exists a unique solution, $\left(\underline{q}_{r}, u_{r}\right) \in \underline{V}_{0}^{r}(\Omega) \times L_{I}^{p}(\Omega)$, to $\left(\mathrm{Q}_{r}\right)$. In addition, we have that

$$
\begin{equation*}
\left|\underline{q}_{r}\right|_{0, r, \Omega} \leq\left\|\underline{q}_{r}\right\|_{\underline{V}^{r}(\Omega)} \leq C_{\star}\left|f_{r}\right|_{0, r, \Omega} \quad \text { and } \quad\left|u_{r}\right|_{0, p, \Omega} \leq \beta k_{\max } C_{\star}^{r-1}\left|f_{r}\right|_{0, r, \Omega}^{r-1}, \tag{2.9a}
\end{equation*}
$$

where $p=\frac{r}{r-1}, C_{\star}=\left(\frac{k_{\max }}{k_{\min }}\right)^{\frac{1}{r}}\left(C_{\Omega}+1\right)$ and $\beta=2\left(C_{\Omega}+1\right)$. Moreover, we have that

$$
\begin{equation*}
\left|\underline{\nabla} u_{r}\right|_{0, p, \Omega} \leq k_{\max } C_{\star}^{r-1}\left|f_{r}\right|_{0, r, \Omega}^{r-1} \tag{2.9b}
\end{equation*}
$$

Proof. For proving the existence and uniqueness of a solution to $\left(\mathrm{Q}_{r}\right)$ and the bounds (2.9a); we adapt the proof of Theorem 2.1 in [17], which is for $\left(\mathrm{Q}_{r}\right)$ with $k \equiv 1$ and $L_{I}^{p}(\Omega)$ and $\underline{V}_{0}^{r}(\Omega)$ replaced by $L^{p}(\Omega)$ and $\underline{V}^{r}(\Omega)$, respectively, and for some positive constant $C_{1}$. Once again, we need to prove the uniformity, as $r \rightarrow 1$, of these constants for later developments in this paper.

Firstly, (2.4) yields that there exists $\underline{q}_{r}^{f} \in \underline{X}\left(f_{r}\right)$ and

$$
\begin{equation*}
\left\|\underline{q}_{r}^{f}\right\|_{\underline{V}^{r}(\Omega)} \leq\left(C_{\Omega}+1\right)\left|f_{r}\right|_{0, r, \Omega} \tag{2.10}
\end{equation*}
$$

Therefore $\left(\mathrm{Q}_{r}\right)$ can be rewritten as: Find $\underline{\underline{q}}_{r}:=\underline{q}_{r}-\underline{q}_{r}^{f} \in \underline{X}(0)$ such that

$$
\begin{equation*}
\left(k\left|\underline{\widehat{q}}_{r}+\underline{q}_{r}^{f}\right|^{r-2}\left(\underline{\widehat{q}}_{r}+\underline{q}_{r}^{f}\right), \underline{v}\right)=0 \quad \forall \underline{v} \in \underline{X}(0) ; \tag{2.11}
\end{equation*}
$$

which, on recalling (2.3) and (2.1), is the Euler-Lagrange equation for the strictly convex minimization problem

$$
\begin{equation*}
\inf _{\underline{v} \in \underline{X}(0)} E\left(\underline{v}+\underline{q}_{r}^{f}\right) . \tag{2.12}
\end{equation*}
$$

Hence there exists a unique $\underline{\widehat{q}}_{r} \in \underline{X}(0)$ solving (2.11). It follows that $\underline{q}_{r}=\underline{q}_{r}^{f}+\underline{\widehat{q}}_{r}$ is unique and satisfies

$$
\begin{equation*}
k_{\min }\left|\underline{q}_{r}\right|_{0, r, \Omega}^{r} \leq k_{\max }\left|\underline{q}_{r}^{f}\right|_{0, r, \Omega}^{r} \quad \text { and } \quad\left|\underline{\nabla} \cdot \underline{q}_{r}\right|_{0, r, \Omega}=\left|\underline{\nabla} \cdot \underline{q}_{r}^{f}\right|_{0, r, \Omega} \tag{2.13}
\end{equation*}
$$

The bounds on $\underline{q}_{r}$ in (2.9a) follow immediately from (2.13) and (2.10).
Let $B \equiv \underline{\nabla} .$, then it follows from (1.11a,b) that $B: \underline{V}_{0}^{r}(\Omega) \rightarrow\left[L_{I}^{p}(\Omega)\right]^{*}$ and the dual operator $B^{*}: L_{I}^{p}(\Omega) \rightarrow$ $\left[\underline{V}_{0}^{r}(\Omega)\right]^{*}$. Moreover, it follows from (2.5), on noting Lemma I.4.1 and Remark I.4.2 in [21], that $B^{*}$ is an isomorphism from $L_{I}^{p}(\Omega)$ onto $Z:=\left\{\underline{v}^{*} \in\left[\underline{V}_{0}^{r}(\Omega)\right]^{*}:\left\langle\underline{v}^{*}, \underline{v}\right\rangle_{\underline{V}_{0}^{r}(\Omega)}=0 \quad \forall \underline{v} \in \operatorname{ker}(B) \subset \underline{V}_{0}^{r}(\Omega)\right\}$, where $\langle\cdot, \cdot\rangle_{\underline{V}_{0}^{r}(\Omega)}$ is the duality pairing on $\left[\underline{V}_{0}^{r}(\Omega)\right]^{*} \times \underline{V}_{0}^{r}(\Omega)$. Hence it follows that there exists a unique solution to $\left(\mathrm{Q}_{r}\right)$. In addition, we have from (2.5) and (2.2b) that

$$
\begin{equation*}
\left|u_{r}\right|_{0, p, \Omega} \leq \beta \sup _{\underline{v} \in \underline{V}_{0}^{r}(\Omega)} \frac{\left(k\left|\underline{q}_{r}\right|^{r-2} \underline{q}_{r}, \underline{v}\right)}{\|\underline{v}\|_{\underline{V}^{r}(\Omega)}} \leq \beta k_{\max }\left|\underline{q}_{r}\right|_{0, r, \Omega}^{r-1}, \tag{2.14}
\end{equation*}
$$

and hence the bound on $u_{r}$ in (2.9a).
Finally, it follows from (2.2b) that

$$
\begin{equation*}
\left|\left(u_{r}, \underline{\nabla} \cdot \underline{v}\right)\right|=\left|\left(k\left|\underline{q}_{r}\right|^{r-2} \underline{q}_{r}, \underline{v}\right)\right| \leq k_{\max }\left|\underline{q}_{r}\right|_{0, r, \Omega}^{r-1}|\underline{v}|_{0, r, \Omega} \quad \forall \underline{v} \in \underline{V}_{0}^{r}(\Omega) . \tag{2.15}
\end{equation*}
$$

This, and (2.9a), immediately yield the desired result (2.9b).
With $f \in \mathcal{M}_{I}(\bar{\Omega})$ being the data for problem (Q); we then choose, for any $r>1$, the corresponding data $f_{r}$ for the regularized problem $\left(\mathrm{Q}_{r}\right)$ as follows. If $f \in L_{I}^{r}(\Omega)$ we set $f_{r} \equiv f$. If $\operatorname{supp}(f) \subset \Omega$ we set

$$
\begin{equation*}
f_{r}(\underline{x}):=\frac{1}{|B(\underline{x}, r-1)|}\langle f, 1\rangle_{C(B(\underline{x}, r-1))} \quad \forall \underline{x} \in \Omega, \tag{2.16}
\end{equation*}
$$

where $B(\underline{x}, \rho)$ is the closed ball in $\mathbb{R}^{n}$ centred at $\underline{x}$ with radius $\rho>0$, and $f$ is extended from $\bar{\Omega}$ to $\mathbb{R}^{n}$ by zero. It follows for $r-1$ sufficiently small that

$$
\begin{equation*}
f_{r} \in L_{I}^{r}(\Omega), \quad \limsup _{r \rightarrow 1}\left|f_{r}\right|_{0, r, \Omega}^{r} \leq \int_{\bar{\Omega}}|f| \quad \text { and } \quad f_{r} \rightarrow f \quad \text { vaguely in } \mathcal{M}(\bar{\Omega}) \quad \text { as } r \rightarrow 1 \tag{2.17}
\end{equation*}
$$

For general $f \in \mathcal{M}_{I}(\bar{\Omega})$, the construction (2.16) can be modified, so that (2.17) still holds. For example, one can partition $\Omega$ into a finite number of strictly star-shaped domains with "centres" $\underline{x}_{\ell}$ and employ the local change of variable $f_{t}(\underline{x})=f\left(\underline{x}_{\ell}+t^{-1}\left(\underline{x}-\underline{x}_{\ell}\right)\right)$ for $t \in(0,1)$ before mollifying; for details see the proof of Lemma 2.4 below, where these techniques are used.
Theorem 2.3. Let the Assumptions (A1) hold. Then there exists a subsequence $\left\{\left(\underline{q}_{r_{j}}, u_{r_{j}}\right)\right\}_{r_{j}>1}$ of $\left\{\left(\underline{q}_{r}, u_{r}\right)\right\}_{r>1}$, where $\left(\underline{q}_{r}, u_{r}\right) \in \underline{V}_{0}^{r}(\Omega) \times L_{I}^{p}(\Omega)$ is the unique solution of $\left(\mathrm{Q}_{r}\right)$ with $f_{r}$ chosen as in (2.16), such that as $r_{j} \rightarrow 1$

$$
\begin{align*}
\underline{q}_{r_{j}} \mid \bar{\Omega}^{(i)} & \rightarrow \underline{q}^{(i)} & & \text { vaguely in }\left[\mathcal{M}\left(\overline{\Omega^{(i)}}\right)\right]^{n}, \quad i=1 \rightarrow N,  \tag{2.18a}\\
\underline{\nabla} \cdot \underline{q}_{r_{j}} & \rightarrow \underline{\nabla} \cdot \underline{q} & & \text { vaguely in } \mathcal{M}(\bar{\Omega}),  \tag{2.18b}\\
u_{r_{j}} & \rightarrow u & & \text { strongly in } C(\bar{\Omega}) . \tag{2.18c}
\end{align*}
$$

