A definition of the background state of the atmosphere using optimal transport

Charlie Egan[*], John Methven[†], David P. Bourne[‡], Mike J. P. Cullen[§]

Abstract

The dynamics of atmospheric disturbances are often described in terms of displacements of air parcels relative to their locations in a notional background state. Modified Lagrangian Mean (MLM) states have been proposed by M. E. McIntyre using the Lagrangian conserved variables potential vorticity and potential temperature to label air masses in the instantaneous 3-D atmospheric state, and then rearranging air masses sandwiched between potential vorticity contours in potential temperature layers so that they are zonally symmetric. Methven and Berrisford used conservation of Kelvin's circulation to define the distribution of MLM state mass in zonal angular momentum (z) and potential temperature (θ) coordinates given 3-D global data. We prove that for any such mass distribution in (z, θ) there exists a mapping to a unique mass distribution in geophysical coordinates (latitude and pressure) which minimises energy and completely defines the MLM state in a single hemisphere. Since the state is an energy minimum, it is suitable for quantification of finite amplitude disturbances and examining atmospheric instability. This state is obtained by solving a free surface problem, which we frame as the minimisation of an optimal transport cost over a class of source measures. The solution consists of a source measure, encoding surface pressure, and an optimal transport map, connecting the distribution of mass in geophysical coordinates to the known distribution of mass in (z, θ). We show that this problem reduces to an optimal transport problem with a known source measure, which has a numerically feasible discretisation, and we present a numerical solution using atmospheric reanalysis data. Additionally, our results hold for a large class of cost functions, and generalise analogous results on free surface variants of the semi-geostrophic equations.

1 Introduction

The dynamics of atmospheric disturbances and geophysical fluid instability are often described in terms of displacements of air parcels relative to a notional background state. Such disturbances include large amplitude waves, vortices, and other weather systems. In [23], Modified Lagrangian Mean (MLM) states were proposed as a suitable choice of background state by using the conserved variables potential vorticity (q) and potential temperature (θ) to label air parcels, thus avoiding the need to calculate 3-D trajectories explicitly to measure disturbance activity. In a chaotic dynamical system, neighbouring trajectories separate exponentially on average, rendering them sensitive to initial conditions and increasingly complex and infeasible to calculate, even numerically, while q

[*]Department of Meteorology, University of Reading, UK; c.egan@reading.ac.uk

[†]Department of Meteorology, University of Reading, UK; j.methven@reading.ac.uk

[‡]David P. Bourne, Maxwell Institute for Mathematical Sciences and Department of Mathematics, Heriot-Watt University, Edinburgh, UK; d.bourne@hw.ac.uk

[§]Met Office, Exeter, UK (retired)

and θ can typically be calculated from instantaneous data. The MLM state is then defined as an adiabatic rearrangement of the 3-D state so that potential vorticity contours on each θ-isosurface are zonally symmetric, reducing dimensions to a latitudinal-vertical cross-section. The authors of [27] obtained a zonally-symmetric MLM state numerically from global atmospheric (3D) data by calculating mass and circulation integrals as functionals of q and θ. Since these integrals are conserved for adiabatic, frictionless flows, the MLM state is obtained as a zonally-symmetric flow with the same integrals as the 3-D state which is sufficient to specify the distribution of mass in the conserved variable coordinates defined by zonal angular momentum (z) and θ. In the current work, we prove the existence, uniqueness and stability (with respect to input data) of the MLM state in geophysical coordinates (latitude and pressure) with the minimum energy, given this distribution of mass in conserved variable coordinates (z, θ). In this way, the MLM state is uniquely defined from the 3-D atmospheric state, as the energy-minimising zonally-symmetric adiabatic rearrangement of fluid parcels. We achieve this by adapting the framework set out in [7, Section 4.5] without restricting our analysis to incompressible flows. This provides a rigorous principle by which to select a unique background state of the atmosphere suitable for the examination of fluid dynamical instability.

We phrase the problem of finding energy-minimising MLM states as the minimisation of an optimal transport cost over a suitable space of source measures (μ) for a given target measure (ν). In the language of optimal transport, the source space corresponds to geophysical coordinates (latitude and pressure) and the target space is defined by the conserved variable coordinates (z, θ). In physical terms, the optimal transport solution maps the mass from the target space (distribution taken as given) into the source space, while also minimising a cost which is shown to be equal to the energy of the state.

We now describe the general mathematical problem that we study, our main results, and their interpretation in the context of finding energy minimising MLM states. For more detail on this interpretation see Section 1.3 and Figure 1. Let $d \in \mathbb{N}$, $d > 1$, and let $B \subset \mathbb{R}^{d-1}$ be compact. Define the source space $X \coloneqq B \times [0, +\infty)$. For a probability density $\bar{p}$ defined on B, define the set

$$X_{\bar{p}} \coloneqq \{(s, p) \in X \,:\, 0 \leq p \leq \bar{p}(s)\}.$$

That is, $X_{\bar{p}}$ is the intersection of the subgraph of $\bar{p}$ with the upper half-space. Let $\mu_{\bar{p}}$ be the restriction of the Lebesgue measure to $X_{\bar{p}}$, and let ν be a probability measure on $\mathbb{R}^d$ with support contained in a compact set $Y \subset \mathbb{R}^d$. For a cost function $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ denote by $\mathcal{T}_c(\mu_{\bar{p}}, \nu)$ the optimal transport cost from $\mu_{\bar{p}}$ to ν for the cost c. We study the minimisation problem

$$\operatorname*{arg\,min}_{\bar{p} \in \mathcal{P}_{\mathrm{ac}}(B)} \mathcal{T}_c(\mu_{\bar{p}}, \nu). \tag{1}$$

Here, $\mathcal{P}_{\mathrm{ac}}(B)$ is the space of probability measures on B that are absolutely continuous with respect to the $(d-1)$-dimensional Lebesgue measure, and we conflate a measure in $\mathcal{P}_{\mathrm{ac}}(B)$ with its density.

In the setting of MLM states, $d = 2$, the coordinates s and p of the source space X represent sine-of-latitude and pressure, respectively, and the set B is a compact interval. The measure ν represents the mass distribution over the space of zonal angular momentum and potential temperature of the atmosphere at an instant in time. An MLM state is determined by a latitude-dependent probability density $\bar{p}$, representing zonally-symmetric surface pressure, and an admissible transport map from $\mu_{\bar{p}}$ to ν, representing an zonally-symmetric mass-preserving rearrangement of the zonal angular momentum and potential temperature of the full state. From a physical perspective, the optimal transport map relates air parcels and their conserved properties, given by the parcel locations in (z, θ), to their distribution in geophysical coordinates. The cost function c defined by (13) represents

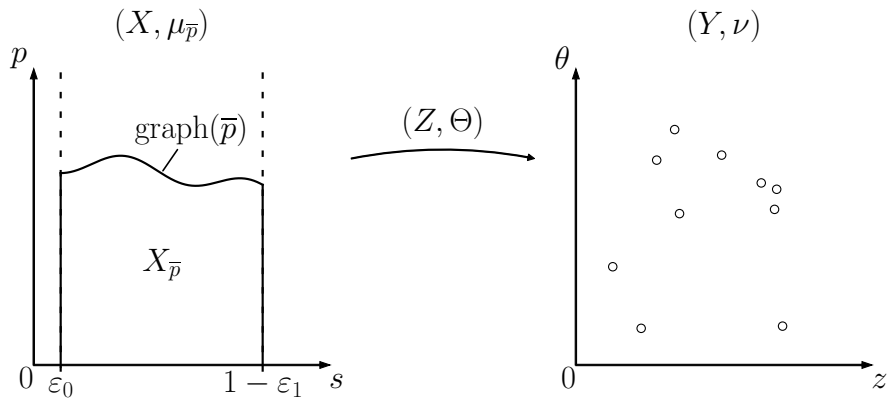


Figure 1: Schematic diagram of an MLM state at an instant in time. The measure ν represents the distribution of mass in the space of zonal angular momentum z and potential temperature θ. It is assumed to have support contained in a compact set $Y \subset (0, +\infty)^2$. The source space $X = [\varepsilon_0, 1-\varepsilon_1] \times [0, +\infty)$ has coordinates $s = \sin(\phi)$, where ϕ is latitude, and pressure p. The constants ε_0, $\varepsilon_1 > 0$ exclude the pole and the equator. The surface pressure $\overline{p}$ is a function of s and is an unknown of the problem. It determines the measure $\mu_{\bar{p}}$ as the restriction of the Lebesgue measure to the set $X_{\overline{p}}$. Zonal angular momentum Z and potential temperature Θ are unknown scalar functions of (s, p). The triple $(\bar{p}, Z, \Theta)$ is an MLM state for ν if (Z, Θ) is an admissible transport map from $\mu_{\bar{p}}$ to ν.

the sum of kinetic energy and enthalpy, and the corresponding transport cost is the energy of the MLM state. Assumptions 2.1 and 2.6 are both satisfied by c, so our results establish a rigorous principle uniquely selecting MLM states based on energy minimisation.

Our main result is that for a large class of cost functions c (Assumption 2.1), the minimisation problem (1) has a unique solution $\bar{p}$, which has continuous density (Theorem 3.10), and can be recovered by solving an optimal transport problem with a known source measure, an adjusted target measure and an extended cost function (see Theorem 5.1 and Figure 2). We use this to derive a numerically tractable discretisation of (1); see Corollary 5.3. Moreover, for target measures ν with common compact support, we show that the set of all such solutions is uniformly bounded and equicontinous, and that solutions are stable with respect to ν (Theorem 4.1). Our existence, uniqueness and stability results generalise analogous results obtained for cost functions defining free surface variants of the semi-geostrophic equations [6, 8, 10] and a model for axisymmetric vortices [13], which rely on the use of convex potentials and their relation to optimal transport with the quadratic cost.

1.1 Motivation and physical problem statement

Considering atmospheric disturbances as perturbations from a background state, wave activity conservation laws, which describe how disturbance amplitude can propagate from one location to another, have been obtained for arbitrary large amplitude disturbances provided that the flow is adiabatic and frictionless and the background state is itself a solution of the equations of motion [24]. The laws are derived using Lagrangian conservation properties of fluid dynamics to label air masses. For global atmospheric dynamics on the sphere, as modelled by the primitive equations, these labels are potential temperature (θ) and Ertel potential vorticity (q).

Modified Lagrangian Mean (MLM) states [23] are background states of the atmosphere defined

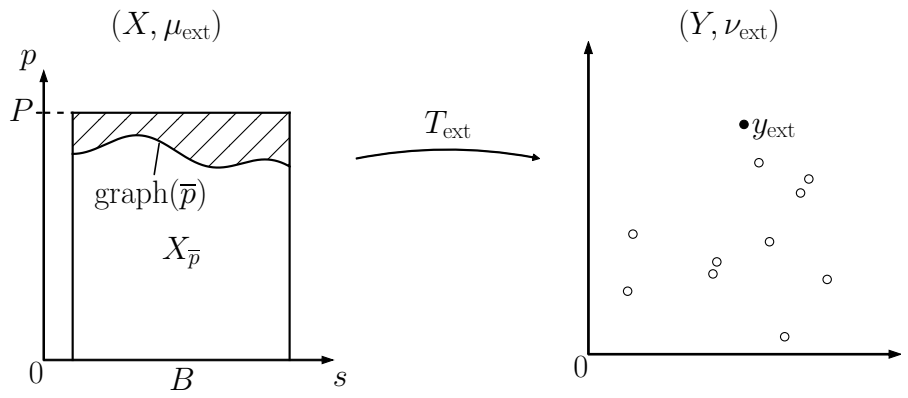


Figure 2: Reduction of Problem 2.4 to the optimal transport problem between the 'extended' source measure $\mu_{\text{ext}} = \mathcal{L}^d \llcorner (B \times [0, P])$ and the 'extended' target measure $\nu_{\text{ext}} = \nu + (P - 1)\delta_{y_{\text{ext}}}$. Here, $\mu_{\text{ext}} = \mathcal{L}^d \llcorner (B \times [0, P])$, where B is a compact subset of $\mathbb{R}^{d-1}$ and the constant $P > 0$ is an upper bound on the minimiser $\bar{p}$ of Problem 2.4. The measure ν is represented by empty circles. The point y_{ext} is represented by a filled circle, and is outside the support of ν. The cost function is extended to y_{ext} by zero. The optimal transport map T_{ext} sends the hatched cell to y_{ext}, the lower boundary of this cell is the graph of $\bar{p}$, and the restriction of T_{ext} to $X_{\bar{p}}$ is the optimal transport map from $\mu_{\bar{p}}$ to ν.

in terms of q and θ labels. Every q-contour in an MLM state is zonally symmetric and the state is a solution of the equations of motion. Atmospheric motions can be described with reference to an MLM state by displacement of q-contours relative to their reference positions. Since MLM states are zonally symmetric and time symmetric (steady) for adiabatic motions, there are two disturbance conservation laws for pseudomomentum and pseudoenergy [18, 26]. This follows generally from Noether's theorem relating conservation laws to symmetries in Hamiltonian dynamical systems [30]. However, the definition of the disturbances themselves depends on the choice of MLM state. This is not satisfactory if you would like to answer questions like whether the atmosphere is unstable and disturbances are expected to grow, so a principle by which to uniquely select an MLM state is desirable.

Several heuristics have been used to define and compute MLM states. In [23], an MLM state is defined by an adiabatic rearrangement of fluid parcels, with θ increasing with height and zonally-symmetric q increasing polewards on surfaces of constant θ. In [27], an MLM state with (approximately) the same mass distribution in coordinates (q, θ) as the full atmospheric flow is defined using two time-invariant integral quantities: mass and Kelvin's circulation. Consider an instantaneous 3-dimensional state of the atmosphere expressed in *isentropic coordinates*, where the three independent variables are sine-of-latitude s, longitude λ and potential temperature θ. Let Q be a function of (s, λ, θ) representing the potential vorticity of the 3-D state. For q-levels $q_l > 0$ and θ-levels $\theta_m > 0$ on an axis-aligned rectangular grid in $\mathbb{R}^2$ with half-level depths $\delta q > 0$ and $\delta\theta > 0$, mass and circulation integrals are computed over each volume

$$\mathcal{V}(q_l, \theta_m) \coloneqq \big\{(s, \lambda, \theta) \in [0, 1] \times [0, 2\pi) \times \mathbb{R} \;\big|\; q_l - \delta q \le Q(s, \lambda, \theta),\ \theta \in [\theta_l - \delta\theta,\, \theta_l + \delta\theta]\big\}.$$

That is, $\mathcal{V}(q_l, \theta_m)$ is the 3-dimensional volume of air bounded laterally by the q contour at $q_l - \delta q$ and vertically by the θ-surfaces at $\theta_l - \delta\theta$ and $\theta_l + \delta\theta$. The corresponding mass and circulation

integrals are

$$\mathcal{M}(q_l, \theta_m) \coloneqq \iiint_{\mathcal{V}(q_l,\theta_m)} r(s, \lambda, \theta) a^2 \, \mathrm{d}s\mathrm{d}\lambda\mathrm{d}\theta, \tag{2}$$

$$\mathcal{C}(q_l, \theta_m) \coloneqq \iiint_{\mathcal{V}(q_l,\theta_m)} \zeta(s, \lambda, \theta) a^2 \, \mathrm{d}s\mathrm{d}\lambda\mathrm{d}\theta, \tag{3}$$

where r is the isentropic mass density and $\zeta = rQ$ is the vertical component of absolute isentropic vorticity, both of the 3-dimensional state, and a is the Earth's radius. A zonally-symmetric MLM state is defined by requiring that its mass and circulation integrals are the same as those computed from the instantaneous 3-dimensional state. In this way, the MLM state is described as an adiabatic rearrangement of the full state because if the flow were adiabatic and frictionless then the full state could be accessible from the background state through adiabatic motions of the fluid, which conserve q, θ, $\mathcal{M}$ and $\mathcal{C}$.

In this paper, we rephrase the constraint that the mass and circulation integrals of the MLM state are the same as those of the full 3-D state as a constraint on the mass density of the MLM state in coordinates (z, θ), where z is zonal angular momentum. We then minimise the integral of kinetic energy and enthalpy subject to this constraint to obtain an energy-minimising MLM state. While z is not a conserved variable of the 3-D state, under the condition of zonal symmetry, it is possible to change variables from (q, θ) to (z, θ). Indeed, for a zonally symmetric state, $\mathcal{C}$ is proportional to z:

$$\mathcal{C}(q, \theta) = 2\pi z. \tag{4}$$

Restricting to positive values of q (in the Northern Hemisphere) means that the vertical component of absolute vorticity $\zeta = rq$ is everywhere positive, so $\mathcal{C}$ is strictly decreasing in q on every θ-level and (4) defines a change of variables from (q, θ) to (z, θ). In particular, the mass density of the 3-D state in (q, θ) can be pushed forward under this change of variables to coordinates (z, θ) and normalised to obtain a normalised mass density (or, more generally, a probability measure) ν. The constraint on the MLM state is satisfied precisely when its normalised mass density in coordinates (z, θ) is equal to ν.

