SEMI-DISCRETE OPTIMAL TRANSPORT TECHNIQUES FOR THE COMPRESSIBLE SEMI-GEOSTROPHIC EQUATIONS

DAVID P. BOURNE*, CHARLIE EGAN†, THÉO LAVIER*, AND BEATRICE PELLONI*

Abstract. We prove existence of weak solutions of the 3D compressible semi-geostrophic (SG) equations with compactly supported measure-valued initial data. These equations model large-scale atmospheric flows. Our proof uses a particle discretisation and semi-discrete optimal transport techniques. We show that, if the initial data is a discrete measure, then the compressible SG equations admit a unique, energy-conserving and global-in-time solution, generated by particle trajectories that are twice continuously differentiable. In general, by discretising the initial measure by particles and sending the number of particles to infinity, we show that for any compactly supported initial measure there exists a global-in-time solution of the compressible SG equations that is Lipschitz in time. This significantly generalises the original results due to Cullen and Maroofi (2003), and it provides the theoretical foundation for the design of numerical schemes using semi-discrete optimal transport to solve the 3D compressible SG equations.

1. Introduction

The semi-geostrophic (SG) equations for a compressible fluid are a simplified model of the formation and evolution of atmospheric fronts, which are important for accurate weather prediction on large scales. The model, incorporating the effect of compressibility, was introduced and analysed in [20].

In this paper we study the system formulated in terms of the so-called *geostrophic variables*. In these variables, the governing evolution equation is a continuity equation for a time-dependent probability measure α_t, usually called the *potential vorticity*.[1] The equation for α_t is given by

$$\partial_t \alpha_t + \nabla \cdot (\alpha_t \mathcal{W}[\alpha_t]) = 0. \tag{1.1}$$

The non-local velocity field $\mathcal{W}[\alpha_t]$ (defined in (1.4) below) is known as the *geostrophic wind*. It is defined at each time t by minimising the *geostrophic energy* given the potential vorticity α_t. Here, and in what follows, the subscript t on a dependent variable denotes evaluation at time t. The PDE (1.1) is derived in [20] by starting from Euler's equations for a compressible fluid, assuming a shallow atmosphere in hydrostatic balance, a rotation-dominated flow, and making the geostrophic approximation, which is valid for large-scale flows, then transforming the resulting equations to geostrophic coordinates.

The main goal of this paper is to construct global-in-time weak solutions of (1.1) as the limit of spatially discrete approximations. This generalises and extends the semi-discrete optimal transport strategy presented in [8] for the incompressible setting, a generalisation that requires substantial changes and poses new technical challenges.

There are several benefits of using semi-discrete optimal transport: this approach highlights an intuitive connection between flows in geostrophic coordinates and corresponding flows in the fluid domain, it provides the means to construct explicit solutions, and it serves as the theoretical foundation for adapting Cullen and Purser's groundbreaking geometric method [21, 23] to solve the compressible SG equations numerically.

(*) Maxwell Institute for Mathematical Sciences and Department of Mathematics, Heriot-Watt University, Edinburgh, UK

(†) Department of Meteorology, University of Reading, Reading, UK

[1] In physical terms, α_t is the inverse of the physical potential vorticity; see [15].

1.1. **Summary of results.** We denote by $\mathcal{X} \subset \mathbb{R}^3$ the compact set representing the physical fluid domain, and by $\mathcal{Y} \subset \mathbb{R}^3$ the *geostrophic domain* where the equations are defined after the change to geostrophic variables. Throughout this paper we assume that $\mathcal{X}$ and $\mathcal{Y}$ satisfy the assumptions stated in Section 2.1.