Moreover, $(\underline{q}, u) \in \underline{V}_{0}^{\mathcal{M}}(\Omega) \times C_{I}(\bar{\Omega})$ solves $(\mathrm{Q}),(1.17 \mathrm{a}, \mathrm{b})$.
Proof. It follows from $(2.9 \mathrm{a}, \mathrm{b})$ and $(2.17)$ that for all $r \in\left(1, \frac{n}{n-1}\right)$

$$
\begin{equation*}
\left|\underline{q}_{r}\right|_{0,1, \Omega}+\left|\underline{\nabla} \cdot \underline{q}_{r}\right|_{0,1, \Omega}+\left\|u_{r}\right\|_{1, p^{\star}, \Omega} \leq C, \tag{2.19}
\end{equation*}
$$

where $p^{\star}>n$. The subsequence convergence results (2.18a-c), where $(\underline{q}, u) \in \underline{V}^{\mathcal{M}}(\Omega) \times C(\bar{\Omega})$ then follow immediately from (2.19), and noting the compact Sobolev embedding $W^{1, p^{\star}}(\Omega) \hookrightarrow \hookrightarrow C(\bar{\Omega})$. As $\underline{q}_{r} \cdot \underline{\nu}=0$ on $\partial \Omega$ and $\left(u_{r}, 1\right)=0$ for all $r>1$, it follows that the limits $(\underline{q}, u) \in \underline{V}_{0}^{\mathcal{M}}(\Omega) \times C_{I}(\bar{\Omega})$.

Passing to the $r_{j} \rightarrow 1$ limit in the $r_{j}$ version (2.2a) with $\eta \in C(\bar{\Omega})$ yields, on noting (2.18b) and (2.17), that (1.17a) holds.

For any $\underline{\xi} \in\left[C_{0}^{\infty}(\Omega)\right]^{n}$, we choose $\underline{v}=\underline{\xi}-\underline{q}_{r_{j}}$ in the $r_{j}$ version of (2.2b) to obtain, on noting (2.1), that

$$
\begin{equation*}
\left(u_{r_{j}}, \underline{\nabla} \cdot\left(\underline{q}_{r_{j}}-\underline{\xi}\right)\right)=\left(k\left|\underline{q}_{r_{j}}\right|^{r_{j}-2} \underline{q}_{r_{j}}, \underline{q}_{r_{j}}-\underline{\xi}\right) \geq \frac{1}{r_{j}} \int_{\Omega} k\left|\underline{q}_{r_{j}}\right|^{r_{j}} \mathrm{~d} \underline{x}-\frac{1}{r_{j}} \int_{\Omega} k|\underline{\xi}|^{r_{j}} \mathrm{~d} \underline{x} . \tag{2.20}
\end{equation*}
$$

For all $\underline{\xi} \in\left[C_{0}^{\infty}(\Omega)\right]^{n}$, we have that

$$
\begin{equation*}
\frac{1}{r} \int_{\Omega} k|\underline{\underline{\xi}}|^{r} \mathrm{~d} \underline{x} \rightarrow \int_{\Omega} k|\underline{\xi}| \mathrm{d} \underline{x} \quad \text { as } r \rightarrow 1 \tag{2.21}
\end{equation*}
$$

and it follows from (2.18b, c) that

$$
\begin{equation*}
\left(\underline{\nabla} \cdot\left(\underline{q}_{r_{j}}-\underline{\xi}\right), u_{r_{j}}\right) \rightarrow\langle\underline{\nabla} \cdot(\underline{q}-\underline{\xi}), u\rangle_{C(\bar{\Omega})} \quad \text { as } r_{j} \rightarrow 1 . \tag{2.22}
\end{equation*}
$$

Next we note from (2.18a), (1.14) and (1.16) that

$$
\begin{equation*}
\liminf _{r_{j} \rightarrow 1} \frac{1}{r_{j}} \int_{\Omega} k\left|\underline{q}_{r_{j}}\right|^{r_{j}} \mathrm{~d} \underline{x} \geq \liminf _{r_{j} \rightarrow 1} \int_{\Omega} k\left|\underline{q}_{r_{j}}\right| \mathrm{d} \underline{x} \geq\langle | \underline{q}|, k\rangle . \tag{2.23}
\end{equation*}
$$

Combining (2.20)-(2.23) yields that

$$
\begin{equation*}
\langle | \underline{\xi}|, k\rangle-\langle | \underline{q}|, k\rangle \geq\langle\underline{\nabla} \cdot(\underline{\xi}-\underline{q}), u\rangle_{C(\bar{\Omega})} \quad \forall \underline{\xi} \in\left[C_{0}^{\infty}(\Omega)\right]^{n} . \tag{2.24}
\end{equation*}
$$

Finally, noting the density of $\left[C_{0}^{\infty}(\Omega)\right]^{n}$ in $\widehat{V}_{0}^{\mathcal{M}}(\Omega)$, recall (1.18c), yields the desired result (1.17b). Hence we have that $(\underline{q}, u) \in \underline{V}_{0}^{\mathcal{M}}(\Omega) \times C_{I}(\bar{\Omega})$ solves $(\mathrm{Q})$.

Let

$$
\begin{equation*}
\underline{X}^{\mathcal{M}}(f):=\left\{\underline{v} \in \underline{V}_{0}^{\mathcal{M}}(\Omega):\langle\underline{\nabla} \cdot \underline{v}, \eta\rangle_{C(\bar{\Omega})}=\langle f, \eta\rangle_{C(\bar{\Omega})} \quad \forall \eta \in C(\bar{\Omega})\right\} . \tag{2.25}
\end{equation*}
$$

If $\underline{\hat{V}}_{0}^{\mathcal{M}}(\Omega) \equiv \underline{V}_{0}^{\mathcal{M}}(\Omega)$, it follows immediately from (1.17b) and (2.25) that the solution $\underline{q} \in \underline{V}_{0}^{\mathcal{M}}(\Omega)$ of (Q) is such that $\underline{q} \in \underline{X}^{\mathcal{M}}(f)$ and

$$
\begin{equation*}
\langle | \underline{q}|, k\rangle \leq\langle | \underline{v}|, k\rangle \quad \forall \underline{v} \in \underline{X}^{\mathcal{M}}(f) ; \tag{2.26}
\end{equation*}
$$

that is, it solves the associated minimization problem. We end this section with the following density result.
Lemma 2.4. If $N=1$, i.e. $k \in C(\bar{\Omega})$, then $\widehat{\hat{V}}_{0}^{\mathcal{M}}(\Omega) \equiv \underline{V}_{0}^{\mathcal{M}}(\Omega) \equiv \underline{\mathcal{V}}_{0}^{\mathcal{M}}(\Omega)$.
Proof. The proof is based on the standard techniques of partition of unity, mollification and change of variable. Let $\left\{G_{\ell}\right\}_{\ell=1}^{L}$ be open sets such that $\bar{\Omega} \subset \bigcup_{\ell=1}^{L} G_{\ell}$ and $G_{\ell} \cap \Omega$ are Lipschitz-continuous and strictly star-shaped with respect to $\underline{x}_{\ell}, \ell=1 \rightarrow L$; see [21], p. 33. Let $\xi_{\ell} \in C_{0}^{\infty}\left(G_{\ell}\right)$ with $\xi_{\ell} \geq 0, \ell=1 \rightarrow L$, and $\sum_{\ell=1}^{L} \xi_{\ell}(\underline{x})=1$ for all $\underline{x} \in \bar{\Omega}$. Given $\underline{v} \in \underline{\mathcal{V}}_{0}^{\mathcal{M}}(\Omega) \equiv \underline{V}_{0}^{\mathcal{M}}(\Omega)$, we extend $\underline{v}$ from $\bar{\Omega}$ by zero to $\mathbb{R}^{n}$ and set $\underline{v}_{\ell}=\xi_{\ell} \underline{v}, \ell=1 \rightarrow L$. The constraint in (1.9a) yields that $\underline{\nabla} \cdot \underline{v}_{\ell} \in \mathcal{M}\left(\mathbb{R}^{n}\right), \ell=1 \rightarrow L$.

For $\ell=1 \rightarrow L$ and for $t \in(0,1), \underline{v}_{\ell, t}(\underline{x})=\underline{v}_{\ell}\left(\underline{x}_{\ell}+t^{-1}\left(\underline{x}-\underline{x}_{\ell}\right)\right)$ for all $\underline{x} \in \mathbb{R}^{n}$ has support in $G_{\ell} \cap \Omega$ and

$$
\begin{align*}
\underline{v}_{\ell, t} & \rightarrow \underline{v}_{\ell} & & \text { vaguely in }[\mathcal{M}(\bar{\Omega})]^{n}
\end{aligned} \quad \begin{aligned}
& \text { as } t \rightarrow 1,  \tag{2.27a}\\
& \underline{\nabla} \cdot \underline{v}_{\ell, t}
\end{align*} \rightarrow \underline{\nabla} \cdot \underline{v}_{\ell} \quad \begin{array}{ll}
\text { vaguely in } \mathcal{M}(\bar{\Omega}) & \tag{2.27b}
\end{array}
$$

Let $j \in C^{\infty}\left(\mathbb{R}^{n}\right)$, with compact support in $B(\underline{0}, 1)$, such that

$$
\int_{B(\underline{0}, 1)} j(\underline{x}) \mathrm{d} \underline{x}=1, \quad j(\underline{x}) \geq 0 \quad \text { and } \quad j(-\underline{x})=j(\underline{x})
$$