In order to calculate the full properties of a background state, and corresponding perturbations, it is necessary and sufficient to obtain its zonal angular momentum and potential temperature as functions of latitude ϕ and pressure p, as well as the surface pressure $\bar{p}$ as a function of latitude. The function $\bar{p}$ defines the extent of the physical domain in pressure coordinates and represents the pressure at the lower boundary of the atmosphere. Let Z denote the MLM zonal angular momentum as a function of $s = \sin(\phi)$ and p, and let Θ denote the MLM potential temperature as a function of s and p. By (4), Z is constrained to take values $z = \mathcal{C}(q, \theta)/2\pi$, where q and θ are values of potential vorticity and potential temperature, respectively, labelling air masses. A zonally-symmetric adiabatic rearrangement of the 3-D circulation specifies the coordinates (s, p) to which these values of z are assigned, thus defining the function Z. Since circulation is regarded as a functional of potential vorticity and potential temperature, this also defines Θ and the MLM state potential vorticity as functions of (s, p). To obtain the MLM state zonal wind (relative to the Earth's surface) u, one may then use the relation

$$z = ua\cos\phi + \Omega a^2 \cos^2\phi, \tag{5}$$

where Ω is the planetary rotation rate. Likewise, the MLM state temperature T may be recovered via the relation

$$\theta = T\left(\frac{p}{p_r}\right)^{-\kappa}, \tag{6}$$

where p_r is a constant reference pressure and $\kappa = 2/7$ is the Poisson constant for a diatomic ideal gas.

Obtaining the background variables Z and Θ as functions of s and p, and $\bar{p}$ as a function of s, is essential to the analysis of perturbation evolution and atmospheric stability, and it is particularly desirable to obtain an MLM state that is an energy minimiser, accounting consistently for the surface pressure. With these variables, it is possible to find any quantities defined as perturbations of q-contours relative to their positions in the background state (called the "equivalent latitudes" of q-contours), as well as perturbation wind and energy. Moreover, finite-amplitude stability results for the atmosphere, which depend on knowing $\overline{p}$ [2], yield upper bounds on disturbance energy.

1.2 Optimal transport

We recall some fundamental concepts and results from optimal transport theory that form the basis of our analysis of MLM states. A more thorough exposition of optimal transport theory can be found in several standard texts such as [25], [28] or [31].

The starting point of optimal transport is the Monge problem: given X, $Y \subset \mathbb{R}^d$, a cost function $c : X \times Y \to \mathbb{R}$, Borel probability measures $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, find

$$\inf \left\{ \int_X c(x, T(x)) \, \mathrm{d}\mu(x) \,\middle|\, T : X \to Y \text{ is Borel measureable and } T_{\#}\mu = \nu \right\}. \tag{7}$$

Here, $T_{\#}\mu$ is the probability measure on Y defined by

$$(T_{\#}\mu)(A) = \mu(T^{-1}(A))$$

for all Borel sets $A \subseteq Y$, which is called the pushforward of μ by T. The spaces X and Y are often referred to as the *source* and *target* spaces, respectively, and the measures μ and ν as the source and target measures. If T satisfies the constraint $T_{\#}\mu = \nu$, then it is said to be an *admissible transport map* from μ to ν. If T attains the infimum in (7) it is said to be an optimal transport map from μ to ν.

Admissible transport maps do not necessarily exist. For example, if μ is a single Dirac mass and ν is a sum of two weighted Dirac masses at distinct locations, then the Monge problem is infeasible. The standard relaxation of the Monge problem is the Kantorovich problem, in which the optimal transport cost is defined by

$$\mathcal{T}_c(\mu, \nu) \coloneqq \inf_{\gamma \in \Gamma(\mu,\nu)} \left\{ \int_X c(x, y) \, \mathrm{d}\gamma(x, y) \right\}, \tag{8}$$

where

$$\Gamma(\mu, \nu) \coloneqq \{\gamma \in \mathcal{P}(X \times Y) \,|\, \pi_{X\#}\gamma = \mu, \ \pi_{Y\#}\gamma = \nu\},$$

and where π_X and π_Y denote the projections onto X and Y, respectively. The convex set $\Gamma(\mu, \nu)$ is called the set of *admissible transport plans* from μ to ν. Under mild assumptions on the cost function c, the Kantorovich problem has a solution γ, which is called an optimal transport plan from μ to ν, and $\mathcal{T}_c(\mu, \nu)$ is the optimal transport cost from μ to ν. The Kantorovich Duality Theorem states that

$$\mathcal{T}_c(\mu, \nu) = \sup_{\psi \in \mathcal{C}(Y)} \left\{ \int_X \psi^c(x) \, \mathrm{d}\mu(x) + \int_Y \psi(y) \, \mathrm{d}\nu(y) \right\}, \tag{9}$$

where $\mathcal{C}(Y)$ is the space of continuous functions on Y, and $\psi^c : X \to \mathbb{R} \cup \{-\infty\}$ is the c-transform of ψ defined by

$$\psi^c(x) = \inf_{y \in Y} \{c(x, y) - \psi(y)\}. \tag{10}$$

(See, e.g., [28, Theorem 1.39] or [31, Theorem 5.10].) Again, under mild assumptions on the cost function c and the sets X and Y, the supremum in (9) is attained. A function ψ that achieves the maximum is called a *Kantorovich potential*.

The Gangbo-McCann Theorem gives conditions under which the Monge problem has a unique solution. (See, for example, [17], [25, Theorem 12], or [31, Theorem 10.28].) Moreover, it ties together the three formulations of optimal transport given above. Indeed, if T is the optimal transport map, then $\gamma \coloneqq (\mathrm{id}_X, T)_{\#}\mu$ is the unique optimal transport plan, and for any Kantorovich potential ψ,

$$c(x, y) = \psi^c(x) + \psi(y) \quad \text{for } \gamma\text{-almost-every } (x, y) \in X \times Y. \tag{11}$$

In particular, the optimal transport map can be recovered from a Kantorovich potential via the expression $T(x) = (\nabla_x c(x, \cdot))^{-1}(\nabla \psi^c(x))$. This expression is well defined when $\nabla_x c(x, \cdot)$ exists and is injective, a condition known as the *twist* condition.

1.3 Relating Modified Lagrangian Mean states to optimal transport

To formulate the definition of an MLM state mathematically, consider again a instantaneous 3-dimensional atmospheric state. Let ν be a probability measure on $\mathbb{R}^2$ representing the normalised mass distribution of this state in zonal angular momentum and potential temperature coordinates (obtained from the 3-D atmospheric state as explained in the introduction). We consider states with positive zonal angular momentum. Since potential temperature is positive, and both zonal angular momentum and potential temperature are bounded, ν has compact support contained in $(0, +\infty)^2$. Let $Y \subset (0, +\infty)^2$ be a compact set containing the support of ν. Let $s = \sin(\phi)$, where ϕ is latitude. Let $p_{\min} > 0$ be a prescribed and arbitrary minimum pressure, and let p be the positive deviation of pressure from $p_{\min}$. We seek zonally-symmetric MLM states defined in pressure coordinates (s, p) on a single hemisphere with a spherical cap from the pole and a small region close to the equator removed. That is, for some ε_0, $\varepsilon_1 \in (0, 1/2)$, we consider the physical domain

$$X \coloneqq [\varepsilon_0, 1 - \varepsilon_1] \times [0, +\infty)$$

containing points $x = (s, p)$. Under the assumption of hydrostatic balance (i.e., that the fluid density multiplied by gravitational acceleration g, is the negative vertical gradient of pressure) the fluid density in pressure coordinates (s, p), often called the pseudodensity, is uniform (and equal to $-1/g$) [19, Chapter 6]. An MLM state is then defined by three maps: surface pressure $\overline{p}$, potential temperature Θ, and zonal angular momentum Z. In the following definition $\mathcal{P}_{\mathrm{ac}}([\varepsilon_0, 1-\varepsilon_1])$ denotes the space of probability measures on $[\varepsilon_0, 1 - \varepsilon_1]$ that are absolutely continuous with respect to the Lebesgue measure, and we conflate such measures with their densities.

Definition 1.1 (MLM state, c.f. Figure 1)**.** *Let $\overline{p} \in \mathcal{P}_{\mathrm{ac}}([\varepsilon_0, 1 - \varepsilon_1])$. Define the domain*

$$X_{\overline{p}} \coloneqq \{(s, p) \in X \,:\, p \le \overline{p}(s)\}.$$

Consider Borel measurable maps $Z, \Theta : X_{\overline{p}} \to \mathbb{R}$, with extension by zero to X. The triple $(\overline{p}, Z, \Theta)$ is an MLM state for ν if and only if

$$(Z, \Theta)_{\#}\mu_{\overline{p}} = \nu, \tag{12}$$

where

$$\mu_{\overline{p}} \coloneqq \mathcal{L}^d \llcorner X_{\overline{p}}$$

is the restriction of the d-dimensional Lebesgue measure to the set $X_{\overline{p}}$*.*

Interpreted physically, the pushforward constraint (12) means that $(\overline{p}, Z, \Theta)$ represents a zonally-symmetric mass-preserving rearrangement of the instantaneous circulation and potential temperature of the 3-D state. Mathematically, it means that (Z, Θ) is an *admissible transport map* from $\mu_{\overline{p}}$ to ν.

We now define a cost function that represents the sum of kinetic energy and enthalpy of a zonally-symmetric state, and we interpret the corresponding transport cost as the energy of this state. Our goal of finding energy-minimising MLM states can then be phrased as minimising the optimal transport cost to ν over all source measures of the form $\mu_{\overline{p}}$ for $\overline{p} : [\varepsilon_0, 1 - \varepsilon_1] \to [0, +\infty)$.

Definition 1.2 (Background state cost function)**.** *Define the cost function* $c : X \times Y \to [0, +\infty)$ *by*

$$c((s,p),(z,\theta)) \coloneqq \frac{1}{2}\left(\frac{z}{a\sqrt{1-s^2}} - \Omega a\sqrt{1-s^2}\right)^2 + c_{\mathrm{p}}\theta\left(\frac{p+p_{\min}}{p_r}\right)^{\kappa}, \tag{13}$$

where $a > 0$ *is the radius of the earth,* $\Omega > 0$ *is the angular velocity of the earth,* $c_{\mathrm{p}} > 0$ *is the specific heat capacity of air at constant pressure,* $p_{\min} > 0$ *is a prescribed and arbitrary minimum pressure,* $p_{\mathrm{r}} > 0$ *is a constant reference pressure, and* $\kappa = R/c_{\mathrm{p}}$*, where* R *is the gas constant so that* $\kappa = 2/7$*.*

Remark 1.3. *Using* $p_{\min} > 0$*, we force pressure to be positive, which ensures that the cost function is locally Lipschitz. This is not restrictive when computing MLM states because pressure tending zero corresponds to leaving the atmosphere and entering outer space, so atmospheric data will only ever extend to small but non-zero pressure. However,* $p = 0$ *would be the natural physical boundary condition, so we hope to consider the case* $p_{\min} = 0$ *in future work.*

Using the relations (5) and (6), we see that

$$c = \frac{1}{2}u^2 + c_{\mathrm{p}}T,$$

which is the sum of kinetic energy and enthalpy. The energy E of an MLM state $(\overline{p}, Z, \Theta)$ is given by the integral of kinetic energy and enthalpy over the whole hemispheric domain:

$$E(\overline{p}, Z, \Theta) = \int_{\varepsilon_0}^{1-\varepsilon_1} \int_0^{\overline{p}(s)} c\big((s,p),(Z(s,p),\Theta(s,p))\big)\,\mathrm{d}p\,\mathrm{d}s. \tag{14}$$

It follows by definition of $\mu_{\overline{p}}$ that

$$E(\overline{p}, Z, \Theta) = \int_X c\big(x,(Z,\Theta)(x)\big)\,\mathrm{d}\mu_{\overline{p}}(x).$$

This is precisely the cost of transporting the measure $\mu_{\overline{p}}$ to the measure ν using the admissible transport map (Z, Θ) with the cost function c. For a given surface pressure $\overline{p}$, the minimal energy over all (Z, Θ) such that $(\overline{p}, Z, \Theta)$ is an MLM state is $\mathcal{T}_c(\mu_{\overline{p}}, \nu)$. To minimise the energy E over *all* MLM states $(\overline{p}, Z, \Theta)$, we therefore minimise $\mathcal{T}_c(\mu_{\overline{p}}, \nu)$ over all surface pressures $\overline{p} \in \mathcal{P}_{\mathrm{ac}}([\varepsilon_0, 1-\varepsilon_1])$.

Problem 1.4 (MLM state energy minimisation)**.** *Given a distribution $\nu \in \mathcal{P}(Y)$ of zonal angular momentum and potential temperature, find a corresponding MLM state $(\overline{p}_*, Z_*, \Theta_*)$ with minimal energy. That is, find*

$$\overline{p}_* \in \operatorname*{arg\,min}_{\overline{p} \in \mathcal{P}_{\mathrm{ac}}([\varepsilon_0, 1-\varepsilon_1])} \mathcal{T}_c(\mu_{\overline{p}}, \nu),$$

and

$$(Z_*, \Theta_*) \in \operatorname*{arg\,min}_{(Z,\Theta)_{\#}\mu_{\overline{p}_*} = \nu} \int_X c\big(x, (Z,\Theta)(x)\big)\, \mathrm{d}\mu_{\overline{p}_*}(x).$$

We prove existence, uniqueness and stability of solutions of a general formulation of Problem 1.4, defined in Problem 2.4, under assumptions on the cost function and the source and target spaces as set out in Section 2. In Proposition 2.7 we show that the cost function (13) defining the energy of an MLM state satisfies these assumptions, meaning that the results that we prove are applicable to Problem 1.4.

1.4 Rearrangements and optimal transport in the atmospheric sciences

The current work builds on the long-standing connection between optimal transport and the atmospheric sciences [7]. The cornerstone of this connection is the interpretation of energy-minimising mass-preserving rearrangements of vector fields as optimal transport maps. We use this perspective to interpret the definition of MLM states in terms of optimal transport, which enables us to prove their existence, uniqueness and stability. Meteorologists Cullen and Purser first used *monotone* rearrangements of vector fields in their study of the *semi-geostrophic* (SG) equations [12], which model the formation of atmospheric fronts, and found them to be minimisers of the *geostrophic energy*. Brenier's celebrated Polar Factorisation Theorem [3] later identified such rearrangements as optimal transport maps for the quadratic cost. Interpreting the geostrophic energy as a transport cost and the energy-minimising rearrangement as an optimal transport map subsequently allowed for the rigorous mathematical analysis of the SG equations in several settings (see [7] for a comprehensive overview).

The 3-dimensional free surface [6, 10], 2-dimensional shallow water [8], and 3-dimensional compressible [11, 15, 9] variants of the SG equations, and the model for forced axisymmetric flows considered in [13], are all continuity equations with velocity defined at each time instant by minimising a transport cost over a space of source measures, similar to Problem 1.4. In the 3-dimensional free surface SG equations, given a target measure $\nu \in \mathcal{P}(\mathbb{R}^3)$, the energy functional is defined for $\bar{p} \in \mathcal{P}_{\mathrm{ac}}(\mathbb{R}^2)$, modelling surface pressure, by

$$\bar{p} \mapsto \mathcal{T}_c(\mu_{\bar{p}}, \nu).$$

Here $\mu_{\bar{p}}$ is the restriction of the 3-dimensional Lebesgue measure to the subgraph of $\bar{p}$ intersected with the upper half space, as in Definition 1.1, and the cost function is given by

$$c(x, y) = \frac{1}{2}|x_1 - y_1|^2 + \frac{1}{2}|x_2 - y_2|^2 + x_3 y_3. \tag{15}$$

For the 3-dimensional compressible SG equations in pressure coordinates [7, Section 3.3.6], the energy functional takes the same form but with the term $x_3 y_3$ in the cost function replaced by $x_3^{\kappa} y_3$, up to a positive multiplicative constant. Both corresponding energy minimisation problems are special cases of Problem 2.4. The vertical coordinate of the target space represents potential temperature, so the support of the target measure ν is contained in the upper half-space. It is straightforward to verify that the cost function (15) satisfies Assumption 2.1 and the twist condition

(Assumption 2.5). A similar argument as used in Proposition 2.7 verifies that these assumptions are also satisfied by the cost function of the 3-D compressible SG equations in pressure coordinates.

The 2-dimensional shallow water SG equations are the special case of the 3-dimensional free surface SG equations where potential temperature is constant. The energy functional then takes the form

$$\rho \mapsto \mathcal{T}_c(\rho, \nu) + \int \rho^\gamma \tag{16}$$

where ρ represents the fluid density, ν is a given target measure, c is a given transport cost, and $\gamma = 2$. This energy minimisation problem is the same as computing the Moreau envelope in the Wasserstein space, as studied in [29]. The energy functional for the compressible SG equations is also of the form (16) but with $\gamma \in (1, 2)$.

Outside of applications in the atmospheric sciences, there are clear parallels between our approach to defining and computing MLM states and the use of optimal transport in the problem of reconstructing the early Universe. There, the aim is to compute a mapping between mass distributions representing the early and late Universe. The cost function is quadratic and the transport cost represents the integrated kinetic energy, while the Kantorovich potential ψ and its c-transform relate to the final and initial gravitational potentials. Not only has optimal transport been used to prove the existence of a unique reconstruction under the condition that the fluid velocity is the gradient of a convex potential [16, 4], it also provides the basis of a numerical method. In [22], a semi-discrete optimal transport algorithm is used to compute a mapping between a uniform distribution of mass, representing the early Universe, onto a discrete mass distribution, where each point represents a galaxy in the late Universe. The efficiency of this method means that it can be applied at scale, enabling the reconstruction of detailed features and making it suitable for practical purposes.

1.5 Towards numerical computation of MLM states

The MLM state in [27] was computed using a numerical technique called equivalent latitude iteration with potential vorticity inversion (ELIPVI). While numerical solutions were obtained, questions of existence and uniqueness of such solutions, as well as convergence of the numerical scheme, were not addressed. It was also not considered whether or not the computed background states were energy minimising.