To define the geostrophic energy of the system, consider an absolutely continuous probability measure $\sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})$. The measure σ is a physical variable, representing the product of the fluid density and potential temperature. For a given probability measure $\alpha_t \in \mathscr{P}(\mathcal{Y})$ representing the potential vorticity, let $\mathcal{T}_c(\sigma, \alpha_t)$ be the *optimal transport cost* between σ and α_t, given below by (2.3), for the cost function $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ given by equation (2.2). The corresponding *geostrophic energy* $E(\sigma, \alpha_t)$ is given by

$$E(\sigma, \alpha_t) = \mathcal{T}_c(\sigma, \alpha_t) + \kappa \int_{\mathcal{X}} \sigma^\gamma \, \mathrm{d}\mathbf{x}, \tag{1.2}$$

where $\gamma \in (1, 2)$ and $\kappa > 0$ are physical constants[2]. The geostrophic energy is the sum of the kinetic energy of the fluid, its gravitational potential energy, and its internal energy.

The geostrophic wind $\mathcal{W}$ that appears in (1.1) is defined by minimising the geostrophic energy over all σ. We show that this minimisation problem has a unique solution

$$\sigma_*[\alpha_t] = \operatorname*{arg\,min}_{\sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})} E(\sigma, \alpha_t), \tag{1.3}$$

which we call the *optimal source measure*. Let T_{α_t} denote the *optimal transport map* from $\sigma_*[\alpha_t]$ to α_t for the cost c. Then the velocity field $\mathcal{W}[\alpha_t] : \mathcal{Y} \to \mathbb{R}^3$ in (1.1) can be formally defined by

$$\mathcal{W}[\alpha_t](\mathbf{y}) = J\left(\mathbf{y} - T_{\alpha_t}^{-1}(\mathbf{y})\right), \qquad J = f_{\mathrm{cor}} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{1.4}$$

where $f_{\mathrm{cor}} > 0$ is the Coriolis parameter, which we take to be constant. The optimal transport map T_{α_t} exists and is unique because the cost function c given by (2.2) satisfies the classical *twist* condition. In general, the transport map T_{α_t} is not invertible, so equation (1.4) is only formal; we give a rigorous meaning to the PDE (1.1) in Definition 2.6.

Equations (1.1)-(1.4) are obtained from the system of equations in physical variables (see the background section below) by the change to *geostrophic variables* given by

$$T_{\alpha_t}(\mathbf{x}) = \left(x_1 + f_{\mathrm{cor}}^{-2} v_2^g(\mathbf{x}, t), x_2 - f_{\mathrm{cor}}^{-2} v_1^g(\mathbf{x}, t), \theta(\mathbf{x}, t)\right),$$

where $\mathbf{v}^g = (v_1^g, v_2^g, 0)$ is the geostrophic velocity of the fluid and θ is its potential temperature [20, equation (5)]. The physical system conserves geostrophic energy, which in these variables has the expression (1.2). The physical requirement that solutions minimise this energy at each fixed time (known as *Cullen's convexity principle*) translates mathematically to the requirement that T_{α_t} be the optimal transport map from $\sigma_*[\alpha_t]$ to α_t. The time evolution in geostrophic variables is governed by equation (1.1), which is coupled to the optimal transport problem through the definition of the velocity $\mathcal{W}[\alpha_t]$. The knowledge of the map T_{α_t} also yields a way to recover the physical quantities $\mathbf{v}^g$ and θ from the components of T_{α_t}. This formulation is analogous to the formulation given for the incompressible system in [5], whereby one obtains an evolution equation such as (1.1) coupled with a Monge-Ampère equation for the convex function P such that $T_{\alpha_t} = \nabla P$. The Monge-Ampère connection is specific to the case of the quadratic cost featuring in the incompressible case. For a detailed derivation of equations (1.1)–(1.4) see [20] and [37].

[2]Interpreted physically, γ is the ratio $\frac{c_{\mathrm{p}}}{c_{\mathrm{v}}}$, where c_{p} is the specific heat at constant pressure and c_{v} is the specific heat at constant volume, and $\kappa = c_v \left(\frac{R_d}{p_0}\right)^{\gamma - 1}$, where R_d is the specific gas constant for dry air, and p_0 is the reference pressure.