Let $J_{\varepsilon}: \mathcal{M}\left(\mathbb{R}^{n}\right) \rightarrow C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$ be such that

$$
\left(J_{\varepsilon} \eta\right)(\underline{x})=\left\langle\eta(\cdot), j_{\varepsilon}(\underline{x}-\cdot)\right\rangle_{C\left(\mathbb{R}^{n}\right)} \quad \forall \underline{x} \in \mathbb{R}^{n}
$$

where $j_{\varepsilon}(\underline{x})=\varepsilon^{-n} j\left(\varepsilon^{-1} \underline{x}\right)$. We extend $J_{\varepsilon}$ in the natural way, so that $\underline{J}_{\varepsilon}:\left[\mathcal{M}\left(\mathbb{R}^{n}\right)\right]^{n} \rightarrow\left[C_{0}^{\infty}\left(\mathbb{R}^{n}\right)\right]^{n}$. Then for $\ell=1 \rightarrow L$ and $t \in(0,1), \underline{v}_{\ell, t, \varepsilon}=\underline{J}_{\varepsilon} \underline{v}_{\ell, t} \in\left[C_{0}^{\infty}\left(\mathbb{R}^{n}\right)\right]^{n}$ has support in $G_{\ell} \cap \Omega$ for $\varepsilon$ sufficiently small and

$$
\begin{align*}
\underline{v}_{\ell, t, \varepsilon} & \rightarrow \underline{v}_{\ell, t} & \text { vaguely in }[\mathcal{M}(\bar{\Omega})]^{n} & \text { as } \varepsilon \rightarrow 0  \tag{2.28a}\\
\underline{\nabla} \cdot \underline{v}_{\ell, t, \varepsilon}=J_{\varepsilon}\left(\underline{\nabla} \cdot \underline{v}_{\ell, t}\right) & \rightarrow \underline{\nabla} \cdot \underline{v}_{\ell, t} & \text { vaguely in } \mathcal{M}(\bar{\Omega}) & \text { as } \varepsilon \rightarrow 0  \tag{2.28b}\\
\limsup _{\varepsilon \rightarrow 0} \int_{\Omega} k\left|\underline{v}_{\ell, t, \varepsilon}\right| \mathrm{d} \underline{x} & \leq \int_{\bar{\Omega}} k\left|\underline{v}_{\ell, t}\right| . & & \tag{2.28c}
\end{align*}
$$

We note that the results (2.27c) and (2.28c) exploit the fact that $k \in C(\bar{\Omega})$. In addition, the introduction of the change of variable involving $t \in(0,1)$ in the above is really only required to ensure that $\underline{v}_{\ell, t, \varepsilon} \in\left[C_{0}^{\infty}(\Omega)\right]^{n}$ for $\varepsilon$ sufficiently small for those $\ell$ such that $G_{\ell} \cap \partial \Omega \neq \emptyset$. Therefore given $\underline{v} \in \underline{\mathcal{V}}_{0}^{\mathcal{M}}(\Omega) \equiv \underline{V}_{0}^{\mathcal{M}}(\Omega)$, then $\underline{v}_{t, \varepsilon}=\sum_{l=1}^{L} \underline{v}_{\ell, t, \varepsilon} \in\left[C_{0}^{\infty}(\Omega)\right]^{n}$ for any $t \in(0,1)$ and $\varepsilon$ sufficiently small. A diagonal subsequence argument yields the desired convergence, recall (1.18c) with $N=1$, as $t \rightarrow 1$ and $\varepsilon \rightarrow 0$ on noting (2.27a-c), (2.28a-c) and that $\sum_{l=1}^{L} \int_{\bar{\Omega}} k\left|\underline{v}_{\ell}\right|=\int_{\bar{\Omega}} k|\underline{v}|$.

Therefore if $N=1$, i.e. $k \in C(\bar{\Omega})$, then (Q) collapses to $(\widetilde{\mathrm{Q}})$; and hence we have existence of a solution to the associated minimization problem. We note, however, that even for a positive lower semicontinuous function $k$ existence of a solution to the associated minimization problem can be proved directly using the Reshetnyak lower semicontinuity theorem; see e.g. [3], Theorem 2.38.

## 3. Finite element approximation of $\left(\mathrm{Q}_{r}\right)$

For ease of exposition, we assume the following.
(A2) Let $n \leq 3$ with $\Omega$ and $\Omega^{(i)}, i=1 \rightarrow N$, polyhedral, if $n \geq 2$. Let $\left\{\mathcal{T}^{h} \equiv \cup_{i=1}^{N} \mathcal{T}_{i}^{h}\right\}_{h>0}$ be a regular family of partitionings of $\Omega$ into disjoint open simplices $\sigma$ with $h_{\sigma}:=\operatorname{diam}(\sigma)$ and $h:=\max _{\sigma \in \mathcal{T}^{h}} h_{\sigma}$, so that

$$
\begin{equation*}
\bar{\Omega}=\cup_{\sigma \in \mathcal{T}^{h}} \bar{\sigma}, \quad \text { with } \quad \overline{\Omega^{(i)}}=\cup_{\sigma \in \mathcal{T}_{i}^{h}} \bar{\sigma}, \quad i=1 \rightarrow N . \tag{3.1}
\end{equation*}
$$

Let $\underline{\nu}_{\partial \sigma}$ be the outward unit normal to $\partial \sigma$, the boundary of $\sigma$. We then introduce

$$
\begin{align*}
& \underline{V}^{h}:=\left\{\underline{v}^{h} \in\left[L^{\infty}(\Omega)\right]^{n}:\left.\underline{v}^{h}\right|_{\sigma}=\underline{a}_{\sigma}+b_{\sigma} \underline{x}, \underline{a}_{\sigma} \in \mathbb{R}^{n}, b_{\sigma} \in \mathbb{R} \quad \forall \sigma \in \mathcal{T}^{h}\right. \text { and } \\
&\left.\underline{v}^{h} \cdot \underline{\nu}_{\partial \sigma} \text { is continuous across simplex boundaries }\right\} \\
&\left.\qquad \subset \underline{v} \in\left[L^{\infty}(\Omega)\right]^{n}: \underline{\nabla} \cdot \underline{v} \in L^{\infty}(\Omega)\right\},  \tag{3.2a}\\
& S^{h}:=\left\{\eta^{h} \in L^{\infty}(\Omega):\left.\eta^{h}\right|_{\sigma}=c_{\sigma} \in \mathbb{R} \quad \forall \sigma \in \mathcal{T}^{h}\right\},  \tag{3.2b}\\
& \underline{V}_{0}^{h}:=\left\{\underline{v}^{h} \in \underline{V}^{h}: \underline{v}^{h} \cdot \underline{\nu}=0 \text { on } \partial \Omega\right\} \quad \text { and } \quad S_{I}^{h}:=\left\{\eta^{h} \in S^{h}:\left(\eta^{h}, 1\right)=0\right\} . \tag{3.2c}
\end{align*}
$$

In order for our finite element approximation to be practical, we introduce $(\underline{v}, \underline{z})^{h}:=\sum_{\sigma \in \mathcal{T}^{h}}(\underline{v}, \underline{z})_{\sigma}^{h}$ with

$$
\begin{equation*}
(\underline{v}, \underline{z})_{\sigma}^{h}:=\frac{1}{n+1}|\sigma| \sum_{j=1}^{n+1} \underline{v}\left(P_{j}^{\sigma}\right) \cdot \underline{z}\left(P_{j}^{\sigma}\right) \quad \forall \underline{v}, \underline{z} \in[C(\bar{\sigma})]^{n}, \quad \forall \sigma \in \mathcal{T}^{h}, \tag{3.3}
\end{equation*}
$$

where $\left\{P_{j}^{\sigma}\right\}_{j=1}^{n+1}$ are the vertices of $\sigma$. Therefore $(\underline{v}, \underline{z})^{h}$ averages the integrand $\underline{v} \cdot \underline{z}$ over each simplex $\sigma$ at its vertices and hence is exact if $\underline{v} \cdot \underline{z}$ is piecewise linear over the partitioning $\mathcal{T}^{h}$. For any $r \geq 1$ and for any
$\underline{v}^{h} \in \underline{V}^{h}$, we have from the equivalence of norms for $\underline{v}^{h}$ and the convexity of $|\cdot|^{r}$ that

$$
\begin{equation*}
C_{r}\left(\left|\underline{v}^{h}\right|^{r}, 1\right)_{\sigma}^{h} \leq \int_{\sigma}\left|\underline{v}^{h}\right|^{r} \mathrm{~d} \underline{x} \leq\left(\left|\underline{h}^{h}\right|^{r}, 1\right)_{\sigma}^{h}:=\frac{1}{n+1}|\sigma| \sum_{j=1}^{n+1}\left|\underline{v}^{h}\left(P_{j}^{\sigma}\right)\right|^{r} \quad \forall \sigma \in \mathcal{T}^{h} \tag{3.4}
\end{equation*}
$$

Our fully practical approximation of $\left(\mathrm{Q}_{r}\right)$ by $\underline{V}_{0}^{h} \times S_{I}^{h}$, on employing (3.3), is then:
$\left(\mathbf{Q}_{r}^{h}\right)$ Find $\underline{q}_{r}^{h} \in \underline{V}_{0}^{h}$ and $u_{r}^{h} \in S_{I}^{h}$ such that

$$
\begin{align*}
\left(\underline{\nabla} \cdot \underline{q}_{r}^{h}, \eta^{h}\right) & =\left(f_{r}, \eta^{h}\right) & & \forall \eta^{h} \in S^{h},  \tag{3.5a}\\
\left(k\left|\underline{q}_{r}^{h}\right|^{r-2} \underline{q}_{r}^{h}, \underline{v}^{h}\right)^{h} & =\left(u_{r}^{h}, \underline{\nabla} \cdot \underline{v}^{h}\right) & & \forall \underline{v}^{h} \in \underline{V}_{0}^{h} . \tag{3.5b}
\end{align*}
$$