We show in Section 5 that the problem of finding energy minimising MLM states (Problem 1.4) reduces to an optimal transport problem where a single weighted Dirac mass has been added to the target measure and the cost to that point is set to zero, as illustrated in Figure 2. This is similar in spirit to the notion of partial optimal transport considered in [21] and used to generate Lagrangian meshes with free boundaries. With a view to solving Problem 1.4 numerically, in Corollary 5.3 we consider the case where the mass distribution ν is a discrete measure. This corresponds precisely to the setting considered in [27]. There, mass and circulation integrals $\mathcal{M}(q_l, \theta_m)$ and $\mathcal{C}(q_l, \theta_m)$ are computed for the full 3-dimensional atmospheric state for points (q_l, θ_m) on an axis-aligned rectangular grid in $\mathbb{R}^2$ with half-level depths δq and $\delta\theta$ (see equations (2) and (3)). Correspondingly, ν would be given by

$$\nu = \frac{1}{M} \sum_{(l,m)} \left(\mathcal{M}(q_l, \theta_m) - \mathcal{M}(q_{l+1}, \theta_m)\right) \delta_{\left(\frac{\mathcal{C}(q,\theta)}{2\pi}, \theta\right)}, \tag{17}$$

where M is a normalisation constant and we recall that circulation is proportional to zonal angular momentum for zonally-symmetric flow (4). By Corollary 5.3, Problem 1.4 then reduces to a

standard semi-discrete optimal transport problem, which is numerically tractable [25, Section 4]. Hence, our analysis not only provides rigorous means by which to select an MLM state given a full atmospheric flow, with guarantees of existence, uniqueness and stability, it also provides a numerical method for computing such states. As a proof of concept, Figure 3 contains a numerically-computed energy-minimising MLM state obtained from climate reanalysis data using this numerical method. A more comprehensive implementation of this method will be the subject of a separate work.

1.6 Notation and conventions

For $d \in \mathbb{N} = \{1, 2, 3, \ldots\}$, let $A \subseteq \mathbb{R}^d$ be a Borel set. We denote the interior of A by $\mathrm{int}(A)$, the boundary of A by ∂A, the diameter of A by $\mathrm{diam}(A)$, and the characteristic function of A by $\mathbb{1}_A$. That is,

$$\mathbb{1}_A(x) = \begin{cases} 1 \text{ if } x \in A, \\ 0 \text{ otherwise.} \end{cases}$$

For a function $f : A \to \mathbb{R}$, we denote by $\mathrm{graph}(f)$ its graph, that is, $\mathrm{graph}(f) \coloneqq \{(x, f(x)) : x \in A\}$. We denote the Lebesgue measure of dimension d by $\mathcal{L}^d$, and its restriction to A by $\mathcal{L}^d \llcorner A$. The space of Borel probability measures on A is denoted by $\mathcal{P}(A)$. We denote by $\mathrm{spt}(\mu)$ the support of a measure $\mu \in \mathcal{P}(A)$. The set of Borel probability measures on A that are absolutely continuous with respect to $\mathcal{L}^d$ is denoted by $\mathcal{P}_{\mathrm{ac}}(A)$. We conflate a measure $\mu \in \mathcal{P}_{\mathrm{ac}}(A)$ with its density with respect to $\mathcal{L}^d$, and the corresponding element of $L^1(A)$. We denote by $\mathcal{C}(A)$ and $\mathcal{C}_b(A)$ the spaces of continuous functions on A and bounded continuous functions on A, respectively. When A is a compact set, we equip $\mathcal{C}(A)$ with the uniform norm, which we denote by $\|\cdot\|_{\mathcal{C}(A)}$. Given Borel sets A, $B \subseteq \mathbb{R}^d$, a cost function $c : A \times B \to \mathbb{R}$, and measures $\mu \in \mathcal{P}(A)$ and $\nu \in \mathcal{P}(B)$, we denote by $\mathcal{T}_c(\mu, \nu)$ the optimal transport cost from μ to ν. For a Borel measureable function $f : A \to B$, the pushforward of μ by f is the measure $f\#\mu \in \mathcal{P}(B)$ defined by $(f\#\mu)(U) \coloneqq \mu(f^{-1}(U))$ for Borel sets $U \subseteq B$. For $\phi \in \mathcal{C}(A)$ and $\psi \in \mathcal{C}(B)$ we denote by $\phi \oplus \psi$ the element of $\mathcal{C}(A \times B)$ defined by $(\phi \oplus \psi)(x, y) = \phi(x) \oplus \psi(y)$. For $\mu \in \mathcal{P}(A)$ and $\nu \in \mathcal{P}(B)$ we denote by $\mu \otimes \nu$ the product measure defined by $(\mu \otimes \nu)(U \times V) = \mu(U)\nu(V)$ for Borel sets $U \subseteq A$ and $V \subseteq B$.

1.7 Outline and main contributions

A precise mathematical statement of the energy minimisation problem, Problem 2.4, studied in this paper is given in Section 2. Our main contributions are:

- Theorem 3.10, which establishes the existence and uniqueness of solutions via duality, and gives optimality conditions relating the solutions of the primal and dual problems.
- Theorem 4.1, which establishes stability of solutions with respect to input data.
- Theorem 5.1 and Corollaries 5.2 and 5.3, which reduce the free surface minimisation problem, Problem 2.4, to a standard optimal transport problem, and treat the case of twisted costs and discrete target measures.

2 Problem statement and assumptions

In this section we precisely state the problem under study in this paper (Problem 2.4), which is a direct generalisation of Problem 1.4 for finding energy-minimising MLM states. We also state

assumptions under which our results hold (Assumptions 2.1, 2.5, and 2.6), and show that they are all satisfied by the cost defining the energy of an MLM state; see Proposition 2.7.

Let $d \geq 2$, and let $B \subset \mathbb{R}^{d-1}$ be a compact set with non-empty interior. Define the source domain $X \coloneqq B \times [0, +\infty)$, and let $Y \subset \mathbb{R}^d$ be a non-empty compact target domain. We denote by $x = (s, p)$ a point in X and by y a point in Y.

Assumption 2.1. *The cost function $c : X \times Y \to [0, +\infty)$ satisfies the following properties:*

A. c is locally Lipschitz.

B. For all $s \in B$ and all $y \in Y$ the function $c((s, \cdot), y) : [0, +\infty) \to [0, +\infty)$ is strictly increasing and unbounded.

C. The set of functions $\{c((\cdot, p), y) : B \to [0, +\infty) \,|\, p \in [0, +\infty),\, y \in Y\}$ is equicontinuous.

Unless otherwise stated, we consider cost functions $c : X \times Y \to [0, +\infty)$ satisfying only Assumption 2.1.

Definition 2.2 (Primal functional)**.** *For $\bar{p} \in \mathcal{P}_{\mathrm{ac}}(B)$ define the set*

$$X_{\bar{p}} \coloneqq \{(s, p) \in X \,|\, p \leq \bar{p}(s)\}.$$

Define the functional

$$\begin{aligned} &\mu : \mathcal{P}_{\mathrm{ac}}(B) \to \mathcal{P}_{\mathrm{ac}}(X), \\ &\bar{p} \mapsto \mu_{\bar{p}} \coloneqq \mathcal{L}^d \llcorner X_{\bar{p}}. \end{aligned} \tag{18}$$

For $\nu \in \mathcal{P}(Y)$ and a cost function $c : X \times Y \to [0, +\infty)$, we define the primal functional

$$\begin{aligned} &F_\nu : \mathcal{P}_{\mathrm{ac}}(B) \to [0, +\infty), \\ &F_\nu(\bar{p}) \coloneqq \mathcal{T}_c(\mu_{\bar{p}}, \nu), \end{aligned}$$

where $\mathcal{T}_c$ is the optimal transport cost from $\mu_{\bar{p}}$ to ν for the cost function c (8).

Remark 2.3. *For each $\bar{p} \in \mathcal{P}_{\mathrm{ac}}(B)$, $X_{\bar{p}}$ is well defined up to $\mathcal{L}^d$-negligible sets, $\mathbb{1}_{X_{\bar{p}}}$ is a well-defined element of $L^1(X)$ and is the Lebesgue density of $\mu_{\bar{p}}$, and $\mu_{\bar{p}}$ is a probability measure. Moreover, by definition, for any $\bar{p}_0$, $\bar{p}_1 \in \mathcal{P}_{\mathrm{ac}}(B)$,*

$$\|\mu_{\bar{p}_0} - \mu_{\bar{p}_1}\|_{L^1(X)} = \|\bar{p}_0 - \bar{p}_1\|_{L^1(B)},$$

and for any $\xi \in \mathcal{C}_b(X)$,

$$\int_X \xi \, \mathrm{d}\mu_{\bar{p}_0} = \int_B \int_0^{\bar{p}_0(s)} \xi(s, p) \, \mathrm{d}p \, \mathrm{d}s.$$

Problem 2.4 (Primal problem)**.** *Given $\nu \in \mathcal{P}(Y)$, find*

$$\bar{p}_* \in \operatorname*{arg\,min}_{\bar{p} \in \mathcal{P}_{\mathrm{ac}}(B)} F_\nu(\bar{p}).$$

When Assumption 2.1 holds, for any $\nu \in \mathcal{P}(Y)$ Problem 1.4 has a unique solution (Theorem 3.10) and this solution has a continuous density. The set of all such solutions is uniformly bounded and equicontinuous (Lemma 3.5), and Problem 1.4 is stable with respect to ν (Theorem 4.1).

Assumption 2.5 (Twist condition)**.** *There exist open sets Ω_X, $\Omega_Y \subset \mathbb{R}^d$ such that $X \subset \Omega_X$ and $Y \subset \Omega_Y$, and an extension of the cost function c onto $\Omega_X \times \Omega_Y$ such that $c \in \mathcal{C}^1(\Omega_X \times \Omega_Y)$ and for all $x_0 \in \Omega_X$ the function $\nabla_x c(x_0, \cdot) : \Omega_Y \to \mathbb{R}^d$ is injective.*

The twist condition is a classical sufficient condition for the existence of transport maps [25, Definition 8]. We show that when both Assumptions 2.1 and 2.5 hold, Problem 2.4 admits a solution $\bar{p}$ and that there exists a unique optimal transport map between $\mu_{\bar{p}}$ and ν, and that this map is stable with respect to ν; see Theorem 4.2.

In Section 5, we show that under Assumption 2.1, Problem 2.4 can be reduced to a single optimal transport problem by appropriate extension of the source and target measures. For the existence of a corresponding optimal transport map, we impose the following condition on the cost function c.

Assumption 2.6 (Extended twist condition)**.** *Assumption 2.5 holds and there exists a constant $\ell > 0$ such that $\min_{y \in Y} \|\nabla_x c(x_0, y)\|_{\mathbb{R}^d} \geq \ell$ for all $x_0 \in \Omega_X$.*

We conclude this section by showing that the cost function (13) defining the background state energy (14) satisfies Assumptions 2.1 and 2.6.

Proposition 2.7. *Let ε_0, $\varepsilon_1 \in (0, 1/2)$, define $X = [\varepsilon_0, 1 - \varepsilon_1] \times [0, +\infty)$, and let $Y \subset (0, +\infty)^2$ be compact and non-empty. The cost function $c : X \times Y \to [0, +\infty)$ defined by (13) satisfies assumptions 2.1 and 2.6.*

Proof. First we show that c satisfies Assumption 2.6. Define $\Omega_X := (\varepsilon_0/2, 1-\varepsilon_1/2) \times (-p_{\min}/2, +\infty)$. Since $Y \subset (0, +\infty)^2$ is compact, there exists $\theta_{\min} > 0$ such that $\Omega_Y := (0, +\infty) \times (\theta_{\min}, +\infty)$ contains Y. Extend c to $\Omega_X \times \Omega_Y$ by the formula (13). Then $c \in \mathcal{C}^1(\Omega_X \times \Omega_Y)$ and for $((s,p),(z,\theta)) \in \Omega_X \times \Omega_Y$,

$$\nabla_{(s,p)} c((s,p),(z,\theta)) = \begin{pmatrix} \dfrac{z^2 s}{a^2 \left(1 - s^2\right)^2} - \Omega^2 a^2 s \\ \dfrac{\kappa c_{\mathrm{p}} \theta}{p_r} \left(\dfrac{p + p_{\min}}{p_r} \right)^{\kappa - 1} \end{pmatrix}. \tag{19}$$

Since $s > 0$ and $p + p_{\min} > 0$, $\nabla_{(s,p)} c((s,p), \cdot)$ is injective on Ω_Y. It is immediate from (19) that for all $(s,p) \in \Omega_X$,

$$\|\nabla_{(s,p)} c((s,p), \cdot)\|_{\mathcal{C}(\Omega_Y)} \geq \frac{\kappa c_{\mathrm{p}} \theta_{\min}}{p_r} \left(\frac{p_{\min}}{2 p_r} \right)^{\kappa - 1} =: \ell > 0,$$

so Assumption 2.6 holds.

Now we show that c satisfies Assumption 2.1. Since $c \in \mathcal{C}^1(\Omega_X \times \Omega_Y)$, it is locally Lipschitz on $X \times Y$, so Assumption 2.1.A. holds. Assumption 2.1.B. holds because κ, c_{p}, $p_{\min}$, $p_r > 0$, so the map

$$p \mapsto c_{\mathrm{p}} \theta \left(\frac{p + p_{\min}}{p_r} \right)^{\kappa}$$

is strictly increasing and unbounded on $[0, +\infty)$ for all $\theta \in (\theta_{\min}, +\infty)$. Since Y is compact, there exists a positive constant $z_{\max}$ such that $Y \subset (0, z_{\max}) \times (0, +\infty)$. Let $((s,p),(z,\theta)) \in X \times Y$. Since $s \leq 1 - \varepsilon_1$ and $\varepsilon_1 \in (0, 1/2)$,

$$a^2 \left(1 - s^2\right)^2 \geq a^2 \left(1 - (1 - \varepsilon_1)^2\right)^2 = a^2 \varepsilon_1^2 (2 - \varepsilon_1^2) \geq a^2 \varepsilon_1^2.$$

It follows from (19) that

$$\left| \frac{\partial}{\partial s} c((s,p),(z,\theta)) \right| \leq \frac{z_{\max}^2}{a^2 \varepsilon_1^2} + \Omega^2 a^2 =: L.$$

The set

$$\left\{c((\cdot,p),(z,\theta)) : [\varepsilon_0, 1-\varepsilon_1] \to [0,+\infty) \;\middle|\; p \in [0,+\infty),\ (z,\theta) \in Y\right\}$$

therefore consists of L-Lipschitz functions so is equicontinuous (Assumption 2.1.C.). □

Since the cost function (13) defining the background state energy (14) satisfies Assumptions 2.1 and 2.6, the existence, uniqueness, and stability of energy-minimising MLM states (i.e., solutions of Problem 1.4), follows from Theorems 3.10, 4.1, and 4.2. Moreover, by Corollary 5.2, any such state can be found by solving the Monge problem (7) with an adjusted target measure and an extended cost function.

3 Existence and uniqueness of optimal surfaces

In this section, we prove that, for any $\nu \in \mathcal{P}(Y)$, Problem 2.4 has a unique solution that is characterised by necessary and sufficient optimality conditions. This result is summarised in Theorem 3.10. We work in the setting described in Section 2 and our working hypothesis is that the cost function c satisfies Assumption 2.1. For brevity, we fix ν and will often drop subscripts ν in the notation. We begin by formally deriving a dual maximisation problem for the minimisation Problem 2.4, as well as a formula for optimal surfaces. The crucial step of the analysis is to prove uniform a priori estimates on the boundedness and continuity of these surfaces (Lemma 3.5), which enables us to prove the existence of a maximiser in the dual problem (Theorem 3.9). Optimality conditions that we prove for the dual problem then link back to the primal problem and allow us to prove our uniqueness and existence result.

We now formally derive the dual problem and thereby introduce the functionals that will be analysed in this section. By Kantorovich duality,

$$F_\nu(\bar{p}) = \sup_{\psi \in \mathcal{C}(Y)} \left\{ \int_B \int_0^{\bar{p}(s)} \psi^c(s,p)\,\mathrm{d}p\,\mathrm{d}s + \int_Y \psi\,\mathrm{d}\nu \right\}. \tag{20}$$

After taking the infimum over $\bar{p}$, the infimum and supremum may be swapped up to inequality to yield weak duality

$$\inf_{\bar{p} \in \mathcal{P}_{\mathrm{ac}}(B)} F_\nu(\bar{p}) \le \sup_{\psi \in \mathcal{C}(Y)} G_\nu(\psi), \tag{21}$$

where we define the dual functional

$$\begin{aligned} &G_\nu : \mathcal{C}(Y) \to \mathbb{R}, \\ &G_\nu(\psi) = \inf_{\bar{p} \in \mathcal{P}_{\mathrm{ac}}(B)} \left\{ \int_B \int_0^{\bar{p}(s)} \psi^c(s,p)\,\mathrm{d}p\,\mathrm{d}s + \int_Y \psi\,\mathrm{d}\nu \right\}. \end{aligned} \tag{22}$$

Define the function

$$\begin{aligned} &f_\psi : B \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}, \\ &f_\psi(s,u) = \begin{cases} \displaystyle\int_0^u \psi^c(s,p)\,\mathrm{d}p & \text{if } u \ge 0, \\ +\infty & \text{otherwise,} \end{cases} \end{aligned}$$

and define the *inner-dual* functional by

$$\begin{aligned} &H_{\nu,\psi} : \mathcal{P}(B) \to \mathbb{R} \cup \{+\infty\}, \\ &H_{\nu,\psi}(\bar{p}) = \begin{cases} \displaystyle\int_B f_\psi(s,\bar{p}(s))\,\mathrm{d}s + \int_Y \psi\,\mathrm{d}\nu & \text{if } \ \bar{p} \in \mathcal{P}_{\mathrm{ac}}(B), \\ +\infty & \text{otherwise.} \end{cases} \end{aligned} \tag{23}$$

By construction,

$$G_\nu(\psi) = \inf_{\bar{p} \in \mathcal{P}(B)} H_{\nu,\psi}(\bar{p}).$$

Notice that, under the assumptions on the cost function and for each fixed $s \in B$, the function $f_\psi(s, \cdot)$ is strictly convex on the interval $[0, +\infty)$ and convex, proper and lower semi-continuous. Indeed, the partial derivative of f_ψ with respect to u on the interval $[0, +\infty)$ is $u \mapsto \psi^c(s, u)$. (Here we take the right-derivative at the left endpoint.) This is the pointwise minimum over a set of strictly increasing functions, which is itself strictly increasing. The properties of f_ψ suggest that the inner-dual functional $H_{\nu,\psi}$ is convex. If this were true, then the primal functional F_ν, as the pointwise supremum over a set of convex functions, would itself be convex. On the other hand, the dual functional G_ν is defined as a pointwise infimum over a set of Kantorovich dual functionals parametrised by $\bar{p}$, each of which is concave. As such, G_ν is concave.