To state the first of our results, we formulate the discrete analogue of the PDE (1.1) by making the discrete ansatz

$$\alpha_t^N = \sum_{i=1}^N m^i \delta_{\mathbf{z}^i(t)}, \tag{1.5}$$

where $\mathbf{z}^i(t) \in \mathcal{Y}$. Let $\mathbf{z}(t) = \big(\mathbf{z}^1(t), \dots, \mathbf{z}^N(t)\big) \in \mathcal{Y}^N$. We show in Proposition 4.5 that this ansatz gives rise to the ODE

$$\dot{\mathbf{z}} = J^N(\mathbf{z} - \mathbf{C}(\mathbf{z})), \tag{1.6}$$

where $J^N \in \mathbb{R}^{3N\times 3N}$ is the block diagonal matrix $J^N := \operatorname{diag}(J, \dots, J)$, and $\mathbf{C}$ is the *centroid map*, which is defined in Definition 4.2. Heuristically, the velocity of the particle $\mathbf{z}^i$ at time t is the evaluation of the velocity field $\mathcal{W}[\alpha_t^N]$ at $\mathbf{z}^i(t)$. However, since α_t^N is discrete, the transport map $T_{\alpha_t^N}$ is not invertible and (1.4) is not well defined. Instead, for each i, a whole region in $\mathcal{X}$ is transported to the single point $\mathbf{z}^i(t)$ and the term involving $T^{-1}_{\alpha_t^N}$ in (1.4) is replaced by the centroid of the region $T^{-1}_{\alpha_t^N}\{\mathbf{z}^i(t)\}$ resulting in the ODE (1.6). See Figure 1. (The sets $T^{-1}_{\alpha_t^N}\{\mathbf{z}^i(t)\}$ are in fact *c-Laguerre cells*, which we denote by L_c^i. See Theorem 2.9.)

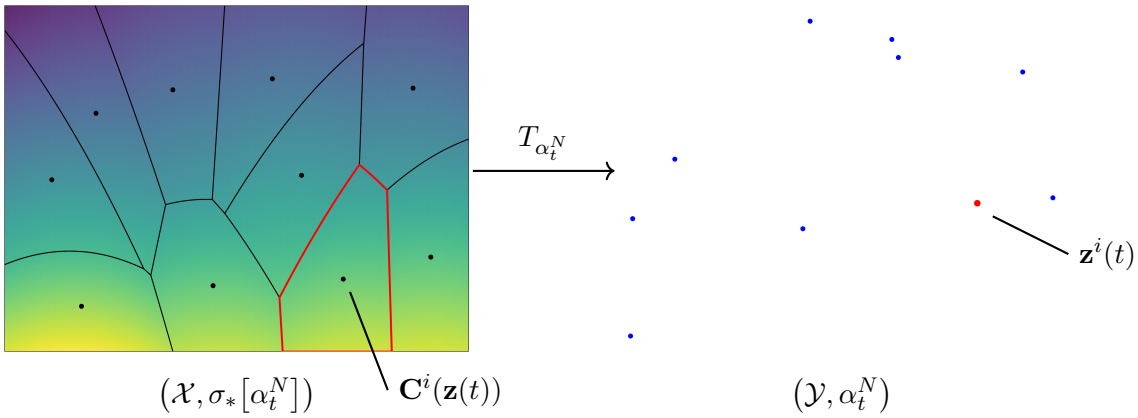


FIGURE 1. On the left is the source space $\mathcal{X}$ coloured by the density of the optimal source measure $\sigma_*[\alpha_t^N]$. The boundaries and centroids $\mathbf{C}^j(\mathbf{z}(t))$ of the regions $T^{-1}_{\alpha_t^N}\{\mathbf{z}^i(t)\}$ are plotted in black. The boundary of the i-th region is highlighted in red. On the right is the target space $\mathcal{Y}$ with the seeds $\mathbf{z}^j(t)$ in blue. The i-th seed, $\mathbf{z}^i(t)$, corresponding to the i-th region is highlighted in red. The union of the seeds is the support of the target measure α_t^N.

We now state our two main results. The notation we use is listed in Section 2. In particular, by $\mathcal{C}^{0,1}([0,\tau];\mathscr{P}(\mathcal{Y}))$ we mean Lipschitz continuity with respect to the W^1 Wasserstein metric on $\mathscr{P}(\mathcal{Y})$ and the standard metric on $[0,\tau]$.