It follows from (2.1) that a solution of $\left(\mathrm{Q}_{r}^{h}\right)$ is such that $\underline{q}_{r}^{h} \in \underline{X}^{h}\left(f_{r}\right):=\left\{\underline{v}^{h} \in \underline{V}_{0}^{h}:\left(\underline{\nabla} \cdot \underline{v}^{h}, \eta^{h}\right)=\left(f_{r}, \eta^{h}\right) \forall \eta^{h} \in\right.$ $\left.S^{h}\right\}$ and

$$
\begin{equation*}
E^{h}\left(\underline{q}_{r}^{h}\right) \leq E^{h}\left(\underline{v}^{h}\right):=\frac{1}{r}\left(k,\left|\underline{v}^{h}\right|^{r}\right)^{h} \quad \forall \underline{v}^{h} \in \underline{X}^{h}\left(f_{r}\right) \tag{3.6}
\end{equation*}
$$

Let $P^{h}: L^{1}(\Omega) \rightarrow S^{h}$ be such that

$$
\begin{equation*}
\left(\left(I-P^{h}\right) z, \eta^{h}\right)=0 \quad \forall \eta^{h} \in S^{h} \tag{3.7}
\end{equation*}
$$

For $r>1$, let $\underline{I}^{h}: \underline{V}_{0}^{r}(\Omega) \cap\left[W^{1, r}(\Omega)\right]^{n} \rightarrow \underline{V}_{0}^{h}$ be the generalised interpolation operator satisfying

$$
\begin{equation*}
\int_{\partial_{i} \sigma}\left(\underline{v}-\underline{I}^{h} \underline{v}\right) \cdot \underline{\nu}_{\partial_{i} \sigma} \mathrm{~d} s=0 \quad i=1 \rightarrow n+1, \quad \forall \sigma \in \mathcal{T}^{h} \tag{3.8}
\end{equation*}
$$

where $\partial \sigma \equiv \cup_{i=1}^{n+1} \partial_{i} \sigma$ and $\underline{\nu}_{\partial_{i} \sigma}$ are the corresponding outward unit normals on $\partial_{i} \sigma$. It follows that

$$
\begin{equation*}
\left(\underline{\nabla} \cdot\left(\underline{v}-\underline{I}^{h} \underline{v}\right), \eta^{h}\right)=0 \quad \forall \eta^{h} \in S^{h} \tag{3.9}
\end{equation*}
$$

In addition, we have for all $\sigma \in \mathcal{T}^{h}$ and any $s \in(1, \infty]$ that

$$
\begin{equation*}
\left|\underline{v}-\underline{I}^{h} \underline{v}\right|_{0, s, \sigma} \leq C_{s} h_{\sigma}|\underline{v}|_{1, s, \sigma} \quad \text { and } \quad\left|\underline{I}^{h} \underline{v}\right|_{1, s, \sigma} \leq C_{s}|\underline{v}|_{1, s, \sigma} \tag{3.10}
\end{equation*}
$$

e.g. see [17], Lemma 3.1 and the proof given there for $s \geq 2$ is also valid for any $s \in(1, \infty]$.

Furthermore, we note from (2.1), (3.3) and (3.10) that for all $\underline{v} \in\left[W^{2, \infty}(\Omega)\right]^{n}$

$$
\begin{align*}
\left.\left|\int_{\Omega}\right| \underline{v}\right|^{r} \mathrm{~d} \underline{x}-\left(\left|\underline{I}^{h} \underline{v}\right|^{r}, 1\right)^{h} \mid & \leq\left.\int_{\Omega}| | \underline{v}\right|^{r}-\left|\underline{I}^{h} \underline{v}\right|^{r}\left|\mathrm{~d} \underline{x}+\left|\int_{\Omega}\right| \underline{I}^{h} \underline{v}\right|^{r} \mathrm{~d} \underline{x}-\left(\left|\underline{I}^{h} \underline{v}\right|^{r}, 1\right)^{h} \mid \\
& \leq r\left(|\underline{v}|_{0, \infty, \Omega}^{r-1}+\left|\underline{I}^{h} \underline{v}\right|_{0, \infty, \Omega}^{r-1}\right)\left[\left|\underline{v}-\underline{I}^{h} \underline{v}\right|_{0,1, \Omega}+\left.h \underline{I}^{h} \underline{v}\right|_{1,1, \Omega}\right] \\
& \leq C_{r} h\|\underline{v}\|_{2, \infty, \Omega}^{r} \tag{3.11}
\end{align*}
$$

We have the following discrete version of Lemma 2.1:
Lemma 3.1. Let the Assumptions (A1) and (A2) hold. Then for all $r \in\left(1, \frac{4}{3}\right)$ with $p=\frac{r}{(r-1)}$, given $\rho^{h} \in S_{I}^{h}$ there exists $\underline{q}_{r}^{\rho^{h}, h} \in \underline{X}^{h}\left(\rho^{h}\right)$ and

$$
\begin{equation*}
\left\|\underline{q}_{r}^{\rho^{h}, h}\right\|_{\underline{V}^{r}(\Omega)} \leq \mu_{r}\left|\rho^{h}\right|_{0, r, \Omega} \tag{3.12}
\end{equation*}
$$

where $\mu_{r} \in \mathbb{R}_{>0}$ is independent of $h$, but possibly dependent on $r$. Hence it follows that

$$
\begin{equation*}
\inf _{\eta^{h} \in S_{I}^{h}} \sup _{\underline{v}^{h} \in \underline{V}_{0}^{h}} \frac{\left(\underline{\nabla} \cdot \underline{v}^{h}, \eta^{h}\right)}{\left\|\underline{v}^{h}\right\|_{\underline{V}^{r}(\Omega)}\left|\eta^{h}\right|_{0, p, \Omega}} \geq\left[2 \mu_{r}\right]^{-1} . \tag{3.13}
\end{equation*}
$$

Proof. Given $\rho^{h} \in S_{I}^{h}$, the construction in the proof of Lemma 2.1 yields a $\underline{q}^{\rho^{h}} \in \underline{X}\left(\rho^{h}\right)$ satisfying the $r$ independent bound (2.4) with $\rho$ replaced by $\rho^{h}$. Unfortunately, $\underline{I}^{h}$ is not well-defined on $\underline{V}_{0}^{r}(\Omega)$. Hence, we use the following alternative construction.

Given $\rho^{h} \in S_{I}^{h}$, let $\widehat{w} \in H_{I}^{1}(\Omega)$ be the unique solution to

$$
\begin{equation*}
(\nabla \widehat{w}, \nabla z)=\left(\rho^{h}, z\right) \quad \forall z \in H^{1}(\Omega) \tag{3.14}
\end{equation*}
$$

As $\Omega$ is polyhedral and $r \in\left(1, \frac{4}{3}\right)$, it follows from elliptic regularity that

$$
\begin{equation*}
\|\widehat{w}\|_{2, r, \Omega} \leq C_{r}\left|\rho^{h}\right|_{0, r, \Omega} \tag{3.15}
\end{equation*}
$$

e.g. see [22], Sections 4 and 8.2. Setting $\widehat{\underline{q}}_{r}^{\rho^{h}}=-\underline{\nabla} \widehat{w}$, it follows from (3.14) and (3.15) that $\underline{\underline{q}}_{r}^{\rho^{h}} \in \underline{X}\left(\rho^{h}\right)$ and $\left\|\underline{\widehat{q}}_{r}^{\rho^{h}}\right\|_{1, r, \Omega} \leq C_{r}\left|\rho^{h}\right|_{0, r, \Omega}$. Hence, on choosing $\underline{q}_{r}^{\rho^{h}, h}=\underline{I}^{h} \underline{\underline{q}}_{r}^{\rho^{h}} \in \underline{X}^{h}\left(\rho^{h}\right)$, it follows from (3.8), (3.9) and (3.10) that the desired bound (3.12) holds.

Given $\eta^{h} \in S_{I}^{h}$, let $\rho^{h}:=(I-f)\left[\left|\eta^{h}\right|^{p-2} \eta^{h}\right]$. It follows that $\rho^{h} \in S_{I}^{h}(\Omega)$ and, similarly to (2.8), that $\left|\rho^{h}\right|_{0, r, \Omega}^{r} \leq$ $2^{r}\left|\eta^{h}\right|_{0, p, \Omega}^{p}$. Hence choosing $\underline{v}^{h}=\underline{I}^{h} \widehat{\underline{q}}_{r}^{\rho^{h}} \in \underline{X}^{h}\left(\rho^{h}\right)$ in (3.13), and noting (3.12) and that $\left(\underline{\nabla} \cdot \underline{v}^{h}, \eta^{h}\right)=\left|\eta^{h}\right|_{0, p, \Omega}^{p}$ yield the desired result (3.13).
Lemma 3.2. Let the Assumptions (A1) and (A2) hold. Then for all $r \in\left(1, \frac{4}{3}\right)$, there exists a unique solution, $\left(\underline{q}_{r}^{h}, u_{r}^{h}\right) \in \underline{V}_{0}^{h} \times S_{I}^{h}$, to $\left(\mathrm{Q}_{r}^{h}\right)$. In addition, we have that

$$
\begin{equation*}
\left\|\underline{q}_{r}^{h}\right\|_{\underline{V}^{r}(\Omega)} \leq C_{r}\left|f_{r}\right|_{0, r, \Omega} \quad \text { and } \quad\left|u_{r}^{h}\right|_{0, p, \Omega} \leq C_{r}\left|f_{r}\right|_{0, r, \Omega}^{r-1}, \tag{3.16}
\end{equation*}
$$

where $p=\frac{r}{r-1}$ and $C_{r}$ is independent of $h$.
Proof. On setting $\underline{\hat{q}}_{r}^{h}:=\underline{q}_{r}^{h}-\underline{q}_{r}^{f, h}$, where $\underline{q}_{r}^{f, h} \in \underline{X}^{h}\left(P^{h} f_{r}\right),\left(\mathrm{Q}_{r}^{h}\right)$ can be rewritten as: Find $\underline{\widehat{q}}_{r}^{h} \in \underline{X}^{h}(0)$ such that

$$
\begin{equation*}
\left(k\left|\underline{\hat{q}}_{r}^{h}+\underline{q}_{r}^{f, h}\right|^{r-2}\left(\underline{\widehat{q}}_{r}^{h}+\underline{q}_{r}^{f, h}\right), \underline{v}^{h}\right)^{h}=0 \quad \forall \underline{v}^{h} \in \underline{X}^{h}(0) ; \tag{3.17}
\end{equation*}
$$

which is the Euler-Lagrange equation for the strictly convex minimization problem

$$
\begin{equation*}
\inf _{\underline{v}^{h} \in \underline{X}^{h}(0)} E^{h}\left(\underline{v}^{h}+\underline{q}_{r}^{f, h}\right) . \tag{3.18}
\end{equation*}
$$