Given convexity of F_ν and concavity of G_ν, the general framework of the Fenchel-Rockafellar Duality Theorem could be applied, at least formally, in a way similar to the proof of the Kantorovich Duality Theorem in [5, Section 6.4.1] to deduce that the inequality in (21) is, in fact, an equality. The corresponding primal-dual optimality conditions, which are necessary and sufficient conditions for $\bar{p}$ and ψ to be optimal in their respective problems, are that $\bar{p}$ minimises the functional $H_{\nu,\psi}$ and ψ is maximal in (20). The latter of these conditions means that ψ is a Kantorovich potential for source measure $\mu_{\bar{p}}$ and target measure ν, and is similar to the result that the subdifferential of the optimal transport cost with respect to the source measure is the set of Kantorovich potentials; see e.g. [28, Proposition 7.17]. We see, then, that to rigorously characterise solutions $\bar{p}$ of the primal problem, it is crucial to analyse the inner-dual functional $H_{\nu,\psi}$ and to characterise its minimisers. By doing so, we arrive at a direct proof of duality and optimality that avoids explicit use of the Fenchel-Rockafellar Duality Theorem and the associated functional framework, but we note that this approach could be pursued to yield similar results.

3.1 Minimisers of the inner-dual functional

In this section, we analyse the inner-dual functional H_ψ for $\psi \in \mathcal{C}(Y)$, where the subscript ν is dropped for brevity. A formal calculation suggests that the first variation (in the sense of [28, Definition 7.12]) of H_ψ at $\bar{p}$ is the function

$$B \ni s \mapsto \frac{\partial f_\psi}{\partial u}(s, \bar{p}(s)) = \psi^c(s, \bar{p}(s)) \in \mathbb{R}.$$

Then, if $\bar{p}$ is a minimiser of H_ψ with continuous density, the first-order necessary optimality conditions stated in [28, Theorem 7.20] would imply that there exists a constant $\tau \in \mathbb{R}$ such that

$$\begin{cases} \psi^c(s, \bar{p}(s)) = \tau & \text{if } \bar{p}(s) > 0, \\ \psi^c(s, \bar{p}(s)) \geq \tau & \text{if } \bar{p}(s) = 0. \end{cases} \tag{24}$$

This is formally equivalent to the condition that the subdifferential of H_ψ at $\bar{p}$ contains the zero element. By definition of the c-transform, if (24) holds then for any $s \in B$,

$$c((s, \bar{p}(s)), y) - \psi(y) \geq \tau \quad \forall\, y \in Y.$$

Since c and ψ are continuous on the compact set Y, if $\bar{p}(s) > 0$ then this holds with equality for some $y \in Y$. By Assumptions 2.1.A.-B., $c((s, \cdot), y)$ is continuous and strictly increasing on $[0, +\infty)$, so it is invertible and its inverse is strictly increasing and non-negative. It follows that

$$\bar{p}(s) \geq c((s, \cdot), y)^{-1}(\psi(y) + \tau) \quad \forall\, y \in Y,$$

which implies that

$$\bar{p}(s) \geq \max_{y \in Y} \left\{ c((s,\cdot),y)^{-1}(\psi(y)+\tau) \right\},$$

with equality if $\bar{p}(s) > 0$. This motivates the following definition.

Definition 3.1. *Define* $q : B \times Y \times \mathbb{R} \to [0,+\infty)$ *by*

$$q(s,y,\sigma) \coloneqq \begin{cases} c((s,\cdot),y)^{-1}(\sigma) & \text{if } \sigma \geq c((s,0),y), \\ 0 & \text{otheriwse.} \end{cases} \tag{25}$$

To formulate our candidate for the unique minimiser of H_ψ, we show in Lemma 3.3 that there exists a unique constant $\tau = \tau_\psi$ such that $\bar{p} = \bar{p}_\psi$ defined by inverting the optimality condition (24) is a probability density. In Theorem 3.4, we will show that $\bar{p}_\psi$ is indeed the unique minimiser of H_ψ. We begin with the following lemma, which will also be used in Lemma 3.5 to show that the set of all functions $\bar{p}_\psi$ for $\psi \in \mathcal{C}(Y)$ is uniformly bounded and equicontinuous.

Lemma 3.2. *For any compact interval* $U \subset \mathbb{R}$*, the set*

$$\mathcal{Q} \coloneqq \{q(\cdot,y,\cdot)|_{B\times U}\}_{y\in Y}$$

is uniformly bounded and equicontinuous.

Proof. Let $U \subset \mathbb{R}$ be a compact interval. We first prove that $\mathcal{Q}$ is uniformly bounded. Note that

$$q(s,y,\sigma) = \max\left\{0,\ \min\{t \in \mathbb{R} : c((s,t),y) \geq \sigma\}\right\}. \tag{26}$$

Let $(p_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ diverging to $+\infty$, for each $n \in \mathbb{N}$ let

$$(s_n, y_n) \in \arg\min\{c((s,p_n),y) :\ s \in B,\ y \in Y\},$$

and let (s_*, y_*) be a cluster point of the sequence $((s_n,y_n))_{n\in\mathbb{N}}$, which exists by compactness of B and Y. Assume for contradiction that there exists a constant $M > 0$ such that

$$c((s_n,p_n),y_n) = \min\{c((s,p_n),y) : s \in B,\ y \in Y\} \leq M \quad \forall\, n \in \mathbb{N}.$$

Then, by continuity of c (Assumption 2.1.A.), for sufficiently large $n \in \mathbb{N}$, $c((s_*,p_n),y_*)$ is bounded by $M+1$. This contradicts Assumption 2.1.B. that $c((s_*,\cdot),y_*)$ is unbounded. Hence,

$$\lim_{n\to\infty} \min\{c((s,p_n),y) :\ s \in B,\ y \in Y\} = +\infty, \tag{27}$$

and, in particular, there exists $P > 0$ such that

$$\min\{c((s,P),y) :\ s \in B,\ y \in Y\} \geq \max_{\sigma\in U} \sigma.$$

Then, by (26), $\mathcal{Q}$ is uniformly bounded by P.

It remains to show that $\mathcal{Q}$ is equicontinuous. Let $P > 0$ be a uniform upper bound for $\mathcal{Q}$. Define $\omega : [0,+\infty) \to [0,+\infty)$ by

$$\omega(\delta) = \min\left\{c((s,p+\delta),y) - c((s,p),y)\ :\ (s,p,y) \in B \times [0,P] \times Y\right\}.$$

We will use ω to construct a common modulus of continuity for functions in $\mathcal{Q}$. By Assumption 2.1.B., $c((s,\cdot),y)$ is strictly increasing for all $(s,y) \in B \times Y$, so ω is strictly increasing. By (27)

and the fact that c is bounded on the compact set $B \times [0, P] \times Y$, ω is unbounded. Next, for arbitrary $n \in \mathbb{N}$, let $A = [0, n]$. By Assumption 2.1.A., c is Lipschitz on $(B \times [0, P + n]) \times Y$, so the restriction of ω to A is continuous; see e.g. [28, Box 1.8]. Since n was arbitrary, this implies that ω is continuous. Since ω is strictly increasing, unbounded, and continuous, it is invertible and its inverse $\omega^{-1} : [0, +\infty) \to [0, +\infty)$ is strictly increasing and continuous. In particular, since $\omega(0) = 0$, ω^{-1} is continuous at 0 and $\lim_{\delta \to 0} \omega^{-1}(\delta) = 0$.

Let $y \in Y$ and $(s_i, \sigma_i) \in B \times U$ for $i \in \{1, 2\}$ and denote $q(s_i, y, \sigma_i)$ by q_i. Then $q_i \leq P$ for $i \in \{1, 2\}$. Since c is locally Lipschitz, there exists a Lipschitz constant $L > 0$ for c on $(B \times [0, P]) \times Y$. Therefore, by construction,

$$\begin{aligned} \omega(|q_2 - q_1|) &\leq |c((s_2, q_2), y) - c((s_2, q_1), y)| \\ &\leq |c((s_2, q_2), y) - c((s_1, q_1), y)| + |c((s_1, q_1), y) - c((s_2, q_1), y)| \\ &\leq |\sigma_2 - \sigma_1| + L\|s_2 - s_1\|_{\mathbb{R}^{d-1}} \\ &\leq (1 + L)\|(s_2, \sigma_2) - (s_1, \sigma_1)\|_{\mathbb{R}^d}. \end{aligned} \tag{28}$$

Since ω is invertible and its inverse is strictly increasing, applying ω^{-1} to both sides of (28) gives

$$|q(s_2, y, \sigma_2) - q(s_1, y, \sigma_1)| = |q_2 - q_1| \leq \omega^{-1}\left((1 + L)\|(s_2, \sigma_2) - (s_1, \sigma_1)\|_{\mathbb{R}^d}\right). \tag{29}$$

Define the function $\tilde{\omega} : [0, +\infty) \to [0, +\infty)$ by $\tilde{\omega}(\delta) = \omega^{-1}((1+L)\delta)$. Then $\lim_{\delta \to 0} \tilde{\omega}(\delta) = 0$, and by (29), $\tilde{\omega}$ is a modulus of continuity for every function in $\mathcal{Q}$, which implies that $\mathcal{Q}$ is equicontinuous. □

We now define the candidate minimiser $\bar{p}_\psi$ of the inner-dual functional H_ψ.

Lemma 3.3. *For each $\psi \in \mathcal{C}(Y)$, there exists a unique constant $\tau_\psi \in \mathbb{R}$ such that the function*

$$\begin{aligned} &\bar{p}_\psi : B \to \mathbb{R} \\ &\bar{p}_\psi(s) = \max_{y \in Y} \{q(s, y, \psi(y) + \tau_\psi)\} \end{aligned} \tag{30}$$

is a probability density, where q is defined by (25).

Proof. Let $\psi \in \mathcal{C}(Y)$, and let τ_0, $\tau_1 \in \mathbb{R}$ satisfy $\tau_0 \leq \tau_1$. Since Y is compact and ψ is continuous, there exists a compact set $U \subset \mathbb{R}$ such that $\psi(y) + \tau \in U$ for all $(y, \tau) \in Y \times [\tau_0, \tau_1]$. By Lemma 3.2, the set

$$\mathcal{Q} \coloneqq \{q(\cdot, y, \cdot)|_{B \times U}\}_{y \in Y}$$

is uniformly bounded and equicontinuous. This implies that the functions defined for each $y \in Y$ by

$$\begin{aligned} &q_{\psi,y} : B \times [\tau_0, \tau_1] \to [0, +\infty), \\ &q_{\psi,y}(s, \tau) = q(s, y, \psi(y) + \tau), \end{aligned}$$

are uniformly bounded and equicontinuous. The function

$$\begin{aligned} &q_\psi : B \times [\tau_0, \tau_1] \to \mathbb{R}, \\ &q_\psi(s, \tau) = \max_{y \in Y} \{q(s, y, \psi(y) + \tau)\}, \end{aligned}$$

is therefore well defined, continuous, and bounded; see for example [28, Box 1.8]. It follows by the Dominated Convergence Theorem that the function $\mathcal{I} : \mathbb{R} \to \mathbb{R}$ defined by

$$\mathcal{I}(\tau) \coloneqq \int_B q_\psi(s, \tau)\, \mathrm{d}s$$

is continuous.

We now find τ_0, $\tau_1 \in \mathbb{R}$ such that $\mathcal{I}(\tau_0) = 0$ and $\mathcal{I}(\tau_1) \geq 1$, and we show that $\mathcal{I}$ is strictly increasing wherever it is positive. Then by the Intermediate Value Theorem, there exists a unique constant τ_ψ such that $\mathcal{I}(\tau_\psi) = 1$, which, by non-negativity of q, is therefore the unique constant such that $\bar{p}_\psi = q_\psi(\cdot, \tau_\psi)$ is a probability density.

Let $r = 1/\mathcal{L}^{d-1}(B)$ and define constants

$$m \coloneqq \min_{(s,y)\in B\times Y} c((s,0),y), \qquad M \coloneqq \max_{(s,y)\in B\times Y} c((s,r),y), \tag{31}$$

and

$$\tau_0 \coloneqq m - \max_{y\in Y} \psi(y), \qquad \tau_1 \coloneqq M - \max_{y\in Y} \psi(y).$$

By construction

$$\psi(y) + \tau_0 \leq c((s,0),y) \qquad \forall\, s \in B,\, y \in Y,$$

so, by definition of q,

$$q_\psi(s,\tau_0) = \max_{y\in Y}\{q(s,y,\psi(y)+\tau_0)\} = 0 \qquad \forall\, s \in B.$$

Therefore, $\mathcal{I}(\tau_0) = 0$. Since $c((s,\cdot),y)$ is strictly increasing for every $s \in B$ and $y \in Y$, and by definition of M and τ_1, for $y \in Y$ maximising ψ,

$$M = \psi(y) + \tau_1 \geq c((s,r),y) \geq c((s,0),y) \qquad \forall\, s \in B.$$

Since $c\big((s,\cdot),y\big)^{-1}$ is increasing, this implies that

$$q(s,y,\psi(y)+\tau_1) = c\big((s,\cdot),y\big)^{-1}(\psi(y)+\tau_1) \geq r \qquad \forall\, s \in B.$$

Then, by definition of r,

$$\mathcal{I}(\tau_1) \geq \int_B q(s,y,\psi(y)+\tau_1)\,\mathrm{d}s \geq \int_B r\,\mathrm{d}s = 1.$$

So $\mathcal{I}(\tau_0) = 0 \leq 1 \leq \mathcal{I}(\tau_1)$, as required.

Let $\tau_\psi \in \mathbb{R}$ satisfy $\mathcal{I}(\tau_\psi) = 1$. We claim that this uniquely defines τ_ψ. By the Intermediate Value Theorem, there exists $\tau \in \mathbb{R}$ such that $0 < \mathcal{I}(\tau) < 1$, so it is sufficient to show that $\mathcal{I}$ is strictly increasing on $[\tau, +\infty)$. By definition of $\mathcal{I}$ and non-negativity of q, there exists a set $A \subseteq B$ with positive $\mathcal{L}^{d-1}$ measure such that

$$0 < q_\psi(s,\tau) \qquad \forall\, s \in A.$$

Let $s \in A$ and take $y \in Y$ to be a maximiser of $q(s,\cdot,\psi(y)+\tau)$ over Y. Then by definition of q, $c((s,0),y) < \psi(y) + \tau$ and

$$q_\psi(s,\tau) = q(s,y,\psi(y)+\tau) = c((s,\cdot),y)^{-1}(\psi(y)+\tau).$$

Since $c\big((s,\cdot),y\big)^{-1}$ is strictly increasing, for any $\delta > 0$

$$q_\psi(s,\tau) = c((s,\cdot),y)^{-1}(\psi(y)+\tau) < c((s,\cdot),y)^{-1}(\psi(y)+\tau+\delta) \leq q_\psi(s,\tau+\delta).$$

It follows that $\mathcal{I}$ is strictly increasing on $[\tau, +\infty)$, as required. □

We now prove that $\bar{p}_\psi$ is the unique minimiser of H_ψ.