Our first theorem asserts the existence and uniqueness of solutions of the ODE (1.6) and correspondingly of the PDE (1.1) with discrete initial data.

Theorem 1.1 (Existence of discrete geostrophic solutions)**.** *Let the assumptions in Section 2.1 hold. Fix $N \in \mathbb{N}$, an arbitrary final time $\tau > 0$, and an initial discrete probability measure $\overline{\alpha}^N = \sum_{i=1}^N m^i \delta_{\overline{\mathbf{z}}^i} \in \mathscr{P}^N(\mathcal{Y})$ that is well-prepared in the sense of (2.9). Then there exists a unique solution $\mathbf{z} \in \mathcal{C}^2([0,\tau];\mathcal{Y}^N)$ of the ODE (1.6) with initial condition $\mathbf{z}(0) = (\overline{\mathbf{z}}^1, \dots, \overline{\mathbf{z}}^N)$. Define $\alpha^N \in \mathcal{C}^{0,1}([0,\tau];\mathscr{P}^N(\mathcal{Y}))$ by*

$$\alpha_t^N = \sum_{i=1}^N m^i \delta_{\mathbf{z}^i(t)}.$$

Then α^N is the unique weak solution of the compressible SG equations (in the sense of Definition 2.6) with discrete initial data $\alpha_0^N = \overline{\alpha}^N$. Moreover, α^N is energy-conserving in the sense that

$$E(\sigma_*[\alpha_t^N], \alpha_t^N) = E(\sigma_*[\alpha_0^N], \alpha_0^N) \quad \forall\, t \in [0, \tau].$$

Our second theorem states the existence of global-in-time weak solutions of the compressible SG equations (1.1) for arbitrary, compactly supported initial data.

Theorem 1.2 (Existence of weak geostrophic solutions)**.** *Let the assumptions in Section 2.1 hold. Let $\tau > 0$ be an arbitrary final time, and let $\overline{\alpha} \in \mathscr{P}_{\mathrm{c}}(\mathcal{Y})$ be an initial compactly-supported probability measure. Then there exists $\alpha \in \mathcal{C}^{0,1}([0,\tau]; \mathscr{P}_{\mathrm{c}}(\mathcal{Y}))$ and $\alpha^N \in \mathcal{C}^{0,1}([0,\tau]; \mathscr{P}^N(\mathcal{Y}))$, $N \in \mathbb{N}$, such that α is a weak solution of the compressible SG equations with initial measure $\overline{\alpha}$ (in the sense of Definition 2.6), α^N is a discrete weak solution for each $N \in \mathbb{N}$, and the sequence (α^N) converges uniformly to α in the following sense:*

$$\lim_{N\to\infty} \sup_{t\in[0,\tau]} W_1(\alpha_t^N, \alpha_t) = 0. \tag{1.7}$$

It is not known whether weak solutions of the compressible SG equations are unique. In particular, Theorem 1.2 only asserts that the weak solution that we construct can be uniformly approximated by discrete weak solutions; it does not assert that *any* weak solution can be uniformly approximated in this way.

We stress that while the incompressible SG system considered in [8] involves the standard quadratic cost and a fixed source measure, the compressible SG system involves a non-standard cost c and a time-dependent source measure, defined by the minimisation problem (1.3). This presents new, significant challenges regarding the regularity of the centroid map $\mathbf{C}$.

Our method of proof gives rise naturally to a numerical method. Indeed, to show that the ODE (1.6) has a unique solution, we first derive a concave dual formulation of the energy minimisation problem (1.3) in the case where α_t is a discrete measure; see Section 3. We show that this dual problem is solvable and that its solution can be used to define the optimal source measure $\sigma_*[\alpha_t]$. While we do not compute numerical solutions here, the dual problem is an unconstrained, finite-dimensional, concave maximisation problem, which is numerically tractable. (The analogous problem with the quadratic cost is solved numerically for example in [42].) Hence the velocity field in the ODE (1.6) can, in principle, be computed numerically. Importantly, solutions of the ODE (1.6) give rise to discrete, energy-conserving solutions of the PDE (1.1). It has been observed in the incompressible setting that numerical solutions inherit this conservation property [23, 36]. Implementing a numerical scheme based on these observations would be a natural extension to the compressible setting of Cullen and Purser's geometric method [21, 23].