Hence there exists a unique $\underline{\widehat{q}}_{r}^{h} \in \underline{X}^{h}(0)$ solving (3.17). It follows from (2.1) that $\underline{q}_{r}^{h}=\underline{q}_{r}^{f, h}+\underline{\hat{q}}_{r}^{h}$ is unique and satisfies

$$
\begin{equation*}
\left(k_{\min },\left|\underline{q}_{r}^{h}\right|^{r}\right)^{h} \leq\left(k_{\max },\left|\underline{q}_{r}^{f, h}\right|^{r}\right)^{h} \quad \text { and } \quad\left|\underline{\nabla} \cdot \underline{q}_{r}^{h}\right|_{0, r, \Omega}=\left|\underline{\nabla} \cdot \underline{q}_{r}^{f, h}\right|_{0, r, \Omega} \tag{3.19}
\end{equation*}
$$

The bounds on $\underline{q}_{r}^{h}$ in (3.16) follow immediately from (3.19), (3.4) and (3.12).
Similarly to the proof of Lemma 2.2, we have that (3.13) yields the existence and uniqueness of $u_{r}^{h}$ solving (3.5b); and hence there exists a unique solution to $\left(\mathrm{Q}_{r}^{h}\right)$. In addition, we have from (3.13), (3.5b) and (3.4) that

$$
\begin{equation*}
\left|u_{r}^{h}\right|_{0, p, \Omega} \leq 2 \mu_{r} \sup _{\underline{v}^{h} \in \underline{V}_{0}^{h}} \frac{\left(k\left|\underline{q}_{r}^{h}\right|^{r-2} \underline{q}_{r}, \underline{v}^{h}\right)^{h}}{\left\|\underline{v}^{h}\right\|_{\underline{V}^{r}}(\Omega)} \leq C_{r}\left[\left(\left|\underline{q}_{r}^{h}\right|^{r}, 1\right)^{h}\right]^{\frac{r-1}{r}} . \tag{3.20}
\end{equation*}
$$

Hence the bound on $u_{r}^{h}$ in (3.16) follows from (3.20), (3.4) and the bound on $\underline{q}_{r}^{h}$ in (3.16).
Theorem 3.3. Let the Assumptions (A1) and (A2) hold. For all $r \in\left(1, \frac{4}{3}\right)$, the unique solution $\left(\underline{q}_{r}^{h}, u_{r}^{h}\right)$ of $\left(Q_{r}^{h}\right)$ is such that as $h \rightarrow 0$

$$
\begin{array}{cl}
\underline{q}_{r}^{h} \rightarrow \underline{q}_{r} & \text { weakly in }\left[L^{r}(\Omega)\right]^{n} \\
\underline{\nabla} \cdot \underline{q}_{r}^{h} \rightarrow \underline{\nabla} \cdot \underline{q}_{r} & \text { weakly in } L^{r}(\Omega) \tag{3.21b}
\end{array}
$$

where $\left(\underline{q}_{r}, u_{r}\right)$ is the unique solution of $\left(Q_{r}\right)$.
Proof. The results (3.21a-c) follow immediately for a subsequence $\left\{\left(\underline{q}_{r}^{h_{j}}, u_{r}^{h_{j}}\right)\right\}_{h_{j}>0}$ of $\left\{\left(\underline{q}_{r}^{h}, u_{r}^{h}\right)\right\}_{h>0}$ from (3.16).
For any $\eta \in L^{p}(\Omega)$, we choose $\eta^{h}=P^{h} \eta \in S^{h}$ in the $h_{j}$ version of (3.5a). Noting (3.21b), and as $P^{h} \eta \rightarrow \eta$ strongly in $L^{p}(\Omega)$ as $h \rightarrow 0$, we obtain the desired result (2.2a).

For any $\underline{\xi} \in\left[C_{0}^{\infty}(\Omega)\right]^{n}$, we choose $\underline{v}^{h}=\underline{I}^{h_{j}} \underline{\xi}-\underline{q}_{r}^{h_{j}}$ in $h_{j}$ version of (3.5b). Hence we have, on noting (2.1), that

$$
\begin{align*}
\left(u_{r}^{h_{j}}, \underline{\nabla} \cdot\left(\underline{q}_{r}^{h_{j}}-\underline{I}^{h_{j}} \underline{\xi}\right)\right) & =\left(k\left|\underline{q}_{r}^{h_{j}}\right|^{r-2} \underline{q}_{r}^{h_{j}}, \underline{q}_{r}^{h_{j}}-\underline{I}^{h_{j}} \underline{\xi}\right)^{h_{j}} \\
& \geq \frac{1}{r}\left(k\left|\underline{q}_{r}^{h_{j}}\right|^{r}, 1\right)^{h_{j}}-\frac{1}{r}\left(k\left|\underline{h}^{h_{j}} \underline{\xi}\right|^{r}, 1\right)^{h_{j}} \tag{3.22}
\end{align*}
$$

For all $\underline{\xi} \in\left[C_{0}^{\infty}(\Omega)\right]^{n}$, we have from (3.11) that

$$
\begin{equation*}
\left(k\left|\underline{\mid}^{h_{j}} \underline{\xi}\right|^{r}, 1\right)^{h_{j}} \rightarrow\left(k \mid \underline{\xi}^{r}, 1\right) \quad \text { as } h_{j} \rightarrow 0 \tag{3.23}
\end{equation*}
$$

and it follows from (3.9) and (3.21c) that

$$
\begin{equation*}
\left(u_{r}^{h_{j}}, \underline{\nabla} \cdot\left(\underline{( }^{h_{j}} \underline{\underline{\xi}}\right)\right)=\left(u_{r}^{h_{j}}, \underline{\nabla} \cdot \underline{\xi}\right) \rightarrow\left(u_{r}, \underline{\nabla} \cdot \underline{\xi}\right) \quad \text { as } h_{j} \rightarrow 0 \tag{3.24}
\end{equation*}
$$

Next we note from (3.5a), (3.21c) and (2.2a) that

$$
\begin{equation*}
\left(u_{r}^{h_{j}}, \underline{\nabla} \cdot \underline{q}_{r}^{h_{j}}\right)=\left(u_{r}^{h_{j}}, f_{r}\right) \rightarrow\left(u_{r}, f_{r}\right)=\left(u_{r}, \underline{\nabla} \cdot \underline{q}_{r}\right) \quad \text { as } h_{j} \rightarrow 0 \tag{3.25}
\end{equation*}
$$

Finally, we note from (3.4), (2.1) and (3.21a) that

$$
\begin{equation*}
\liminf _{h_{j} \rightarrow 0}\left(k\left|\underline{q}_{r}^{h_{j}}\right|^{r}, 1\right)^{h_{j}} \geq\left(k\left|\underline{q}_{r}\right|^{r}, 1\right) \tag{3.26}
\end{equation*}
$$

Combining (3.22)-(3.26), and noting the density of $\left[C_{0}^{\infty}(\Omega)\right]^{n}$ in $\underline{V}_{0}^{r}(\Omega)$, yields that

$$
\begin{equation*}
\frac{1}{r}\left(k,|\underline{\xi}|^{r}-\left|\underline{q}_{r}\right|^{r}\right) \geq\left(u_{r}, \underline{\nabla} \cdot\left(\underline{\xi}-\underline{q}_{r}\right)\right) \quad \forall \underline{\xi} \in \underline{V}_{0}^{r}(\Omega) \tag{3.27}
\end{equation*}
$$

Therefore choosing $\underline{\xi}=\underline{q}_{r} \pm \varepsilon \underline{v}$ with $\underline{v} \in \underline{V}_{0}^{r}(\Omega)$ in (3.27), and noting (2.1), yields the desired result (2.2b), on letting $\varepsilon \rightarrow 0$. Hence we have that $\left(\underline{q}_{r}, u_{r}\right) \in \underline{V}_{0}^{r}(\Omega) \times L_{I}^{p}(\Omega)$ solves $\left(\mathrm{Q}_{r}\right)$. As the solution $\left(\underline{q}_{r}, u_{r}\right)$ of $\left(\mathrm{Q}_{r}\right)$ is unique, the whole sequence $\left\{\left(\underline{q}_{r}^{h}, u_{r}^{h}\right)\right\}_{h>0}$ converges in (3.21a-c).
Remark 3.4. On combining Theorems 2.3 and 3.3 we have convergence of a subsequence of $\left\{\left(\underline{q}_{r}^{h}, u_{r}^{h}\right)\right\}_{(r>1, h>0)}$ as $r_{j} \rightarrow 1$ and $h_{j} \rightarrow 0$ to $(\underline{q}, u) \in \underline{V}_{0}^{\mathcal{M}}(\Omega) \times C_{I}(\bar{\Omega})$, which is a solution of $(\mathrm{Q})$.