Theorem 3.4. *For $\psi \in \mathcal{C}(Y)$, $\bar{p}_\psi$ defined by (30) is the unique minimiser of H_ψ over $\mathcal{P}_{\mathrm{ac}}(B)$. In particular, $\bar{p} = \bar{p}_\psi$ satisfies (24) with $\tau = \tau_\psi$.*

Proof. By Lemma 3.3, $\bar{p}_\psi$ is a well-defined probability density. By construction, $\bar{p} = \bar{p}_\psi$ satisfies (24) with $\tau = \tau_\psi$. We now verify that $\bar{p}_\psi$ is the unique minimiser of H_ψ over $\mathcal{P}(B)$. Since $f_\psi(s,\cdot)$ is strictly convex on $[0,+\infty)$ and its partial derivative with respect to u is $\psi^c(s,\cdot)$, for any distinct $\tilde{u}$, $u \in [0,+\infty)$,

$$f_\psi(s,\tilde{u}) > f_\psi(s,u) + \psi^c(s,u)(\tilde{u}-u). \tag{32}$$

Let $\bar{p}_0 \in \mathcal{P}_{\mathrm{ac}}(B)$. Suppressing the argument s, (24) and (32) give

$$\begin{aligned}
H_\psi(\bar{p}_0) &= \int_B f_\psi(\cdot,\bar{p}_0)\,\mathrm{d}s \\
&> \int_B \left[f_\psi(\cdot,\bar{p}) + \psi^c(\cdot,\bar{p})(\bar{p}_0-\bar{p})\right]\,\mathrm{d}s \\
&= \int_B f_\psi(\cdot,\bar{p})\,\mathrm{d}s + \tau\int_{B\cap\{\bar{p}>0\}}(\bar{p}_0-\bar{p})\,\mathrm{d}s + \int_{B\cap\{\bar{p}=0\}}\psi^c(\cdot,\bar{p})\bar{p}_0\,\mathrm{d}s \\
&\geq \int_B f_\psi(\cdot,\bar{p})\,\mathrm{d}s + \tau\int_{B\cap\{\bar{p}>0\}}(\bar{p}_0-\bar{p})\,\mathrm{d}s + \int_{B\cap\{\bar{p}=0\}}\tau\bar{p}_0\,\mathrm{d}s \\
&= \int_B f_\psi(\cdot,\bar{p})\,\mathrm{d}s + \tau\int_B(\bar{p}_0-\bar{p})\,\mathrm{d}s \\
&= \int_B f_\psi(\cdot,\bar{p})\,\mathrm{d}s \\
&= H_\psi(\bar{p}).
\end{aligned}$$

Here, the final inequality holds because $\bar{p}_0$ is non-negative, and the penultimate equality holds because $\bar{p}$ and $\bar{p}_0$ are probability densities. □

We conclude this subsection by proving that the set of all functions $\bar{p}_\psi$ for $\psi \in \mathcal{C}(Y)$ is uniformly bounded and equicontinuous. This result will be used to restrict the dual functional to a subset of $\mathcal{C}(Y)$ that has compact closure, which allows us to prove the existence of maximisers of the dual functional; see Lemma 3.6 and Theorem 3.9. It will also be used in Section 4 to prove stability of minimisers of Problem 2.4.

Lemma 3.5. *The set*

$$\mathcal{S} \coloneqq \{\bar{p}_\psi \mid \psi \in \mathcal{C}(Y)\}$$

is uniformly bounded and equicontinuous.

Proof. Let m and M be defined as in (31). We show in the proof of Lemma 3.3 that

$$\tau_\psi \leq M - \max_{y\in Y}\psi(y)$$

This implies the uniform bound

$$\psi(y) + \tau_\psi \leq M \qquad \forall\, y \in Y\ \forall\, \psi \in \mathcal{C}(Y).$$

Moreover, by definition of q,

$$q(\cdot,y,\sigma) = 0 \qquad \forall\, y \in Y\ \forall\, \sigma \leq m.$$

Therefore

$$\mathcal{S}_0 \coloneqq \{q(\cdot, y, \psi(y) + \tau_\psi) \,|\, y \in Y,\, \psi \in \mathcal{C}(Y)\} \subseteq \{q(\cdot, y, \sigma) \,|\, y \in Y,\, \sigma \in [m, M]\} \cup \{0\},$$

which is uniformly bounded and equicontinuous by Lemma 3.2. Since each $\bar{p}_\psi$ is defined as the pointwise maximum over $y \in Y$ of functions in $\mathcal{S}_0$, the set $\mathcal{S}$ is also uniformly bounded and equicontinuous with the same bound and modulus of continuity; see e.g. [28, Box 1.8]. □

3.2 Maximisers of the dual functional

We now show that maximisers of the dual functional G_ν (22) exist and satisfy a necessary and sufficient first-order optimality condition; see Theorem 3.9. For brevity, we drop the subcript ν. Recall that

$$G(\psi) = \inf_{\bar{p} \in \mathcal{P}_{\mathrm{ac}}(B)} K(\,\cdot\,; \mu_{\bar{p}}, \nu),$$

where $K(\,\cdot\,; \mu_{\bar{p}}, \nu)$ is the Kantorovich functional for source and target measures $\mu_{\bar{p}}$ and ν, respectively. Each Kantorovich functional is concave and upper semi-continuous. As such, G is concave and upper-semi continuous. The existence of a maximiser of G_ν then follows by restriction to the set $A(Y)$ defined in Lemma 3.6, which we show to have compact closure. To establish the optimality condition, we apply [32, Theorem 2.4.18] to obtain an expression for the superdifferential of G_ν. We see that ψ is a maximiser of G_ν if and only if it is a Kantorovich potential from $\mu_{\bar{p}}$ to ν with $\bar{p} = \bar{p}_\psi$.

Let $P > 0$ be a uniform bound for the set $\mathcal{S} = \{\bar{p}_\psi \,|\, \psi \in \mathcal{C}(Y)\}$, which exists by Lemma 3.5. For $\varphi \in \mathcal{C}(B \times [0, P])$, define $\varphi^{\bar{c}} : Y \to \mathbb{R}$ by

$$\varphi^{\bar{c}}(y) \coloneqq \min_{x \in B \times [0,P]} \{c(x, y) - \varphi(x)\}.$$

We call $\varphi^{\bar{c}}$ the $\bar{c}$-transform of φ. (The notation $\bar{c}$ is simply used to indicate that the minimum is taken over the source space $B \times [0, P]$, distinguishing the $\bar{c}$-transform from the the c-transform (10), which is defined by minimisation over the target space Y.) We say that a function $\psi \in \mathcal{C}(Y)$ is $\bar{c}$-concave if there exists $\varphi \in \mathcal{C}(B \times [0, P])$ such that $\psi = \varphi^{\bar{c}}$.

Lemma 3.6. *Define*

$$A(Y) \coloneqq \{\psi \in \mathcal{C}(Y) \mid \psi \text{ is } \bar{c}\text{-concave, and } \min_{y \in Y} \psi(y) = 0\}.$$

Then $A(Y)$ is non-empty and

$$\sup_{\psi \in \mathcal{C}(Y)} G(\psi) = \sup_{\psi \in A(Y)} G(\psi). \tag{33}$$

Proof. Let $\psi \in \mathcal{C}(Y)$. Since $B \times [0, P]$ is compact and c is locally Lipschitz (Assumption 2.1.A.), the functions $c(\cdot, y) - \psi(y)$ for $y \in Y$ are Lipschitz on $B \times [0, P]$ and all have the same Lipschitz constant. As the pointwise minimum of these functions, ψ^c is Lipschitz on $B \times [0, P]$. Likewise, $\psi^{c\bar{c}} : Y \to \mathbb{R}$ defined by

$$\psi^{c\bar{c}}(y) \coloneqq \min_{x \in B \times [0,P]} \{c(x, y) - \psi^c(x)\}$$

is also Lipschitz. In particular, letting $\lambda = \min_{y \in Y} \psi^{c\bar{c}}(y)$,

$$\tilde{\psi} \coloneqq \psi^{c\bar{c}} - \lambda = (\psi + \lambda)^{c\bar{c}}$$

is continuous, $\bar{c}$-concave, and has minimum value zero, so is an element of $A(Y)$. By [28, Proposition 1.34], $\psi^{c\bar{c}c} = \psi^c$ on $B \times [0, P]$ and $\psi^{c\bar{c}} \geq \psi$. Then, for any $\bar{p} \in \mathcal{P}_{\mathrm{ac}}(B)$ with continuous density such that $\|\bar{p}\|_{\mathcal{C}(B)} \leq P$, the support of $\mu_{\bar{p}}$ is contained in $B \times [0, P]$ and

$$\begin{aligned} K(\psi; \mu_{\bar{p}}, \nu) &= \int_X \psi^c \, \mathrm{d}\mu_{\bar{p}} + \int_Y \psi \, \mathrm{d}\nu \\ &\leq \int_X \psi^{c\bar{c}c} \, \mathrm{d}\mu_{\bar{p}} + \int_Y \psi^{c\bar{c}} \, \mathrm{d}\nu. \\ &= \int_X \left(\psi^{c\bar{c}c} + \lambda\right) \, \mathrm{d}\mu_{\bar{p}} + \int_Y \left(\psi^{c\bar{c}} - \lambda\right) \, \mathrm{d}\nu \\ &= \int_X \tilde{\psi}^c \, \mathrm{d}\mu_{\bar{p}} + \int_Y \tilde{\psi} \, \mathrm{d}\nu \\ &= K(\tilde{\psi}; \mu_{\bar{p}}, \nu). \end{aligned} \tag{34}$$

Recall that by definition $K(\psi; \mu_{\bar{p}}, \nu) = H_\psi(\bar{p})$, similarly for $\tilde{\psi}$. By Theorem 3.4, each of the functionals H_ψ and $H_{\tilde{\psi}}$ has a unique minimiser over $\mathcal{P}_{\mathrm{ac}}(B)$, and by Lemma 3.5 this minimiser is continuous and bounded by P. Taking the minimum over $\bar{p}$ in (34) therefore gives

$$G(\psi) = \min_{\bar{p} \in \mathcal{P}_{\mathrm{ac}}(B)} K(\psi; \mu_{\bar{p}}, \nu) \leq \min_{\bar{p} \in \mathcal{P}_{\mathrm{ac}}(B)} K(\tilde{\psi}; \mu_{\bar{p}}, \nu) = G(\tilde{\psi}).$$

So, for any $\psi \in \mathcal{C}(Y)$ we have found $\tilde{\psi} \in A(Y)$ such that $G(\psi) \leq G(\tilde{\psi})$, which establishes (33). □

In order to derive optimality conditions using [32, Theorem 2.4.18], we now show that the map $\bar{p} \mapsto K(\psi, \mu_{\bar{p}}, \nu) = H_\psi(\bar{p})$ is lower semi-continuous for every $\psi \in \mathcal{C}(Y)$ by applying the following special case of [1, Theorem 3.3].

Proposition 3.7 ([1, Theorem 3.3])**.** *Let $f : B \times \mathbb{R} \to [0, +\infty]$ be such that, for every $s \in B$, $f(s, \cdot)$ is proper, convex, lower semi-continuous, and equal to $+\infty$ on $(-\infty, 0)$. Suppose also that $\lim_{t \to +\infty} \frac{f(\cdot,t)}{t} = +\infty$ uniformly on B. Then the functional defined on $\mathcal{P}(B)$ by*

$$\bar{p} \mapsto \begin{cases} \int_B f(s, \bar{p}(s)) \, \mathrm{d}s & \text{if } \bar{p} \in \mathcal{P}_{\mathrm{ac}}(B) \\ +\infty & \text{otherwise,} \end{cases}$$

is lower semi-continuous with respect to the weak- topology on $\mathcal{P}(B)$.*

Lemma 3.8. *For any $\nu \in \mathcal{P}(Y)$ and $\psi \in \mathcal{C}(Y)$ the functional H_ψ is lower semi-continuous with respect to the weak-* topology on $\mathcal{P}(B)$.*

Proof. It is sufficient to show that f_ψ satisfies the hypotheses of Proposition 3.7. For every $s \in B$, $f_\psi(s, \cdot)$ is proper, convex, and lower semi-continuous. Since $H_\psi(\bar{p}) = H_{\psi+\lambda}(\bar{p})$ for all $\lambda \in \mathbb{R}$ and Y is compact, we assume without loss of generality that $\psi(y) < 0$ for all $y \in Y$. Then, since c non-negative and ψ is negative, f_ψ is non-negative and equal to $+\infty$ on the interval $(-\infty, 0)$.

It remains to show that $\lim_{t \to +\infty} \frac{f_\psi(\cdot,t)}{t} = +\infty$ uniformly on B. That is

$$\forall\, M > 0,\ \exists\, T > 0 \ \text{ such that } \ t \geq T \implies \frac{f_\psi(x, t)}{t} \geq M \quad \forall\, s \in B.$$

Let $M > 0$ and let $r \in \mathbb{R}$ satisfy

$$r > \max_{(s,y)\in B\times Y} q(s, y, M + \psi(y)) \geq 0.$$

This maximum is well-defined and non-negative because q is non-negative and continuous by Lemma 3.2 and $B \times Y$ is compact. By construction and Assumption 2.1.B.,

$$c((s, r), y) > M + \psi(y) \quad \forall\, s \in B,\, y \in Y,$$

which implies that

$$-M + \min_{s\in B} \psi^c(s, r) = -M + \min_{s\in B} \min_{y\in Y} \{c((s, r), y) - \psi(y)\} > 0.$$

Define the positive constant

$$T := \frac{r \max_{s\in B} \psi^c(s, r)}{-M + \min_{s\in B} \psi^c(s, r)}.$$

By (32), for any $\tilde{s} \in B$ and $t \geq T$,

$$\frac{f_\psi(\tilde{s}, t)}{t} \geq \frac{f_\psi(\tilde{s}, r) + (t - r)\dfrac{\partial f_\psi}{\partial t}(\tilde{s}, r)}{t} = \frac{f_\psi(\tilde{s}, r) - r\psi^c(\tilde{s}, r)}{t} + \psi^c(\tilde{s}, r).$$

By definition of f_ψ and positivity of ψ^c,

$$f_\psi(\tilde{s}, r) - r\psi^c(\tilde{s}, r) = \int_0^r \left(\psi^c(\tilde{s}, \tilde{r}) - \psi^c(\tilde{s}, r)\right) \mathrm{d}\tilde{r} > -r \max_{s\in B} \psi^c(s, r)$$

It follows that

$$\frac{f_\psi(\tilde{s}, t)}{t} \geq \frac{-r \max_{s\in B} \psi^c(s, r)}{T} + \min_{s\in B} \psi^c(s, t) = M,$$

as required. Applying Proposition 3.7 with $f = f_\psi$ completes the proof. □

We now state and prove the main result of this section.

Theorem 3.9. *There exists a maximiser of G. In particular, ψ maximises G if and only if ψ is a Kantorovich potential from $\mu_{\bar{p}}$ to ν with $\bar{p} = \bar{p}_\psi$.*

Proof. We first prove existence of a maximiser of G. By Lemma 3.6, it is sufficient to prove that the closure of $A(Y)$ is a compact subset of $\mathcal{C}(Y)$, and that G is proper, concave, upper semi-continuous, and bounded above on $A(Y)$.

Let $\psi \in A(Y)$. Then there exists $\varphi \in \mathcal{C}(B \times [0, P])$ such that $\psi = \varphi^{\overline{c}}$. By Assumption 2.1.A., the cost c is Lipschitz on the compact set $(B \times [0, P]) \times Y$ with some Lipschitz constant $L > 0$. As the pointwise minimum over a family of L-Lipschitz functions, ψ is L-Lipschitz on Y. Then, since $\min_{y\in Y} \psi(y) = 0$,

$$|\psi(y)| \leq L\mathrm{diam}(Y) \qquad \forall\, y \in Y, \tag{35}$$

which is finite since Y is compact. As such, $A(Y)$ is uniformly bounded and equicontinuous, so by the Arzela-Ascoli Theorem, its closure is compact in $\mathcal{C}(Y)$ with respect to the uniform topology.

Next we show that G is concave and upper semi-continuous with respect to the uniform topology. Let $\bar{p} \in \mathcal{P}_{\mathrm{ac}}(B)$ and let ψ_0, $\psi_1 \in \mathcal{C}(Y)$. Let $x \in X$, and let $y_i \in Y$ be such that

$$\psi_i^c(x) = c(x, y_i) - \psi_i(y_i)$$

for each $i \in \{0, 1\}$. Then, using the definition of the c-transform,

$$\begin{aligned}
\psi_1^c(x) - \psi_0^c(x) &= \min_{y \in Y}\{c(x,y) - \psi_1(y)\} - (c(x, y_0) - \psi_0(y_0)) \\
&\leq (c(x, y_0) - \psi_1(y_0)) - (c(x, y_0) - \psi_0(y_0)) \\
&= \psi_1(y_0) - \psi_0(y_0) \\
&\leq \|\psi_1 - \psi_0\|_{\mathcal{C}(Y)}.
\end{aligned}$$

By symmetry, it follows that

$$|\psi_1^c(x) - \psi_0^c(x)| = \max\left\{\psi_1^c(x) - \psi_0^c(x),\, \psi_0^c(x) - \psi_1^c(x)\right\} \leq \|\psi_1 - \psi_0\|_{\mathcal{C}(Y)}.$$

Since $x \in X$ was arbitrary,

$$|K(\psi_1; \mu_{\bar{p}}, \nu) - K(\psi_0; \mu_{\bar{p}}, \nu)| = \left| \int_X (\psi_1^c - \psi_0^c)\, \mathrm{d}\mu_{\bar{p}} + \int_Y (\psi_1 - \psi_0)\, \mathrm{d}\nu \right| \leq 2\, \|\psi_1 - \psi_0\|_{\mathcal{C}(Y)}\,.$$

So $K(\,\cdot\,; \mu_{\bar{p}}, \nu)$ is 2-Lipschitz on $\mathcal{C}(Y)$. In particular, it is upper semi-continuous. Moreover, for $\lambda \in (0, 1)$ and $x \in X$,

$$\begin{aligned}
((1-\lambda)\psi_1 + \lambda\psi_0)^c(x) &= \min_{y \in Y}\left\{c(x,y) - (1-\lambda)\psi_1(y) - \lambda\psi_0(y)\right\} \\
&\geq (1-\lambda) \min_{y \in Y}\left\{c(x,y) - \psi_1(y)\right\} - \lambda \min_{y \in Y}\left\{c(x,y) - \psi_0(y)\right\} \\
&= (1-\lambda)\psi_1^c(x) + \lambda\psi_0^c(x).
\end{aligned}$$

Therefore,

$$\begin{aligned}
K((1-\lambda)\psi_1 + \lambda\psi_0; \mu_{\bar{p}}, \nu) &= \int_X ((1-\lambda)\psi_1 + \lambda\psi_0)^c\, \mathrm{d}\mu_{\bar{p}} + \int_Y (1-\lambda)\psi_1 + \lambda\psi_0\, \mathrm{d}\nu \\
&\geq (1-\lambda)\left(\int_X \psi_1^c\, \mathrm{d}\mu_{\bar{p}} + \int_Y \psi_1\, \mathrm{d}\nu\right) + \lambda\left(\int_X \psi_0^c\, \mathrm{d}\mu_{\bar{p}} + \int_Y \psi_0\, \mathrm{d}\nu\right) \\
&= (1-\lambda)K(\psi_1; \mu_{\bar{p}}, \nu) + \lambda K(\psi_0; \mu_{\bar{p}}, \nu),
\end{aligned}$$

so $K(\,\cdot\,; \mu_{\bar{p}}, \nu)$ is concave. As the pointwise minimum of a family of concave 2-Lipschitz functions, G is concave and 2-Lipschitz. In particular, it is upper semi-continuous.