1.2. **Background.** The Navier-Stokes equations are the fundamental mathematical model used by meteorologists to simulate the dynamics of the atmosphere and oceans. Since the viscosity of air is low, the atmosphere is often modelled as inviscid, so that the diffusion coefficient in the Navier-Stokes equations can be set to zero to obtain the Euler equations. At length scales on the order of tens of kilometres, such flows develop singularities known as atmospheric fronts. These can be modelled using the SG approximation, which is a good approximation for shallow, rotationally-dominated flows (small Rossby numbers). This approximation was first formulated in 1949 by Eliassen [24] and rediscovered in the 1970s by Hoskins [34]. Since then it has been studied extensively (see for example the book [15] and the bibliography therein) and the SG equations have been used as a diagnostic tool at the UK Met Office [14].

The SG approximation is based on two balances that dictate the dynamics of the system. The first is *hydrostatic balance*, namely the assumption that the vertical component of the pressure gradient is balanced by the gravitational force. The second is *geostrophic balance*, which assumes that the first two components of the pressure gradient are balanced by the Coriolis force [24]. The corresponding *geostrophic wind* (in physical variables) is a large-scale, two-dimensional approximation of the fluid velocity and is directed along isobars.

From a physical point of view, the model described by the SG equations conserves the *geostrophic energy*: the sum of kinetic, potential and internal energy. However, mathematically, the equations do not appear to be in a conservative form, as is the case for the well-known *quasi-geostrophic* reduction. This inconsistency was considered by Hoskins [34], who clarified the mathematical structure of the incompressible SG equations by an ingenious change of variables to the so-called *geostrophic* coordinates, sometimes called *dual coordinates*.

1.3. **Previous results.** The change of coordinates to geostrophic variables was made mathematically rigorous in the groundbreaking work of Benamou and Brenier [5], where the incompressible SG equations are formulated in geostrophic variables as a transport equation coupled to a Monge-Ampère equation. Brenier's solution of the optimal transport problem for the quadratic cost in [10] was used to prove the existence of global-in-time weak solutions of the incompressible SG equations in geostrophic variables. This seminal work provided the foundation for much of the subsequent analysis of the SG equations. Indeed, similar reformulations are established in [17] for the SG shallow water equations and in [20] for the compressible SG equations, which we study in this paper, where the template of the proof given in [5] is used to prove global existence of weak solutions.

In [20] the existence of global-in-time solutions of the compressible SG equations is established for compactly-supported initial potential vorticity $\overline{\alpha} \in L^r(\mathcal{Y})$ for $r \in (1, \infty)$. The case $r = 1$ was added by Faria in [26]. We extend this result to $\overline{\alpha} \in \mathscr{P}_c(\mathcal{Y})$. However, in [20] the term in the cost function c representing the gravitational potential energy is taken to be any C^2 function whose derivative with respect to the vertical variable is non-zero, which is more general than the case we treat. The same result is proved in [19] using the general theory for Hamiltonian ODEs in Wasserstein space developed by Ambrosio and Gangbo in [4]. Additionally, [19] presents a proof of the existence of weak Lagrangian solutions in physical space, analogous to the results of Cullen, Feldman & Tudorascu for the incompressible SG equations [16, 27, 29, 30].