## 4. Numerical algorithm and mesh adaption

The nonlinear algebraic system (3.5a,b) was solved iteratively using the following successive over-relaxation algorithm. Given an initial approximation $\underline{Q}^{0}$ to $\underline{q}_{r}^{h}$, for $j \geq 0$ find $\underline{Q}^{j+\frac{1}{2}} \in \underline{V}_{0}^{h}$ and $U^{j+1} \in S_{I}^{h}$ such that

$$
\begin{align*}
\left(\underline{\nabla} \cdot \underline{Q}^{j+\frac{1}{2}}, \eta^{h}\right) & =\left(f_{r}, \eta^{h}\right) & & \forall \eta^{h} \in S^{h},  \tag{4.1a}\\
\left(k\left[\left|\underline{Q}^{j}\right|^{r-2} \underline{Q}^{j}+\left|\underline{Q}^{j}\right|_{\varepsilon}^{r-2}\left(\underline{Q}^{j+\frac{1}{2}}-\underline{Q}^{j}\right)\right], \underline{v}^{h}\right)^{h} & =\left(U^{j+1}, \underline{\nabla} \cdot \underline{v}^{h}\right) & & \forall \underline{v}^{h} \in \underline{V}_{0}^{h} ; \tag{4.1b}
\end{align*}
$$

where $|\underline{a}|_{\varepsilon}:=\left(|\underline{a}|^{2}+\varepsilon^{2}\right)^{\frac{1}{2}}$ for a chosen positive $\varepsilon \ll 1$. This regularization ensures that all terms in the linear system (4.1a,b) are well-defined. Uniqueness, and hence existence, follows on recalling (3.3) and (3.13). Then set

$$
\begin{equation*}
\underline{Q}^{j+1}=\gamma \underline{Q}^{j+\frac{1}{2}}+(1-\gamma) \underline{Q}^{j} \tag{4.2}
\end{equation*}
$$

where $\gamma$ is an over-relaxation parameter. When successive iterations of $\underline{Q}^{j}$ converge to a prescribed tolerance, set $\underline{q}_{r}^{h}=\underline{Q}^{j+1}$ and $u_{r}^{h}=U^{j+1}$.

In all the examples below, we chose $r=1+10^{-7}, \varepsilon=10^{-7}, \gamma=1.3$ or 1.5 , and the stopping criteria to be $\left|\underline{Q}^{j+1}-\underline{Q}^{j}\right|_{0, \infty, \Omega} /\left|\underline{Q}^{j+1}\right|_{0, \infty, \Omega} \leq 10^{-3}$. Although this is a relatively crude criterion, for the examples considered in Section 5 it was more than sufficient to make the results of further iterations graphically indistinguishable in the presented plots.

The number of iterations was not sensitive to the values of $r$ and $\varepsilon$. Typically, more iterations were required on finer meshes, although this dependence was weak. Using adaptive mesh refinement, and computing the initial approximation by interpolating the solution from the crude to the fine mesh (we took $\underline{Q}^{0}=0$ for the initial mesh); the above algorithm required between 20 to 50 iterations for the initial meshes with about 1000 triangles, and between 20 to 80 iterations for meshes with more than 50000 triangles.

Adaptive mesh refinement is important because the optimal flux $q$ is often singular, e.g. if the measures $f^{ \pm}$ are not absolutely continuous with respect to the Lebesgue measure. The transport flux may also be singular if the domain $\Omega$ is non-convex, see [7,19,25], or is inhomogeneous, see [7,25]. The following iterative refinement/coarsening procedure allowed us to significantly decrease "smearing" of the numerical solution near singularities, see [7]. To simplify the notation, below we often omit the indices $h$ and $r$ in $\left(\underline{q}_{r}^{h}, u_{r}^{h}\right)$ and $\mathcal{T}^{h}$.

Let $\mathcal{T}$ be the initial (crude) regular partitioning of $\Omega$ into disjoint open simplices, which is used to compute our starting approximation $\left(\underline{q}^{(0)}, u^{(0)}\right)$ to the final approximate solution. On the $(m+1)$-st iteration of the adaption procedure, the same initial mesh, $\mathcal{T}$, is refined sequentially $(m+1)$ times using the approximate solution $\left(\underline{q}^{(m)}, u^{(m)}\right)$ from the previous iteration and a chosen refinement criterion. Hence a sequence of partitionings, $\left\{\mathcal{T}^{(m+1,0)}:=\mathcal{T}, \mathcal{T}^{(m+1,1)}, \ldots, \mathcal{T}^{(m+1, m+1)}\right\}$ is generated. The finest of them, $\mathcal{T}^{(m+1, m+1)}$, is then used to compute the next approximation $\left(\underline{q}^{(m+1)}, u^{(m+1)}\right)$.

Typically, a fine finite element mesh should be used in those regions, where the solution is changing rapidly. A natural refinement strategy is to find those simplex faces across which the flux jump, $\left[\underline{q}_{r}^{h}\right]$, is large; and to mark the adjacent simplices for refinement. Using the iterative re-meshing procedure described above we can, on the $(m+1)$-st iteration, compile a list of, say, centres of $\mathcal{T}^{(m, m)}$ simplex faces with large $\underline{q}^{(m)}$ jumps. We then obtain the partitionings $\mathcal{T}^{(m+1, i+1)}, i=0 \rightarrow m$, by marking for refinement those simplices of $\mathcal{T}^{(m+1, i)}$ whose closures contain at least one point from the list. This was our main refinement strategy, although we modified it in some of the examples below.

## 5. Numerical experiments

To test our numerical scheme we solved several two-dimensional MK problems with $\Omega$ being the unit square. Three iterations of mesh adaption were usually made. The Matlab Partial Differential Equation Toolbox was used for the generation and the refinement of finite element meshes. We refer to [6] for the Matlab realization of the lowest order Raviart-Thomas element in $\mathbb{R}^{2}$.

## Example 1. Inhomogeneous domain

Let $f^{+}$and $f^{-}$be uniformly distributed in the left and right ellipses $\Omega^{ \pm}$, respectively, (see Fig. 1) and have the same total masses. We define $\Omega_{1}$ to be the slanted ellipse in the middle and $\Omega_{2}=\Omega \backslash \bar{\Omega}_{1}$. Numerical results are shown for $\left.k\right|_{\Omega_{2}} \equiv 1$ and $\left.k\right|_{\Omega_{1}} \equiv k_{1}$ with $k_{1}=0.01,1,3$, and $\infty$ (from the top row to the bottom row, respectively). The case $k_{1}=\infty$ corresponds to $\Omega_{1}$ being an obstacle, and was solved in $\Omega=\Omega_{2}$ with the no-flux boundary condition, $\underline{q} \cdot \underline{\nu}=0$, on both components of its boundary.

The initial mesh $\overline{\mathcal{T}}$ contained less than $10^{3}$ triangles and was refined adaptively according to the procedure described above. More precisely, let $\mathcal{E}:=\mathcal{E}\left(\mathcal{T}^{(m, m)}\right)$ be the set of all internal edges of the mesh $\mathcal{T}^{(m, m)}$ used to compute $\left(\underline{q}^{(m)}, u^{(m)}\right)$. The absolute value of the flux jump, $\delta_{e}=\left|\left[\underline{q}^{(m)}\right]_{e}\right|$, is then estimated for each edge $e \in \mathcal{E}$ as the average of the $\left|\left[\underline{\underline{q}}^{(m)}\right]\right|$ values at the edge ends. The minimal set of edges, $\mathcal{E}_{\kappa} \subset \mathcal{E}$, satisfying, for a given threshold $\kappa$,

$$
\begin{equation*}
\sum_{e \in \mathcal{E}_{\kappa}} \delta_{e} \geq \kappa \sum_{e \in \mathcal{E}} \delta_{e} \tag{5.1}
\end{equation*}
$$

is determined. The midpoints of the edges from this set are then used to construct $\mathcal{T}^{(m+1, m+1)}$ via $m+1$ consequent refinements of the initial partitioning $\mathcal{T}$. The value of threshold parameter used in this example was $\kappa=0.95$. This led to the meshes, after three refinement iterations, containing approximately 25000 triangles; which corresponds to approximately 60000 degrees of freedom (the number of triangles plus the number of internal edges) to determine the solution $\left(\underline{q}_{r}^{h}, u_{r}^{h}\right)$ on such a mesh.

We note these meshes are fitted to the ellipses in the domain $\Omega$; e.g. the ellipses $\Omega^{ \pm}$are approximated by polygons $\Omega^{ \pm, h}$, which are a union of triangles. Therefore $f^{ \pm}$have to be modified slightly in order to maintain a balanced problem under this perturbation of domain, i.e. $f^{ \pm, h}=\left|\Omega^{ \pm}\right| f^{ \pm} /\left|\Omega^{ \pm, h}\right|$.

## Example 2. Discrete case

Efficient numerical algorithms have been derived for discrete MK problems, where $f^{+}(\underline{x})=\sum_{i=1}^{l} a_{i} \delta\left(\underline{x}-\underline{y}_{i}\right)$, $f^{-}(\underline{x})=\sum_{j=1}^{n} b_{j} \delta\left(\underline{x}-\underline{x}_{j}\right)$ with $\sum_{i=1}^{l} a_{i}=\sum_{j=1}^{n} b_{j}$. A continuous approach can compete with these schemes only if there are so many discrete sources that it is preferable to approximate them by regularized sources, e.g. through convolution. This is certainly not the case for the two simple examples presented in Figure 2. However, these discrete problems can be considered a challenging test for our numerical scheme, because the optimal flux is concentrated upon a set of dimension one and is, therefore, especially singular.

In both examples we set $a_{i}=1 / l, b_{j}=1 / n$ for all $i, j$; the points $\underline{y}_{i}, \underline{x}_{j}$ were randomly chosen in $\Omega=[0,1]^{2}$; and we took $k \equiv 1$. The optimal transport plans (see the upper two plots in Fig. 2), calculated by means of a linear programming subroutine, show much simpler structure for the first example $(l=9, n=6)$, where each point source is connected by transport rays to exactly two sinks, than for the second example $(l=11, n=7)$. The total cost of transportation in these examples is 0.2148 and 0.2615 , respectively.