We now show that G is real valued (in particular proper) and bounded above on $A(Y)$. Indeed, let $\psi \in A(Y)$ and let $\bar{p} = \bar{p}_\psi$. By Theorem 3.4, $\bar{p}_\psi$ minimises H_ψ, so by definition of G,

$$G(\psi) = K(\psi; \mu_{\bar{p}}, \nu). \tag{36}$$

By Lemma 3.5, $\|\bar{p}\|_{\mathcal{C}(B)} \leq P$, where $P > 0$ is a constant independent of ψ. Since c is Lipschitz on the compact set $(B \times [0, P]) \times Y$, its maximum over that set, which we denote by $c_{\max}$, is attained. Using the bound (35) and the fact that $\min_{y \in Y} \psi(y) = 0$,

$$\begin{aligned}
K(\psi; \mu_{\bar{p}}, \nu) &= \int_X \psi^c\, \mathrm{d}\mu_{\bar{p}} + \int_Y \psi\, \mathrm{d}\nu \\
&\leq \left(c_{\max} - \min_{y \in Y} \psi(y)\right) \int_B \bar{p}(s)\, \mathrm{d}s + L\mathrm{diam}(Y) \\
&= c_{\max} + L\mathrm{diam}(Y),
\end{aligned}$$

which is an upper bound for G on $A(Y)$ by (36). In particular, since $\bar{p}$ is bounded and ψ and ψ^c are continuous, $G(\psi)$ is real valued so G is proper. Since this holds for any $\psi \in \mathcal{C}(Y)$, G is real valued. This concludes the proof that a maximiser of G over $\mathcal{C}(Y)$ exists.

Finally we prove the necessary and sufficient optimality condition. Since G is concave and real valued, for every $\psi \in \mathcal{C}(Y)$ the superdifferential of G at ψ, denoted by $\partial^+ G(\psi)$, is non-empty and contains 0 if and only if ψ is a maximiser of G. By Lemma 3.8, for every $\psi \in \mathcal{C}(Y)$ the functional H_ψ is lower semi-continuous with respect to the weak-* topology on $\mathcal{P}(B)$. By definition

$$G(\psi) = \inf_{\bar{p} \in \mathcal{P}(Y)} H_\psi(\bar{p}).$$

Let $\psi \in \mathcal{C}(Y)$ and let $\bar{p} = \bar{p}_\psi$ be the unique minimiser of H_ψ, which exists by Theorem 3.4. Then by [32, Theorem 2.4.18],

$$\partial^+ G(\psi) = \partial^+ \left[K(\,\cdot\,; \mu_{\bar{p}}, \nu) \right](\psi).$$

Hence, ψ is a maximiser of G if and only if $0 \in \partial^+ \left[K(\,\cdot\,; \mu_{\bar{p}}, \nu) \right](\psi)$, which holds if and only if ψ is a Kantorovich potential from $\mu_{\bar{p}}$ to ν. □

3.3 Minimisers of the primal functional

We now combine the results of this section to prove that the primal problem (Problem 2.4) has a unique solution, to prove that duality holds between the primal and dual problems, and to derive an optimality condition relating both problems.

Theorem 3.10. *There exists a unique minimiser of F. In particular,*

$$\min_{\bar{p} \in \mathcal{P}_{\mathrm{ac}}(B)} F(\bar{p}) = \max_{\psi \in \mathcal{C}(Y)} G(\psi), \tag{37}$$

and $\bar{p}$ and ψ are optimal in their respective problems if and only if $p = p_\psi$ and ψ is a Kantorovich potential from $\mu_{\bar{p}}$ to ν.

Proof. First we prove duality and existence of a minimiser of F (37). Let $\psi \in \mathcal{C}(Y)$ be a maximiser of G, which exists by Theorem 3.9. By Theorem 3.4, $\bar{p} = \bar{p}_\psi$ is the unique minimiser of H_ψ, and the optimality condition stated in Theorem 3.9 implies that ψ is a Kantorovich potential from $\mu_{\bar{p}}$ to ν. Hence, by the Kantorovich Duality Theorem,

$$G(\psi) = \inf_{\mathcal{P}_{\mathrm{ac}}(B)} H_{\nu,\psi} = K(\psi; \mu_{\bar{p}}, \nu) = \max_{\mathcal{C}(Y)} K(\,\cdot\,; \mu_{\bar{p}}, \nu) = \mathcal{T}_c(\mu_{\bar{p}}, \nu) = F(\bar{p}).$$

Then, using weak duality (21),

$$\max_{\mathcal{C}(Y)} G = G(\psi) = F(\bar{p}) \geq \inf_{\mathcal{P}_{\mathrm{ac}}(B)} F \geq \sup_{\mathcal{C}(Y)} G, \tag{38}$$

So all inequalities in (38) must be equalities, meaning that $\bar{p}$ is a minimiser of the primal functional F over $\mathcal{P}_{\mathrm{ac}}(B)$. This establishes the existence of a minimiser of F over $\mathcal{P}_{\mathrm{ac}}(B)$, and proves that (37) holds.

By Theorem 3.4, $\operatorname*{arg\,min}_{\mathcal{P}_{\mathrm{ac}}(B)} H_\psi = \{\bar{p}_\psi\}$ is a singleton for every $\psi \in \mathcal{C}(Y)$. To prove that the minimiser of F is unique, it is therefore sufficient to prove that

$$\operatorname*{arg\,min}_{\mathcal{P}_{\mathrm{ac}}(B)} F = \operatorname*{arg\,min}_{\mathcal{P}_{\mathrm{ac}}(B)} H_\psi \qquad \forall\, \psi \in \operatorname*{arg\,max}_{\mathcal{C}(Y)} G. \tag{39}$$

Let ψ be a maximiser of G and let $\bar{p} = \bar{p}_\psi$ be the unique minimiser of H_ψ. By Theorem 3.9, ψ is a Kantorovich potential from $\mu_{\bar{p}}$ to ν. Then by (37) and the Kantorovich Duality Theorem,

$$\min_{\mathcal{P}_{\mathrm{ac}}(B)} F = \max_{\mathcal{C}(Y)} G = G(\psi) = \min_{\mathcal{P}_{\mathrm{ac}}(B)} H_\psi = K(\psi; \mu_{\bar{p}}, \nu) = \mathcal{T}_c(\mu_{\bar{p}}, \nu) = F(\bar{p}).$$

Hence

$$\arg\min_{\mathcal{P}_{\mathrm{ac}}(B)} H_\psi \subseteq \arg\min_{\mathcal{P}_{\mathrm{ac}}(B)} F. \tag{40}$$

For a contradiction, suppose that there exists $\bar{p} \in \arg\min_{\mathcal{P}_{\mathrm{ac}}(B)} F \setminus \arg\min_{\mathcal{P}_{\mathrm{ac}}(B)} H_\psi$. By the Kantorovich Duality Theorem,

$$\min_{\mathcal{P}_{\mathrm{ac}}(B)} F = F(\bar{p}) = \mathcal{T}_c(\mu_{\bar{p}}, \nu) \geq K(\psi; \mu_{\bar{p}}, \nu) = H_\psi(\bar{p}). \tag{41}$$

By assumption,

$$H_\psi(\bar{p}) > \min_{\mathcal{P}_{\mathrm{ac}}(B)} H_\psi = G_\nu(\psi) = \max_{\mathcal{C}(Y)} G. \tag{42}$$

Together, (41) and (42) imply that

$$\min_{\mathcal{P}_{\mathrm{ac}}(B)} F > \max_{\mathcal{C}(Y)} G,$$

which contradicts (37). Hence, it must hold that

$$\arg\min_{\mathcal{P}_{\mathrm{ac}}(B)} H_\psi \supseteq \arg\min_{\mathcal{P}_{\mathrm{ac}}(B)} F. \tag{43}$$

Together, (40) and (43) establish (39), as required.

We conclude the proof by establishing the necessary and sufficient optimality conditions. By (39), $\bar{p}$ and ψ are optimal in their respective problems if and only if

$$\bar{p} = \bar{p}_\psi \qquad \text{and} \qquad \psi \in \arg\max_{\mathcal{C}(Y)} G.$$

By Theorem 3.9, this holds if and only if ψ is a Kantorovich potential from $\mu_{\bar{p}}$ to ν, as required. □

4 Stability

In this section, we establish stability of Problem 2.4 with respect to the target measure ν. We work in the setting described in Section 2 and cost functions $c : X \times Y \to [0, +\infty)$ satisfying only Assumption 2.1, unless otherwise specified. Using the stability of the optimal transport cost $\mathcal{T}_c$ with respect to source and target measures [31, Theorem 5.20], we show that for ν_n converging to ν in $\mathcal{P}(Y)$, the sequence of minimisers $\bar{p}_n$ of the functionals F_{ν_n} converges uniformly to the minimiser $\bar{p}$ of F_ν (Theorem 4.1). If the cost c additionally satisfies the twist condition (Assumption 2.5) then there exists a corresponding sequence of optimal transport maps T_n between $\mu_{\bar{p}_n}$ and ν_n, and we show that this sequence converges to the optimal transport map T between $\mu_{\bar{p}}$ and ν (Theorem 4.2).

Theorem 4.1. *The map* $F : \big(\mathcal{P}_{\mathrm{ac}}(B) \cap L^\infty(B), \|\cdot\|_{L^\infty(B)}\big) \times \mathcal{P}(Y) \to \mathbb{R}$ *given by*

$$F(\bar{p}, \nu) \coloneqq F_\nu(\bar{p}) = \mathcal{T}_c(\mu_{\bar{p}}, \nu)$$

is continuous, where $\mathcal{P}(Y)$ is equipped with the weak topology. In particular, suppose that $\nu_n \rightharpoonup \nu$ in $\mathcal{P}(Y)$, and let $\bar{p}_n$ (respectively $\bar{p}$) be the unique minimiser of F_{ν_n} (respectively F_ν) over $\mathcal{P}_{\mathrm{ac}}(B)$ for each $n \in \mathbb{N}$. Then $F_{\nu_n} \xrightarrow{\Gamma} F_\nu$ on $\big(\mathcal{P}_{\mathrm{ac}}(B) \cap L^\infty(B), \|\cdot\|_{L^\infty(B)}\big)$, $\bar{p}_n \to \bar{p}$ in $\mathcal{C}(B)$, and $\mu_{\bar{p}_n} \to \mu_{\bar{p}}$ strongly in $L^1(X)$. Moreover, if γ_n (respectively γ) is an optimal transport plan from $\mu_{\bar{p}_n}$ to ν_n (respectively from $\mu_{\bar{p}}$ to ν) for each $n \in \mathbb{N}$, then up to a subsequence $\gamma_n \rightharpoonup \gamma$ in $\mathcal{P}(X \times Y)$.

Proof. We first show that F is continuous. Let $((\bar{p}_n, \nu_n))_{n\in\mathbb{N}}$ be a sequence converging in $\big(\mathcal{P}_{\mathrm{ac}}(B) \cap L^\infty(B), \|\cdot\|_{L^\infty(B)}\big) \times \mathcal{P}(Y)$ to a point $(\bar{p}, \nu)$. Since B is compact, $\|\bar{p}_n - \bar{p}\|_{L^1(B)} \to 0$, and by Remark 2.3, $\|\mu_{\bar{p}_n} - \mu_{\bar{p}}\|_{L^1(X)} \to 0$. This implies that

$$\mu_{\bar{p}_n} \rightharpoonup \mu_{\bar{p}}. \tag{44}$$

Since the sequence $(\bar{p}_n)_{n\in\mathbb{N}}$ converges in $L^\infty(B)$, it is uniformly bounded, so $\mu_{\bar{p}}$ and $\mu_{\bar{p}_n}$ have common compact support, say $A \subset X$. By continuity of c (Assumption 2.1.A.) and compactness of $A \times Y$,

$$\mathcal{T}_c(\mu_{\bar{p}}, \nu) < +\infty \quad \text{and} \quad \mathcal{T}_c(\mu_{\bar{p}_n}, \nu_n) < +\infty \quad \forall\, n \in \mathbb{N}. \tag{45}$$

For each $n \in \mathbb{N}$ let γ_n be an optimal transport plan from $\mu_{\bar{p}_n}$ to ν_n. Since $\mu_{\bar{p}_n} \rightharpoonup \mu_{\bar{p}}$ and $\nu_n \rightharpoonup \nu$ and the corresponding transport costs are finite (45), by [31, Theorem 5.20], there exists an optimal transport plan γ from $\mu_{\bar{p}}$ to ν such that, up to a subsequence, $\gamma_n \rightharpoonup \gamma$. For all $n \in \mathbb{N}$, the support of γ_n and γ is necessarily contained in $A \times Y$, on which c is a bounded continuous function, so, up to a subsequence,

$$F(\bar{p}_n, \nu_n) = \mathcal{T}_c(\mu_{\bar{p}_n}, \nu_n) = \int_{A\times Y} c \,\mathrm{d}\gamma_n \to \int_{A\times Y} c \,\mathrm{d}\gamma = \mathcal{T}_c(\mu_{\bar{p}}, \nu) = F(\bar{p}, \nu).$$

Since the limit is independent of the chosen subsequence, the whole sequence converges to this limit. This establishes the claimed continuity of F, which immediately implies the claimed Γ-convergence.

Finally, suppose that $\nu_n \to \nu$ in $\mathcal{P}(Y)$, let $\bar{p}_n$ be the unique minimiser of F_{ν_n} over $\mathcal{P}_{\mathrm{ac}}(B)$ for each $n \in \mathbb{N}$, and let $\bar{p}$ be the unique minimiser of F_ν over $\mathcal{P}_{\mathrm{ac}}(B)$. By Theorem 3.10, $(\bar{p}_n)_{n\in\mathbb{N}} \subset \mathcal{S}$, which is uniformly bounded and equicontinuous by Lemma 3.5. By the Ascoli-Arzela Theorem, any subsequence has a further subsequence that converges uniformly, and therefore with respect to the strong L^∞ topology. Since $F_{\nu_n} \xrightarrow{\Gamma} F_\nu$ on $(\mathcal{P}_{\mathrm{ac}}(B) \cap L^\infty(B), \|\cdot\|_{L^\infty(B)})$, the limit of this further subsequence must be equal to $\bar{p}$. Since the limit is independent of the choice of subsequence, the whole sequence converges to this limit. By Remark 2.3, $\mu_{\bar{p}_n} \to \mu_{\bar{p}}$ strongly in $L^1(X)$. Letting γ_n (respectively γ) by an optimal transport plan from $\mu_{\bar{p}_n}$ to ν_n (respectively from $\mu_{\bar{p}}$ to ν) for each $n \in \mathbb{N}$, applying [31, Theorem 5.20] as above gives that, up to a subsequence, $\gamma_n \rightharpoonup \gamma$ in $\mathcal{P}(X \times Y)$. □

Theorem 4.1 establishes stability of minimisers of Problem 2.4 with respect to ν. We now show that when the cost c is twisted (Assumption 2.5), if $\bar{p}$ is the minimiser of F_ν then there is a unique optimal transport map T from $\mu_{\bar{p}}$ to ν, and T is also stable with respect to ν.

Theorem 4.2 (Existence and stability of transport maps)**.** *Suppose that the cost $c : X \times Y \to \mathbb{R}$ satisfies Assumption 2.5. Then for any $\bar{p} \in \mathcal{P}_{\mathrm{ac}}(B) \cap L^\infty(B)$ and $\nu \in \mathcal{P}(Y)$ there exists a unique optimal transport map T from $\mu_{\bar{p}}$ to ν, which is defined $\mu_{\bar{p}}$ almost everywhere. Moreover, suppose that $\nu_n \rightharpoonup \nu$ in $\mathcal{P}(Y)$, let $(\bar{p}_n)_{n\in\mathbb{N}}$ be the corresponding sequence of minimisers of Problem 2.4, and for each $n \in \mathbb{N}$ let T_n (respectively T) be the unique optimal transport map from $\mu_{\bar{p}_n}$ to ν_n (respectively $\mu_{\bar{p}}$ to ν) that is zero outside the support of $\mu_{\bar{p}_n}$ (respectively $\mu_{\bar{p}}$). Then*

$$\|T_n - T\|_{L^r(X)} \to 0 \qquad \forall\, r \in [1, +\infty).$$

Proof. For any $\bar{p} \in \mathcal{P}_{\mathrm{ac}}(B) \cap L^\infty(B)$ and $\nu \in \mathcal{P}(Y)$, both $\mu_{\bar{p}}$ and ν have compact support, and $\mu_{\bar{p}}$ is absolutely continuous with respect to $\mathcal{L}^d$. Since c is twisted (Assumption 2.5), by the Gangbo-McCann Theorem (see for example [17] or [25, Theorem 12]), there exists a unique optimal transport map T from $\mu_{\bar{p}}$ to ν, which is defined $\mu_{\bar{p}}$ almost everywhere. Moreover, the measure $\gamma \coloneqq (\mathrm{id}_X, T)_\#(\mu_{\bar{p}})$ is the unique optimal transport plan from $\mu_{\bar{p}}$ to ν.