In [8] a *semi-discrete* optimal transport approach is used to prove the global existence of weak solutions for the *incompressible* SG equations for measure-valued initial data. Exact solutions of the incompressible SG equations are constructed starting from spatially discrete initial data. These discrete solutions are then used to approximate solutions for arbitrary initial data. Similar existence results were already established in increasing levels of generality in [5, 31, 39, 28]. The approach of [8] is distinct from those of earlier papers in its use of a spatial discretisation based on semi-discrete optimal transport, rather than a combination of time discretisation and spatial mollification as used in [5]. It is also distinct from the fully discrete approach used to study a variant of the incompressible SG equations in [18], where both physical and geostrophic variables are spatially discretised. Recently entropy-regularised, fully discrete optimal transport was used to prove global existence of weak solutions for the incompressible SG equations and convergence of a numerical method [12]. While we cannot give here an exhaustive reference list to what is now a very substantial body of work, other important contributions to the analysis of the SG equations include the works of Ambrosio et al. [1, 2] and recent work on the incompressible SG equations with non-constant Coriolis parameter [43]. The incompressible SG equations with non-constant Coriolis parameter was first considered in [13]. In this paper we assume that the Coriolis parameter is constant.

Numerical methods for the SG equations date back to the development of the geometric method of Cullen & Purser [21]. More recently, major developments in the field of numerical optimal transport have given rise to efficient numerical methods for solving the SG equations. Semi-discrete optimal transport has been used to obtain solutions of the 2-dimensional SG Eady slice equations [23] and the 3D incompressible SG equations [36], while entropy-regularised discrete optimal transport has been used to solve the Eady slice equations [7, 12] and the semigeostrophic shallow water equations [6]. Note that $\mathcal{W}[\alpha_t]$ can be formally rewritten as $\mathcal{W}[\alpha_t](\mathbf{y}) = -f_{\mathrm{cor}}^{-2}\, y_3\, J\nabla\psi_t(\mathbf{y})$, where (φ_t, ψ_t) is an optimal Kantorovich potential pair for transporting $\sigma_*[\alpha_t]$ to α_t with cost c [37]. This expression could be convenient for applying the techniques from [6, 7, 12] to the compressible SG equations. Other applications of semi-discrete

optimal transport in fluid dynamics are studied for example in [32, 33]. The finite element method has recently been used to solve a 2D version of the compressible SG equations [46].

1.4. **Outline of the paper.** Section 2 includes notation, background material, and the definition of a weak solution of (1.1). In Section 3 we derive a dual formulation of the energy minimisation problem (1.3) and in Section 4 we prove Theorem 1.1. Our main theorem, Theorem 1.2, is proved in Section 5. In Section 6 we give two explicit solutions of the compressible SG equations, namely an absolutely continuous steady state solution and a discrete solution with a single particle. In Appendix A we discuss a technical assumption on the fluid domain. In Appendix B we prove a technical lemma, Lemma 3.6, which is important for establishing the regularity of the centroid map $\mathbf{C}$.

2. Preliminaries

2.1. **Assumptions and notation.** Throughout this paper we make the following assumptions:

- $g > 0$ is the acceleration due to gravity and $f_{\text{cor}} > 0$ is the Coriolis parameter, which we assume to be constant.
- $\mathcal{X} \subset \mathbb{R}^3$ denotes the fluid domain. We assume that $\mathcal{X}$ is nonempty, connected, compact, and that it coincides with the closure of its interior. Moreover, we assume that the set $\Phi^{-1}(\mathcal{X})$ is convex, where $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by
$$\Phi(\mathbf{x}) = \left(f_{\text{cor}}^{-2}\, x_1,\, f_{\text{cor}}^{-2}\, x_2,\, g^{-1}\left(x_3 - \tfrac{1}{2} f_{\text{cor}}^{-2}\left(x_1^2 + x_2^2\right)\right)\right). \tag{2.1}$$
In Appendix A we give examples of domains satisfying this assumption and show that it is equivalent to assuming that $\mathcal{X}$ is *c-convex* [35, Definition 1.2], which is an important notion in regularity theory for semi-discrete optimal transport. We require this technical assumption to prove Lemma 3.6.
- $\mathcal{Y} = \mathbb{R}^2 \times (\delta, \frac{1}{\delta})$ denotes the geostrophic domain, where $\delta \in (0, 1)$. This is equivalent to assuming that the initial potential temperature of the fluid is bounded away from zero and infinity.