The refining/coarsening algorithm based on the criterion (5.1) was able to produce the finite element meshes refined mainly in the vicinity of transport rays; however, for these problems slightly better meshes were produced with another criterion based on the flux itself and not upon its edge jumps. Let $\underline{\phi}_{e}^{(m)}(\underline{x})$ be the basis vector function associated with the edge $e \in \mathcal{E}:=\mathcal{E}\left(\mathcal{T}^{(m, m)}\right)$ in the Raviart-Thomas finite element space, see e.g. [6], and $\underline{q}^{(m)}=\sum_{e \in \mathcal{E}} q_{e}^{(m)} \underline{\phi}_{e}^{(m)}$. For these problems $\mathcal{T}^{(m+1, m+1)}$ was built, by means of the sequence of refinements of $\mathcal{T}$ described above, using the midpoints of the edges from the set $\mathcal{E}_{\kappa}=\left\{e \in \mathcal{E}:\left|q_{e}^{(m)}\right| \geq \kappa \max \left(\left|q_{e}^{(m)}\right|\right)\right\}$ with $\kappa=0.002$. With probability one the points $\left\{\underline{y}_{i}\right\}_{i=1}^{l}$ and $\left\{\underline{x}_{j}\right\}_{j=1}^{n}$ lied in the interior of a triangle, so no smoothing of these discrete sources was required.

The two lower plots in Figure 2 are the finite element meshes $\mathcal{T}^{(m, m)}$ and the contour plots of $\left|\underline{q}^{(m)}\right|$ after several iterations of mesh adaption. Here $m=3$ and $m=4$ for the first and second examples, respectively; because better resolution is required to approximate the more complicated transport plan structure in the second example. The meshes contain 16445 and 96466 triangles, respectively. Even using almost $10^{5}$ triangles, we could obtain only a similar, but not exactly the optimal, connection of sources and sinks in the second example.


Figure 1. Optimal transport in an inhomogeneous domain. $f^{+}$and $f^{-}$are uniform in their supports (left and right ellipses, respectively), $k=1$ outside the slanted ellipse between them. Inside this ellipse $k=0.01,1,3, \infty$ (from the top row to the bottom row). Shown for each value: left - the finite element mesh obtained after three adaptions, middle - the transport flux $\underline{q}_{r}^{h}$ and $u_{r}^{h}$ levels, right - contour plot of $\left|\underline{q}_{r}^{h}\right|$.


Figure 2. Two discrete problems: $f^{+}=\sum_{i=1}^{l} \frac{1}{l} \delta\left(\underline{x}-\underline{y}_{i}\right), f^{-}=\sum_{i=1}^{n} \frac{1}{n} \delta\left(\underline{x}-\underline{x}_{i}\right)$ with randomly chosen $\underline{y}_{i}$ (red stars) and $\underline{x}_{i}$ (blue dots). Left column: $l=9, n=6$; right column: $l=11, n=7$. Upper row: exact optimal transport rays found by linear programming; the total transportation costs are 0.2148 and 0.2615 , respectively. Lower row: same problems solved by our continuous method. To indicate the main transport rays, level contours of $\left|\underline{q}_{r}^{h}\right|$ are shown in each plot. The estimated transportation costs are 0.2149 and 0.2624 , respectively.

Nevertheless, the plans obtained are very close to optimal in both cases; and the calculated transportation cost,

$$
\int_{\bar{\Omega}}|\underline{q}| \approx \sum_{\sigma \in \mathcal{T}^{(m, m)}}|\sigma|\left|\underline{q}^{(m)}\left(\underline{o}_{\sigma}\right)\right|,
$$

where $\underline{o}_{\sigma}$ is the center of element $\sigma$, is 0.2149 for the first, and 0.2624 for the second example. These are surprisingly close to the true costs found by linear programming (relative errors are approximately $0.05 \%$ and $0.3 \%$ ).

## Example 3. Optimal couplings

Suppose $k \equiv 1$, the measure $f^{+}=1 /|\Omega|$ is uniform and $f^{-}(\underline{x})=\sum_{i=1}^{n} \alpha_{i} \delta\left(\underline{x}-\underline{x}_{i}\right)$, with $\underline{x}_{i} \in \Omega$ and $\sum_{i=1}^{n} \alpha_{i}=1$, is a discrete distribution. For each sink $\underline{x}_{i}$, the solution of the MK problem determines a subset $A_{i} \subset \Omega$ "coupled" to $\underline{x}_{i}$, i.e. the set from which $f^{+}$should be transported into $\underline{x}_{i}$. A general characterization of such solutions given by Rüschendorf and Uckelmann, see [26], for the linear cost function reads:

Let $F(\underline{x})=\max _{i=1 \rightarrow n}\left\{-\left|\underline{x}-\underline{x}_{i}\right|+a_{i}\right\}$ be a function determined by $n$ parameters $a_{i}$ and $A_{i}=\{\underline{x} \in \Omega$ : $\left.F(\underline{x})=-\left|\underline{x}-\underline{x}_{i}\right|+a_{i}\right\}$. The coupling $\left\{A_{i}\right\}$ is optimal if and only if the parameters $a_{i}$ are such that $\left|A_{i}\right|=\alpha_{i}$ for all $i$.

Using this characterization, it is easy to solve the inverse MK problem: for given values of the parameters $a_{i}$, determine the sets $A_{i}$ and calculate the coefficients $\alpha_{i}=\left|A_{i}\right|$ such that $\left\{A_{i}\right\}$ is the optimal coupling for $f^{-}(\underline{x})=\sum_{i=1}^{n} \alpha_{i} \delta\left(\underline{x}-\underline{x}_{i}\right)$.

If all $a_{i}=0$, we obtain a partitioning of $\Omega$ for which the points are coupled to the closest sinks; and this determines the sets $A_{i}$. In the case of closest sinks, these sets are polygons. Estimating the areas of each $A_{i}$ numerically and imposing that $\sum_{i=1}^{n}\left|A_{i}\right|=1$, we then solved the direct MK problem with the discrete measure $f^{-}(\underline{x})=\sum_{i=1}^{n}\left|A_{i}\right| \delta\left(\underline{x}-\underline{x}_{i}\right)$ to compare the approximate and exact couplings. It is convenient to present the couplings by plotting the normalized "directions" of transportation, $\underline{q}_{r}^{h} /\left(\left|\underline{q}_{r}^{h}\right|+\varepsilon\left|\underline{q}_{r}^{h}\right|_{0, \infty, \Omega}\right)$, at equally spaced points in $\Omega$, see Figure 3, where we chose $\varepsilon=0.35 \times 10^{-3}$. The exact boundaries of the sets $A_{i}$ are also shown.

Note that the flux $q$ is zero on the boundary of each set $A_{i}$. Since these boundaries are the main interest in the coupling problems, we used the singularity-oriented refinement criterion (5.1) with a lower threshold, $\kappa=0.5$; but, in order to refine the mesh near these boundaries, we added to the list of refinement points also those midpoints of the edges $e \in \mathcal{T}^{(m, m)}$ with $\left|q_{e}^{(m)}\right|<0.0025 \max \left(\left|q_{e}^{(m)}\right|\right)$. After three refinements the mesh contained about 50000 triangles. Once again, we did not smooth the discrete sinks.

We solved also the coupling problem for $f^{-}$with the sinks $\left\{\underline{x}_{i}\right\}_{i=1}^{n}$ placed at the same points as in the previous example, see Figure 3, but now with $\alpha_{i}=1 / n$ for all $i$. Now the sets $A_{i}$ should have equal areas and are not necessarily polygons. The exact boundaries are now more difficult to find, because determining the coefficients $a_{i}$ in the solution characterization is not a simple task. In this example the boundaries were approximated by the contours $\left|\underline{q}_{r}^{h}\right|=\varepsilon\left|\underline{q}_{r}^{h}\right|_{0, \infty, \Omega}$. Similarly to the above, a mesh with about 50000 triangles was produced after three refinements.

## Appendix A: Duality of the potential and flux formulations

Let $f \in \mathcal{M}(\bar{\Omega})$ and $k: \bar{\Omega} \rightarrow \mathbb{R}_{>0}$ be a continuous function. Instead of the problem

$$
\begin{equation*}
\max \left\{\langle f, u\rangle_{C(\bar{\Omega})}: u \in \operatorname{Lip}_{k}(\bar{\Omega})\right\} \tag{0}
\end{equation*}
$$

where $\operatorname{Lip}_{k}(\bar{\Omega})$ is defined as in (1.7), for which a solution exists; we consider the minimization problem

$$
\begin{equation*}
\inf \{F(u)+G(\Lambda u): u \in V\}, \tag{P}
\end{equation*}
$$

where $\Lambda u:=\underline{\nabla} u$ is a linear operator from $V:=C^{1}(\bar{\Omega})$ to $\underline{Y}:=[C(\bar{\Omega})]^{n}$ and $F: V \rightarrow \mathbb{R}$ and $G: \underline{Y} \rightarrow \overline{\mathbb{R}}$ are convex functions defined as follows:

$$
F(u):=-\langle f, u\rangle_{C(\bar{\Omega})} \quad \text { and } \quad G(\underline{v}):= \begin{cases}0 & \text { if }|\underline{v}(\underline{x})| \leq k(\underline{x}) \quad \forall \underline{x} \in \bar{\Omega}, \\ \infty & \text { otherwise. }\end{cases}
$$

Although $(\mathcal{P})$ usually has no solutions, one can show that if $k \in C(\bar{\Omega})$ then $\inf (\mathcal{P})=-\max \left(\mathcal{P}_{0}\right)$. Clearly in general one has that $\inf (\mathcal{P}) \geq-\max \left(\mathcal{P}_{0}\right)$. The stated equality follows for $k \in C(\bar{\Omega})$, as in this case $\operatorname{Lip}_{k}(\bar{\Omega}) \equiv\left\{u \in W^{1, \infty}(\Omega):|\underline{\nabla} u(\underline{x})| \leq k(\underline{x})\right.$ a.e. $\left.\underline{x} \in \Omega\right\}$; and given any $u \in \operatorname{Lip}_{k}(\bar{\Omega})$, one can find a sequence $\left\{u_{n}\right\}_{n=1}^{\infty} \in V \cap \operatorname{Lip}_{k}(\bar{\Omega})$ such that $\lim _{n \rightarrow \infty}\left|u-u_{n}\right|_{0, \infty, \Omega}=0$, which follows from a simplified version of [12], Proposition 4.2.