Suppose that $\nu_n \rightharpoonup \nu$ in $\mathcal{P}(Y)$, let $(\bar{p}_n)_{n\in\mathbb{N}}$ be the corresponding sequence of minimisers of Problem 2.4, and for each $n \in \mathbb{N}$ let T_n (respectively T) be the unique optimal transport map from $\mu_{\bar{p}_n}$ to ν_n (respectively $\mu_{\bar{p}}$ to ν) extended by zero onto X. Define $\gamma_n \coloneqq (\mathrm{id}_X, T_n)_\#(\mu_{\bar{p}_n})$ and $\gamma \coloneqq (\mathrm{id}_X, T)_\#(\mu_{\bar{p}})$. Let $r \in [1, +\infty)$ and $\varepsilon \in (0, 1)$. We aim to find $N \in \mathbb{N}$ such that for $n \geq N$, $\|T - T_n\|^r_{L^r(X)} \leq \varepsilon$.

By Theorem 3.10, $\bar{p}_n$, $\bar{p} \in \mathcal{S}$ for all $n \in \mathbb{N}$, and by Lemma 3.5, $\mathcal{S}$ is equicontinuous and uniformly bounded. Let $P > 0$ be a uniform bound for $\mathcal{S}$. Define the compact set

$$X_P \coloneqq \{(s, p) \in X \,:\, p \leq P\},$$

and for each $n \in \mathbb{N}$ define

$$\Delta_{n,\varepsilon} \coloneqq \left\{x \in X_P \,:\, \|T(x) - T_n(x)\| \geq \frac{\varepsilon}{3}\right\}.$$

We first show that with

$$\delta = \frac{\varepsilon}{3\mathrm{diam}(Y)^r}$$

there exists $N_0 \in \mathbb{N}$ such that for $n \geq N_0$, $\mu_{\bar{p}}(\Delta_{n,\varepsilon}) \leq \delta$.

By Theorem 4.1, $\|\mu_{\bar{p}} - \mu_{\bar{p}_n}\|_{L^1(X)} \to 0$ and $\gamma_n \rightharpoonup \gamma$, where the convergence is along the whole sequence by uniqueness of γ. Hence, there exists $N \in \mathbb{N}$ such that for all $n \geq N$

$$\|\mu_{\bar{p}} - \mu_{\bar{p}_n}\|_{L^1(X)} \leq \frac{\delta}{4}. \tag{46}$$

By Lusin's Thoerem, there exists a compact set $A \subset X_P$ such that $T|_A$ is continuous and

$$\mu_{\bar{p}}(X \setminus A) \leq \frac{\delta}{4}. \tag{47}$$

In particular, the set

$$A_\varepsilon \coloneqq \left\{(x, y) \in A \times Y \,:\, \|T(x) - y\| \geq \frac{\varepsilon}{3}\right\}$$

is closed. Since γ is concentrated on the graph of T, $\gamma(A_\varepsilon) = 0$, so weak convergence of γ_n to γ and closedness of A_ε implies that there exists $N_1 \in \mathbb{N}$ such that for $n \geq N_1$

$$\gamma_n(A_\varepsilon) \leq \frac{\delta}{4}. \tag{48}$$

We have,

$$\begin{aligned}
\mu_{\bar{p}_n}(\Delta_{n,\varepsilon} \cap A) &= \mu_{\bar{p}_n}\left(\left\{x \in A \,:\, \|T(x) - T_n(x)\| \geq \frac{\varepsilon}{3}\right\}\right) \\
&= \gamma_n\left(\left\{x \in A \,:\, \|T(x) - T_n(x)\| \geq \frac{\varepsilon}{3}\right\} \times Y\right) \\
&= \gamma_n\left(\left\{(x, y) \in A \times Y \,:\, \|T(x) - y\| \geq \frac{\varepsilon}{3}\right\}\right) \\
&= \gamma_n(A_\varepsilon),
\end{aligned}$$

where the second equality holds because γ_n has first marginal $\mu_{\bar{p}_n}$, and the third equality holds because $\mathrm{spt}(\gamma_n) \subseteq \mathrm{graph}(T_n)$. For all $n \geq N_2 \coloneqq \max\{N_0, N_1\}$,

$$\begin{aligned}
\gamma_n(A_\varepsilon) = \mu_{\bar{p}_n}(\Delta_{n,\varepsilon} \cap A) &= \mu_{\bar{p}_n}(\Delta_{n,\varepsilon} \setminus (X_P \setminus A)) \\
&\geq \mu_{\bar{p}_n}(\Delta_{n,\varepsilon}) - \mu_{\bar{p}_n}(X_P \setminus A) \\
&= \mu_{\bar{p}_n}(\Delta_{n,\varepsilon}) + (\mu_{\bar{p}}(X_P \setminus A) - \mu_{\bar{p}_n}(X_P \setminus A)) - \mu_{\bar{p}}(X_P \setminus A) \\
&\geq \mu_{\bar{p}_n}(\Delta_{n,\varepsilon}) - \|\mu_{\bar{p}} - \mu_{\bar{p}_n}\|_{L^1(X)} - \mu_{\bar{p}}(X_P \setminus A).
\end{aligned}$$

Then by (46), (47), and (48),

$$\mu_{\bar{p}_n}(\Delta_{n,\varepsilon}) \leq \frac{3\delta}{4}. \tag{49}$$

Then by (46) and (49), for all $n \geq N_2$,

$$\mu_{\bar{p}}(\Delta_{n,\varepsilon}) = \mu_{\bar{p}}(\Delta_{n,\varepsilon}) - \mu_{\bar{p}_n}(\Delta_{n,\varepsilon}) + \mu_{\bar{p}_n}(\Delta_{n,\varepsilon}) \leq \|\mu_{\bar{p}} - \mu_{\bar{p}_n}\|_{L^1(X)} + \mu_{\bar{p}_n}(\Delta_{n,\varepsilon}) \leq \delta,$$

as claimed.

Since the optimal transport maps T_n and T map into Y

$$\|T - T_n\|_{L^\infty(X)} \leq \mathrm{diam}(Y).$$

By definition of $\Delta_{n,\varepsilon}$, and since $\mu_{\bar{p}}$ is a probability measure on X, it follows that for all $n \geq N_2$,

$$\begin{aligned}
\|T - T_n\|^r_{L^r(X,\mu_{\bar{p}})} &= \int_X \|T - T_n\|^r_{\mathbb{R}^d} \, \mathrm{d}\mu_{\bar{p}} \\
&= \int_{X \setminus \Delta_{n,\varepsilon}} \|T - T_n\|^r_{\mathbb{R}^d} \, \mathrm{d}\mu_{\bar{p}} + \int_{\Delta_{n,\varepsilon}} \|T - T_n\|^r_{\mathbb{R}^d} \, \mathrm{d}\mu_{\bar{p}} \\
&\leq \left(\frac{\varepsilon}{3}\right)^r + \mu_{\bar{p}}(\Delta_{n,\varepsilon}) \mathrm{diam}(Y)^r \\
&< \frac{\varepsilon}{3} + \delta \mathrm{diam}(Y)^r \\
&= \frac{2\varepsilon}{3}.
\end{aligned}$$

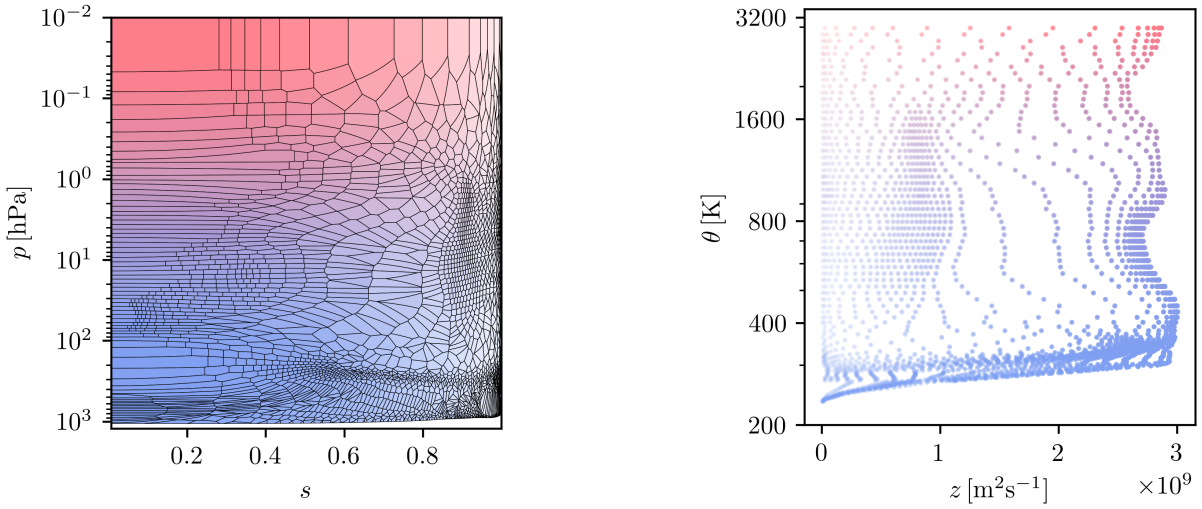
Finally, since Y is compact, there exists $R > 0$ such that $Y \subset B_R(0) \subset \mathbb{R}^d$. Uniform convergence of $\bar{p}_n$ to $\bar{p}$ implies that there exists $N_3 \in \mathbb{N}$ such that for all $n \geq N_3$

$$\|\bar{p} - \bar{p}_n\|_{L^1(B)} \leq \frac{\varepsilon}{3R^r}.$$

Let $N = \max\{N_2, N_3\}$. Since T is zero outside of $\mathrm{spt}(\mu_{\bar{p}})$, and T_n is zero outside of $\mathrm{spt}(\mu_{\bar{p}_n})$, for $n \geq N$

$$\begin{aligned}
\|T - T_n\|^r_{L^r(X)} &= \int_X \|T(x) - T_n(x)\|^r_{\mathbb{R}^d} \, \mathrm{d}x \\
&= \|T - T_n\|^r_{L^r(X,\mu_{\bar{p}})} + \int_X \|T_n(x)\|^r_{\mathbb{R}^d} \mathbb{1}_{\{\bar{p}(s) \leq p \leq \bar{p}_n(s)\}}(x) \, \mathrm{d}x \\
&\leq \frac{2\varepsilon}{3} + R^r \|\bar{p} - \bar{p}_n\|_{L^1(B)} \\
&= \varepsilon,
\end{aligned}$$

which completes the proof. □



(a) Optimal c_{ext}-Laguerre tessellation of X

(b) Discrete mass density ν on Y

Figure 3: The energy-minimising MLM state computed from ERA-Interim data [14] at 1800UTC on 22 January 2010. Right: the discrete mass density ν defined by (17) using the 3-D mass and circulation integrals (2)-(3) as computed in [27]. Left: the optimal c_{ext}-Laguerre tessellation computed using the damped Newton method of [20]. The pressure axis is reversed so that the fictitious cell L_{n+1}, which represents the ground, extends across the bottom of the domain. Its top boundary coincides with the graph of the energy-minimising surface pressure $\bar{p}$. Here, $p_{\min} = 0.01\,\text{hPa}$. The colour scheme is defined so that potential temperature increases from blue to red, and zonal angular momentum increases with opacity. We see that Z decreases poleward and Θ decreases with pressure.

5 Reduction to an optimal transport problem

We now use the optimality conditions of Theorem 3.10 to show that under Assumption 2.1 Problem 2.4 reduces to a standard optimal transport problem. This reduction is illustrated in Figure 2. In addition, we show that if the cost function c is twisted (Assumption 2.5) and its gradient with respect to x is bounded away from zero (Assumption 2.6), then the corresponding Monge problem admits a unique solution (Corollary 5.2). In the case where ν is discrete, we derive the form of this solution in terms of *Laguerre cells* (Corollary 5.3), which can be found by solving a semi-discrete optimal transport problem [25, Section 4]. This result gives rise a numerical method for computing energy-minimising MLM states from data. An example numerical solution using the background state cost function (13) is illustrated in Figure 3. The discrete target measure ν is defined by (17) using the 3-D mass and circulation integrals (2) and (3) computed from ERA-Interim data [14] at 1800UTC on 22 January 2010, as in [27], making Figure 3 comparable to [27, Figure 4]. We solve the semi-discrete optimal transport problem numerically using the damped Newton method of [20].

Theorem 5.1. *Let $y_{\text{ext}} \in \mathbb{R}^d \setminus Y$ and define $c_{\text{ext}} : X \times Y \cup \{y_{\text{ext}}\} \to [0, +\infty)$ by*

$$c_{\text{ext}}(x, y) = \begin{cases} c(x, y) & \text{if} \quad y \in Y \\ 0 & \text{if} \quad y = y_{\text{ext}}. \end{cases}$$

Let $P > 0$ be a uniform upper bound on the set $\mathcal{S}$, and define

$$\mu_{\text{ext}} \coloneqq \mathcal{L}^d \llcorner (B \times [0, P]), \qquad \nu_{\text{ext}} \coloneqq \nu + (P-1)\delta_{y_{\text{ext}}}.$$

For $\nu \in \mathcal{P}(Y)$,

$$\left.\begin{aligned} &\bar{p} \in \arg\min F_\nu \\ &\psi \in \arg\max G_\nu \\ &\gamma \in \operatorname*{arg\,min}_{\tilde{\gamma} \in \Gamma(\mu_{\bar{p}}, \nu)} \int_{X \times Y} c \, \mathrm{d}\tilde{\gamma} \end{aligned}\right\} \iff \begin{cases} \bar{p} = \max\limits_{y \in Y} q(\cdot, y, \psi_{\text{ext}}(y) - \psi_{\text{ext}}(y_{\text{ext}})) \\ \psi_{\text{ext}} \text{ is a Kantorovich potential from } \mu_{\text{ext}} \text{ to } \nu_{\text{ext}} \\ \gamma_{\text{ext}} \in \operatorname*{arg\,min}\limits_{\tilde{\gamma} \in \Gamma(\mu_{\text{ext}}, \nu_{\text{ext}})} \int_{X \times Y} c_{\text{ext}} \, \mathrm{d}\tilde{\gamma} \end{cases} \tag{50}$$

where

$$\gamma_{\text{ext}} \coloneqq \gamma + \left(\mathcal{L}^d \llcorner \{(s,p) \,:\, s \in B,\, \bar{p}(s) \le p \le P\} \right) \otimes \delta_{y_{\text{ext}}}, \tag{51}$$

and

$$\psi_{\text{ext}}(y) = \begin{cases} \psi(y) & \text{if} \quad y \in Y, \\ -\tau_\psi & \text{if} \quad y = y_{\text{ext}}. \end{cases} \tag{52}$$

Proof. Let $\nu \in \mathcal{P}(Y)$. The optimality conditions for Problem 2.4 given in Theorem 3.10 state that

$$\left.\begin{aligned} &\bar{p} \in \arg\min F_\nu \\ &\psi \in \arg\max G_\nu \end{aligned}\right\} \iff \begin{cases} \bar{p} = \bar{p}_\psi \\ \psi \text{ is a Kantorovich potential from } \mu_{\bar{p}} \text{ to } \nu. \end{cases}$$

Combining this with the classical optimality conditions for optimal transport relating optimal transport plans to Kantorovich potentials (see for example [25, Proposition 8]), and definition of $\bar{p}_\psi$, gives

$$\left.\begin{aligned} &\bar{p} \in \arg\min F_\nu \\ &\psi \in \arg\max G_\nu \\ &\gamma \in \operatorname*{arg\,min}_{\tilde{\gamma} \in \Gamma(\mu_{\bar{p}}, \nu)} \int_{X \times Y} c \, \mathrm{d}\tilde{\gamma} \end{aligned}\right\} \iff \begin{cases} \bar{p} = \max\limits_{y \in Y} q(\cdot, y, \psi(y) + \tau_\psi) \\ \psi^c \oplus \psi = c \quad \gamma\text{-a.e} \end{cases} \tag{53}$$

where τ_ψ is the unique constant such that $\bar{p}_\psi$ is a probability measure on B. Conversely, again by [25, Proposition 8],

$$\left.\begin{aligned} &\bar{p} = \max_{y \in Y} q(\cdot, y, \psi_{\text{ext}}(y) - \psi_{\text{ext}}(y_{\text{ext}})) \\ &\psi_{\text{ext}} \text{ is a Kantorovich potential from } \mu_{\text{ext}} \text{ to } \nu_{\text{ext}} \\ &\gamma_{\text{ext}} \in \operatorname*{arg\,min}_{\tilde{\gamma} \in \Gamma(\mu_{\text{ext}}, \nu_{\text{ext}})} \int_{X \times Y} c_{\text{ext}} \, \mathrm{d}\tilde{\gamma} \end{aligned}\right\} \iff \begin{cases} \bar{p} = \max\limits_{y \in Y} q(\cdot, y, \psi_{\text{ext}}(y) - \psi_{\text{ext}}(y_{\text{ext}})) \\ \psi_{\text{ext}}^{c_{\text{ext}}} \oplus \psi_{\text{ext}} = c_{\text{ext}} \quad \gamma_{\text{ext}}\text{-a.e.} \end{cases}$$

Therefore, to establish (50), it is sufficient to show that

$$\psi^c \oplus \psi = c \quad \gamma\text{-a.e} \iff \psi_{\text{ext}}^{c_{\text{ext}}} \oplus \psi_{\text{ext}} = c_{\text{ext}} \quad \gamma_{\text{ext}}\text{-a.e} \tag{54}$$

where $\psi \in \mathcal{C}(Y)$, $\tau_\psi \in \mathbb{R}$, and $\psi_{\text{ext}} \in \mathcal{C}(Y \times \{y_{\text{ext}}\})$ are related via (52), and γ_{ext} is defined by (51). In particular, since $B \times \{0\}$ is $\mathcal{L}^d$-negligible, the set $N \coloneqq (B \times \{0\}) \times Y$ has zero γ measure and zero γ_{ext} measure, and it is sufficient to show that

$$\psi^c \oplus \psi = c \quad \text{on } \operatorname{spt}(\gamma) \setminus N \iff \psi_{\text{ext}}^{c_{\text{ext}}} \oplus \psi_{\text{ext}} = c_{\text{ext}} \quad \text{on } \operatorname{spt}(\gamma_{\text{ext}}) \setminus N. \tag{55}$$

First, we show that

$$\psi_{\text{ext}}^{c_{\text{ext}}}(x) + \psi_{\text{ext}}(y) = c_{\text{ext}}(x,y) \quad \forall\, (x,y) \in \text{spt}(\gamma_{\text{ext}}) \setminus (N \cup \text{spt}(\gamma)), \tag{56}$$

regardless of any hypotheses on ψ and γ. Let , $(x,y) = ((s,p),y) \in \text{spt}(\gamma_{\text{ext}}) \setminus (N \cup \text{spt}(\gamma))$. By construction $p \geq \bar{p}(s) > 0$, $y = y_{\text{ext}}$, and $\tau_\psi = -\psi_{\text{ext}}(y) = c_{\text{ext}}(x,y) - \psi_{\text{ext}}(y)$. Let $\tilde{y} \in Y$. On the one hand, if $c((s,0),\tilde{y}) \leq \psi(\tilde{y}) + \tau_\psi$, since $c((s,\cdot),\tilde{y})$ is increasing,

$$c((s,p),\tilde{y}) \geq c((s,\bar{p}(s)),\tilde{y}) \geq \psi(\tilde{y}) + \tau_\psi,$$

where the final inequality holds by definition of q. On the other hand, if $c((s,0),\tilde{y}) \geq \psi(\tilde{y}) + \tau_\psi$, then

$$c((s,p),\tilde{y}) \geq c((s,0),\tilde{y}) \geq \psi(\tilde{y}) + \tau_\psi.$$

In either scenario,

$$c_{\text{ext}}(x,\tilde{y}) - \psi(\tilde{y}) \geq \tau_\psi = c_{\text{ext}}(x,y) - \psi_{\text{ext}}(y),$$

so (56) holds.

To complete the proof, let $(x,y) = ((s,p),y) \in \text{spt}(\gamma) \setminus N$. Then $y \in Y$ and $0 < p \leq \bar{p}(s)$. In particular, there exists $\tilde{y} \in Y$ such that

$$c((s,\bar{p}(s),\tilde{y}) - \psi(\tilde{y}) \leq \tau_\psi,$$

otherwise $\bar{p}(s)$ would be zero by definition. Since $c((s,\cdot),\tilde{y})$ is increasing (Assumption 2.1.B.),

$$\psi^c(x) = \psi^c((s,p)) \leq c((s,p),\tilde{y}) - \psi(\tilde{y}) \leq c((s,\bar{p}(s)),\tilde{y}) - \psi(\tilde{y}) \leq \tau_\psi.$$

By construction

$$\tau_\psi = c_{\text{ext}}((s,\bar{p}(s)),y_{\text{ext}}) - \psi_{\text{ext}}(y_{\text{ext}}).$$

It follows that $\psi^c(x) = \psi_{\text{ext}}^{c_{\text{ext}}}(x)$. Therefore,

$$\psi^c(x) + \psi(y) = c(x,y) \iff \psi_{\text{ext}}^{c_{\text{ext}}}(x) + \psi_{\text{ext}}(y) = c_{\text{ext}}(x,y). \tag{57}$$

Since $\text{spt}(\gamma) \subseteq \text{spt}(\gamma_{\text{ext}})$, this establishes the backwards implication in (55). Together, (56) and (57) establish the forwards implication in (55), which completes the proof. □

If Assumption 2.6 holds, then the cost function c_{ext} is twisted. The following corollary is therefore a direct consequence of Theorem 5.1 and the Gangbo-McCann Theorem (see for example [17] or [25, Theorem 12]).

Corollary 5.2. *Let c_{ext}, μ_{ext}, ν_{ext}, ψ_{ext}, γ_{ext}, and P be defined as in Theorem 5.1. If Assumption 2.6 holds, then there exists a unique optimal transport map T_{ext} from μ_{ext} to ν_{ext} for the cost c_{ext}, and it satisfies $\gamma_{\text{ext}} = (\text{id}_{B\times[0,P]}, T_{\text{ext}})\#\mu_{\text{ext}}$, and*

$$\nabla_x c_{\text{ext}}(x, T_{\text{ext}}(x)) = \nabla \psi_{\text{ext}}^c(x) \qquad \forall\, x \in B \times [0,P].$$

In particular, if $\bar{p} \in \arg\min F_\nu$ is the unique solution of Problem 2.4 then $T := T_{\text{ext}}|_{X_{\bar{p}}}$ is the optimal transport map from $\mu_{\bar{p}}$ to ν for the cost c, where we recall that $X_{\bar{p}} = \{(s,p) \in X \,|\, p \leq \bar{p}(s)\}$.

With a view to solving Problem 2.4 numerically, we now write down the formulation of Corollary 5.2 in the case that ν is a finitely supported measure. This reduces Problem 2.4 to a standard semi-discrete optimal transport problem. (See [25, Section 4] for an overview of semi-discrete optimal transport). By [25, Proposition 37] the corresponding Monge problem has a solution T expressed in (59) in terms of so-called *Laguerre cells* L_i (see (58) and more generally [25, Definition 17]). The intersection of any two Laguerre cells is $\mathcal{L}^d$-negligible [25, Proposition 37], so T is well defined and corresponds to a labelled partition of the set $X_{\bar{p}}$.

Corollary 5.3. *Suppose Assumption 2.6 holds. Let $n \in \mathbb{N}$, and for $i \in \{1, \dots, n\}$ let $y_i \in Y$ be distinct and $m_i > 0$ be such that $\sum_{i=1}^n m_i = 1$. Let*

$$\nu = \sum_{i=1}^n m_i \delta_{y_i},$$

and let $\bar{p} \in \arg\min F_\nu$ be the unique solution of Problem 2.4. Let y_{ext}, c_{ext}, μ_{ext}, and ν_{ext} be defined as in Theorem 5.1, and let ψ_{ext} be a Kantorovich potential between μ_{ext} and ν_{ext} for the cost c_{ext}. Define $\Psi := (\psi_{\text{ext}}(y_i))_{i=1}^{n+1}$ and $\mathbf{y} = (y_i)_{i=1}^{n+1}$, with the convention that $y_{n+1} = y_{\text{ext}}$. For $i \in \{1, \dots, n+1\}$, define

$$L_i(\Psi, \mathbf{y}) := \left\{ x \in B \times [0, P] \;\middle|\; c_{\text{ext}}(x, y_i) - \Psi_i \leq c_{\text{ext}}(x, y_j) - \Psi_j \quad \forall\, j \in \{1, \dots, n+1\} \right\}. \tag{58}$$

Then

$$\text{graph}(\bar{p}) \subseteq \partial L_{n+1}(\Psi, \mathbf{y}),$$

and $T : X_p \to Y$ defined by

$$T := \sum_{i=1}^n y_i \mathbb{1}_{L_i(\Psi, \mathbf{y})} \tag{59}$$

is the optimal transport map from $\mu_{\bar{p}}$ to ν for the cost c.

6 Conclusion

In this work, we have given a rigorous definition of a Modified Lagrangian Mean (MLM) state in a single hemisphere in latitude-pressure coordinates (Definition 1.1). Such a state consists of a surface pressure function $\overline{p}$ determining the extent of the source domain, and an admissible transport map $T = (Z, \Theta)$ between the uniform density on the source domain and a mass distribution ν in the space of zonal angular momentum and potential temperature. We have proved that if ν has compact support in $(0, +\infty)^2$, then there exists a unique energy-minimising MLM state defined on a domain excluding the equator and the pole (Theorem 3.10). Optimality conditions imply that this state is in fact the solution of a single optimal transport problem (Theorem 5.1). Our results hold not only for the cost function defining the energy of an MLM state (13) but for a wide class of cost functions satisfying mild assumptions. These include the cost function (15) defining free-surface variants of the semi-geostrophic equations [6, 9, 8].

Regarding energy-minimising MLM states specifically, our results come with three main caveats. First, the exclusion of the extreme latitudes in the given hemisphere. The pole is excluded to ensure that the cost function is sufficiently regular (locally Lipschitz; Assumption 2.1.A.). The equator is excluded to ensure that the cost function is twisted (Assumption 2.5) and admits a twisted extension (Assumption 2.6). This guarantees the existence of optimal transport maps (rather than merely optimal transport plans), which can be interpreted as energy-minimising rearrangements of zonal angular momentum and potential temperature. Second, and for the same two reasons, an arbitrary minimum pressure is assumed and built into the cost function (13). Third, the zonal angular momentum is assumed to be positive and bounded away from zero to ensure that the cost function is twisted.

In addition to the general existence and uniqueness result, we have shown that to compute an energy-minimising MLM state from a discrete mass distribution in zonal angular momentum and potential temperature coordinates (target space), as in the setting of [27], it is sufficient to solve

a semi-discrete optimal transport problem with a twisted cost (Corollary 5.3). This reduces to maximising the Kantorovich dual functional, which is a concave, finite dimensional, unconstrainted maximisation problem [25, Section 4], and is numerically tractable using, for example, the damped Newton method of [20]. Figure 3 illustrates an example where the discrete mass distribution in target space is specified from integrals calculated from 3-D global atmospheric data and the semi-discrete optimal transport solution obtains the corresponding distribution of mass in the source space corresponding to a meridional cross-section in latitude-pressure coordinates.

Acknowledgements

The authors thank Thilo D. Stier (University of Göttingen) for implementing the numerical semi-discrete optimal transport solver that was used to generate Figure 3. The authors would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the satellite programme 'Geophysical Fluid Dynamics; from Mathematical Theory to Operational Prediction', hosted at the University of Reading, where work on this paper was initiated, and which was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/R014604/1. CE gratefully acknowledges the support of: the EPSRC via the grant EP/W522570/1; the German Research Foundation (DFG) via the CRC 1456 'Mathematics of Experiment', project A03; and the University of Reading Advancing the Frontiers of Earth System Prediction (AFESP) Programme (award number A3720200). DPB acknowledges financial support from the EPSRC grant EP/V00204X/1.

Data Availability

The ECMWF reanalysis data used to produce Figure 3 is available from the ECMWF MARS archive (https://confluence.ecmwf.int/display/CKB/The+family+of+ERA5+datasets). Note that the full resolution native ECMWF model output in terms of spectral coefficients on terrain-following model levels was used to compute the mass and circulation integrals (rather than the gridded, pressure level data available at the Copernicus Data Store). Here, ERA-Interim data was used, but ERA5 data gives very similar results for the mass distribution ν in (z, θ) coordinates.

References

[1] G. Bouchitté and G. Buttazzo, *New lower semicontinuity results for nonconvex functionals defined on measures*, Nonlinear Analysis: Theory, Methods & Applications, 15 (1990), pp. 679–692.

[2] J. Bowman and T. Shepherd, *Nonlinear symmetric stability of planetary atmospheres*, J. Fluid Mech., 296 (1995), pp. 391–407.

[3] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Communications on Pure and Applied Mathematics, 44 (1991), pp. 375–417.

[4] Y. Brenier, U. Frisch, M. Hénon, G. Loeper, S. Matarrese, R. Mohayaee, and A. Sobolevski, *Reconstruction of the early Universe as a convex optimization problem*, Monthly Notices of the Royal Astronomical Society, 346 (2003), pp. 501–524.

[5] G. Carlier, *Classical and modern optimization*, World Scientific, 2022.

[6] J. Cheng, *Semigeostrophic equations in physical space with free upper boundary*, Calculus of Variations and Partial Differential Equations, 56 (2016), pp. 55–149.

[7] M. J. P. Cullen, *The Mathematics of Large-Scale Atmosphere and Ocean*, World Scientific, 2021.

[8] M. J. P. Cullen and W. Gangbo, *A variational approach for the 2-dimensional semi-geostrophic shallow water equations*, Archive for Rational Mechanics and Analysis, 156 (2001), pp. 241–273.

[9] M. J. P. Cullen, D. Gilbert, and B. Pelloni, *Solutions of the fully compressible semi-geostrophic system*, Communications in Partial Differential Equations, 39 (2014), pp. 591–625.

[10] M. J. P. Cullen, T. Kuna, B. Pelloni, and M. Wilkinson, *The stability principle and global weak solutions of the free surface semi-geostrophic equations in geostrophic coordinates*, Proceedings of the Royal Society A, 475 (2019), p. 20180787.

[11] M. J. P. Cullen and H. Maroofi, *The fully compressible semi-geostrophic system from meteorology*, Arch. Rational Mech. Anal., 167 (2003), p. 309–336.

[12] M. J. P. Cullen and R. J. Purser, *An extended Lagrangian theory of semi-geostrophic frontogenesis*, Journal of the Atmospheric Sciences, 41 (1984), pp. 1477–1497.

[13] M. J. P. Cullen and M. Sedjro, *On a model of forced axisymmetric flows*, SIAM Journal on Mathematical Analysis, 46 (2014), pp. 3983–4013.

[14] D. P. Dee, S. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. Balmaseda, G. Balsamo, d. P. Bauer, et al., *The ERA-Interim reanalysis: Configuration and performance of the data assimilation system*, Quarterly Journal of the royal meteorological society, 137 (2011), pp. 553–597.

[15] J. Faria, *On the existence and weak stability of solutions to the compressible semigeostrophic equations*, Journal of Mathematical Analysis and Applications, 406 (2013), pp. 447–463.

[16] U. Frisch, S. Matarrese, R. Mohayaee, and A. Sobolevski, *A reconstruction of the initial conditions of the universe by optimal mass transportation*, Nature, 417 (2002), pp. 260–262.

[17] W. Gangbo and R. J. McCann, *The geometry of optimal transportation*, Acta Mathematica, 177 (1996), pp. 113–161.

[18] P. Haynes, *Forced, dissipative generalisations of finite-amplitude wave-activity conservation relations for zonal and non-zonal basic flows*, J. Atmos. Sci., 45 (1988), pp. 2352–2362.

[19] B. J. Hoskins and I. N. James, *Fluid dynamics of the mid-latitude atmosphere*, John Wiley & Sons, 2014.

[20] J. Kitagawa, Q. Mérigot, and B. Thibert, *Convergence of a Newton algorithm for semi-discrete optimal transport*, Journal of the European Mathematical Society, 21 (2019), pp. 2603–2651.

[21] B. Lévy, *Partial optimal transport for a constant-volume lagrangian mesh with free boundaries*, Journal of Computational Physics, 451 (2022), p. 110838.

[22] B. Lévy, R. Mohayaee, and S. von Hausegger, *A fast semidiscrete optimal transport algorithm for a unique reconstruction of the early Universe*, Monthly Notices of the Royal Astronomical Society, 506 (2021), pp. 1165–1185.

[23] M. McIntyre, *Towards a Lagrangian-mean description of stratospheric circulations and chemical transports*, Phil. Trans. R. Soc. Lond., A296 (1980), pp. 129–148.

[24] M. McIntyre and T. Shepherd, *An exact local conservation theorem for finite-amplitude disturbances to non-parallel shear flows, with remarks on Hamiltonian structure and on Arnold's stability theorems*, J. Fluid Mech., 181 (1987), pp. 527–565.

[25] Q. Mérigot and B. Thibert, *Optimal transport: discretization and algorithms*, in Handbook of numerical analysis, vol. 22, Elsevier, 2021, pp. 133–212.

[26] J. Methven, *Wave activity for large amplitude disturbances described by the primitive equations on the sphere*, J. Atmos. Sci., 70 (2013), pp. 1616–1630.

[27] J. Methven and P. Berrisford, *The slowly evolving background state of the atmosphere*, Quarterly Journal of the Royal Meteorological Society, 141 (2015), pp. 2237–2258.

[28] F. Santambrogio, *Optimal Transport for Applied Mathematicians*, Springer, 2015.

[29] C. Sarrazin, *Lagrangian discretization of variational mean field games*, SIAM Journal on Control and Optimization, 60 (2022), pp. 1365–1392.

[30] T. Shepherd, *Symmetries, conservation laws, and Hamiltonian structure in geophysical fluid dynamics*, Adv. Geophys., 32 (1990), pp. 287–338.

[31] C. Villani, *Optimal transport: old and new*, Springer, 2008.

[32] C. Zalinescu, *Convex analysis in general vector spaces*, World Scientific Publishing Co., Inc, 2002.