Figure 3. Couplings to five point sinks (blue dots). Left: the weights of the sinks are determined from the inverse problem and are such that coupling is to the nearest sink; shown are the field of transportation directions (numerical solution) and the exact boundaries (red lines). Right: equal sinks placed at the same points, the sets coupled to each sink have the same area; the exact set boundaries are approximated by the contours $\left|\underline{q}_{r}^{h}\right|=3.5 \cdot 10^{-4}\left|\underline{q}_{r}^{h}\right|_{0, \infty, \Omega}$ ) (red lines).

For $u_{0}=0$ we obtain that $F\left(u_{0}\right)=0$ and the function $G(\underline{v})$ is continuous at $\Lambda u_{0}=0$. Hence it follows from [13], Chapter III, Theorem 4.1, that problem $(\mathcal{P})$ is stable; that is, the dual problem

$$
\begin{equation*}
\sup \left\{-F^{*}\left(\Lambda^{*} \underline{q}\right)-G^{*}(-\underline{q}): \underline{q} \in \underline{Y}^{*}\right\} \tag{*}
\end{equation*}
$$

has a solution and $\inf (\mathcal{P})=\sup \left(\mathcal{P}^{*}\right)$. Here $\underline{Y^{*}} \equiv[\mathcal{M}(\bar{\Omega})]^{n}$ is the dual of $\underline{Y}, \Lambda^{*}: \underline{Y}^{*} \rightarrow V^{*}$, the dual of $V$, is the transpose of $\Lambda: V \rightarrow \underline{Y}$ and $F^{*}: V^{*} \rightarrow \overline{\mathbb{R}}, G^{*}: \underline{Y}^{*} \rightarrow \overline{\mathbb{R}}$ are the conjugate functions of $F$ and $G$.

Hence for any $u^{*} \in V^{*}$, we have that

$$
F^{*}\left(u^{*}\right):=\sup \left\{\left\langle u^{*}, u\right\rangle_{C^{1}(\bar{\Omega})}+\langle f, u\rangle_{C(\bar{\Omega})}: u \in V\right\}= \begin{cases}0 & \text { if } u^{*}+f=0 \\ \infty & \text { otherwise }\end{cases}
$$

Therefore $-F^{*}\left(\Lambda^{*} \underline{q}\right)>-\infty \Leftrightarrow \Lambda^{*} \underline{q}+f=0$. This means that the supremum of $\left(\mathcal{P}^{*}\right)$ should be sought in the convex set $\left.\underline{Z}:=\left\{\underline{q} \in \underline{Y}^{*}: \bar{q}, \underline{\nabla} u\right\rangle_{C(\bar{\Omega})}+\langle f, u\rangle_{C(\bar{\Omega})}=0 \quad \forall u \in C^{1}(\bar{\Omega})\right\} \equiv\left\{\underline{q} \in \underline{\mathcal{V}}_{0}^{\mathcal{M}}(\Omega):\langle\underline{\nabla} \cdot \underline{q}-\right.$ $\left.f, \eta\rangle_{C(\bar{\Omega})}=0 \quad \forall \eta \in C(\bar{\Omega})\right\}$, on recalling (1.9a,b). Furthermore as $k \in C(\bar{\Omega})$ and positive,

$$
\begin{aligned}
G^{*}(\underline{q}) & :=\sup \left\{\langle\underline{q}, \underline{v}\rangle_{C(\bar{\Omega})}: \underline{v} \in \underline{Y} \text { and }|\underline{v}(\underline{x})| \leq k(\underline{x}) \quad \forall \underline{x} \in \bar{\Omega}\right\} \\
& \equiv \sup \left\{\langle\underline{q}, k \underline{v}\rangle_{C(\bar{\Omega})}: \underline{v} \in \underline{Y} \text { and }|\underline{v}(\underline{x})| \leq 1 \quad \forall \underline{x} \in \bar{\Omega}\right\} \equiv\langle | \underline{q}|, k\rangle_{C(\bar{\Omega})} .
\end{aligned}
$$

Therefore the flux formulation $\left(\mathcal{P}^{*}\right)$ can now be rewritten as $\max \left\{-\langle\mid \underline{q}, k\rangle_{C(\bar{\Omega})}: \underline{q} \in \underline{Z}\right\}$, and so if $k \in C(\bar{\Omega})$ then

$$
\min \left\{\langle | \underline{q}|, k\rangle_{C(\bar{\Omega})}: \underline{q} \in \underline{Z}\right\}=\max \left\{\langle u, f\rangle_{C(\bar{\Omega})}: u \in \operatorname{Lip}_{k}(\bar{\Omega})\right\}
$$

Finally, we note that the above equivalence has been shown in a far more general context, including $k$ being a positive lower semicontinuous function, using more sophisticated mathematical machinery; see [23].

## References

[1] L. Ambrosio, Optimal transport maps in Monge-Kantorovich problem, Proceedings of the ICM (Beijing, 2002) III. Higher Ed. Press, Beijing (2002) 131-140.
[2] L. Ambrosio, Lecture notes on optimal transport, in Mathematical Aspects of Evolving Interfaces, L. Ambrosio et al. Eds., Lect. Notes in Math. 1812 (2003) 1-52.
[3] L. Ambrosio, N. Fusco and D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems. Clarendon Press, Oxford (2000).
[4] S. Angenent, S. Haker and A. Tannenbaum, Minimizing flows for the Monge-Kantorovich problem. SIAM J. Math. Anal. 35 (2003) 61-97.
[5] G. Aronson, L.C. Evans and Y. Wu, Fast/slow diffusion and growing sandpiles. J. Diff. Eqns. 131 (1996) 304-335.
[6] C. Bahriawati and C. Carstensen, Three Matlab implementations of the lowest-order Raviart-Thomas MFEM with a posteriori error control. Comput. Methods Appl. Math. 5 (2005) 333-361.
[7] J.W. Barrett and L. Prigozhin, Dual formulations in critical state problems. Interfaces Free Boundaries 8 (2006) $347-368$.
[8] J.W. Barrett and L. Prigozhin, Partial L ${ }^{1}$ Monge-Kantorovich problem: variational formulation and numerical approximation. (Submitted).
[9] J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. Numer. Math. 84 (2000) 375-393.
[10] G. Bouchitté, G. Buttazzo and P. Seppecher, Shape optimization solutions via Monge-Kantorovich equation. C.R. Acad. Sci. Paris 324-I (1997) 1185-1191.
[11] L.A. Caffarelli and R.J. McCann, Free boundaries in optimal transport and Monge-Ampère obstacle problems. Ann. Math. (to appear).
[12] R. De Arcangelis and E. Zappale, The relaxation of some classes of variational integrals with pointwise continuous-type gradient constraints. Appl. Math. Optim. 51 (2005) 251-277.
[13] I. Ekeland and R. Temam, Convex Analysis and Variational Problems. North-Holland, Amsterdam (1976).
[14] L.C. Evans, Weak Convergence Methods for Nonlinear Partial Differential Equations, C.B.M.S. 74. AMS, Providence RI (1990).
[15] L.C. Evans, Partial differential equations and Monge-Kantorovich mass transfer, Current Developments in Mathematics. Int. Press, Boston (1997) 65-126.
[16] L.C. Evans and W. Gangbo, Differential Equations Methods for the Monge-Kantorovich Mass Transfer Problem. Mem. Amer. Math. Soc. 137 (1999).
[17] M. Farhloul, A mixed finite element method for a nonlinear Dirichlet problem. IMA J. Numer. Anal. 18 (1998) $121-132$.
[18] M. Farhloul and H. Manouzi, On a mixed finite element method for the p-Laplacian. Can. Appl. Math. Q. 8 (2000) 67-78.
[19] M. Feldman, Growth of a sandpile around an obstacle, in Monge Ampere Equation: Applications to Geometry and Optimization, L.A Caffarelli and M. Milman Eds., Contemp. Math. 226, AMS, Providence (1999) 55-78.
[20] G.B. Folland, Real Analysis: Modern Techniques and their Applications (Second Edition). Wiley-Interscience, New York (1984).
[21] V. Girault and P.-A. Raviart, Finite Element Methods for Navier-Stokes Equations. Springer-Verlag, Berlin (1986).
[22] P. Grisvard, Elliptic Problems in Nonsmooth Domains. Pitman, Massachusetts (1985).
[23] A. Pratelli, Equivalence between some definitions for the optimal mass transport problem and for transport density on manifolds. Ann. Mat. Pura Appl. 184 (2005) 215-238.
[24] L. Prigozhin, Variational model for sandpile growth. Eur. J. Appl. Math. 7 (1996) 225-235.
[25] L. Prigozhin, Solutions to Monge-Kantorovich equations as stationary points of a dynamical system. arXiv:math.0C/0507330, http://xxx.tau.ac.il/abs/math.0C/ 0507330 (2005).
[26] L. Rüschendorf and L. Uckelmann, Numerical and analytical results for the transportation problem of Monge-Kantorovich. Metrika 51 (2000) 245-258.
[27] G. Strang, $L^{1}$ and $L^{\infty}$ approximation of vector fields in the plane. Lecture Notes in Num. Appl. Anal. 5 (1982) $273-288$.
[28] C. Villani, Topics in Optimal Transportation, Graduate Studies in Mathematics 58. AMS, Providence RI (2003).


[^0]:    Keywords and phrases. Monge-Kantorovich problem, optimal transportation, mixed methods, finite elements, existence, convergence analysis.
    ${ }^{1}$ Department of Mathematics, Imperial College London, London SW7 2AZ, UK. jwb@ic.ac.uk
    ${ }^{2}$ Department of Solar Energy and Environmental Physics, Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Sede Boqer Campus, 84990, Israel.