We also use the following standard notation:

- $\mathscr{D}(\mathcal{Y} \times \mathbb{R})$ denotes the space of real-valued infinitely-differentiable and compactly supported test functions on $\mathcal{Y} \times \mathbb{R}$.
- $\mathcal{H}^d$ denotes the d-dimensional Hausdorff measure.
- For a Borel set $A \subset \mathcal{X}$, $\mathbb{1}_A : \mathcal{X} \to \{0, 1\}$ denotes the characteristic function of the set A. That is
$$\mathbb{1}_A(\mathbf{x}) = \begin{cases} 1 & \text{if} \quad \mathbf{x} \in A, \\ 0 & \text{otherwise.} \end{cases}$$
- $\mathscr{P}(A)$ denotes the set of Borel probability measures on $A \subseteq \mathbb{R}^d$.
- $\mathscr{P}_{\text{ac}}(A)$ denotes the set of Borel probability measures on $A \subseteq \mathbb{R}^d$ that are absolutely continuous with respect to the Lebesgue measure. Throughout this paper we identify absolutely continuous measures with their density functions.
- Given a Borel map $S : A \to B$ and a measure $\mu \in \mathscr{P}(A)$, the pushforward of μ by S is the measure $S_{\#}\mu \in \mathscr{P}(B)$ defined by $S_{\#}\mu(U) = \mu(S^{-1}(U))$ for all Borel sets $U \subseteq B$.
- $\mathscr{P}_{\text{c}}(\mathcal{Y})$ denotes the set of Borel probability measures on $\mathcal{Y}$ with compact support. In this paper we always equip $\mathscr{P}_{\text{c}}(\mathcal{Y})$ with the W_1 Wasserstein metric, which is defined by $W_1 : \mathscr{P}_{\text{c}}(\mathcal{Y}) \times \mathscr{P}_{\text{c}}(\mathcal{Y}) \to [0, \infty)$,
$$W_1(\alpha, \beta) = \inf\left\{\int_{\mathcal{Y}\times\mathcal{Y}} \|\mathbf{y} - \tilde{\mathbf{y}}\| \,\mathrm{d}\gamma(\mathbf{y}, \tilde{\mathbf{y}}) \,:\, \gamma \in \mathscr{P}(\mathcal{Y} \times \mathcal{Y}),\, \pi_{1\#}\gamma = \alpha,\, \pi_{2\#}\gamma = \beta\right\},$$
where $\pi_1 : \mathcal{Y} \times \mathcal{Y} \to \mathcal{Y}$ is defined by $\pi_1(\mathbf{y}, \tilde{\mathbf{y}}) = \mathbf{y}$ and $\pi_2 : \mathcal{Y} \times \mathcal{Y} \to \mathcal{Y}$ is defined by $\pi_1(\mathbf{y}, \tilde{\mathbf{y}}) = \tilde{\mathbf{y}}$.
- $\mathscr{P}^N(\mathcal{Y})$ denotes the set of discrete probability measures on $\mathcal{Y}$ with support of cardinality N; see equation (2.8).

2.2. **The compressible SG equations.** In this section we define weak solutions of (1.1).

Definition 2.1 (The cost function)**.** The cost function $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ associated with the compressible SG equations is given by

$$c(\mathbf{x}, \mathbf{y}) = \frac{1}{y_3}\left(\frac{f_{\mathrm{cor}}^2}{2}(x_1 - y_1)^2 + \frac{f_{\mathrm{cor}}^2}{2}(x_2 - y_2)^2 + g x_3\right). \tag{2.2}$$

It is easy to check that c is *twisted*, namely that c is differentiable in $\mathbf{x}$ and that the map $\mathbf{y} \mapsto \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{y})$ is injective for all $\mathbf{x} \in \mathcal{X}$ [41, Definition 1.16].

Note that, while the term gx_3 in (2.2) is physically motivated, it could be replaced by $\phi(\mathbf{x})$ for any function $\phi \in \mathcal{C}^2(\mathcal{X})$ such that $\partial_{x_3}\phi \neq 0$ (so that c is twisted). In [20] the existence of weak solutions of the compressible SG equations is proved in this more general setting, while the assumptions on the initial data in [20], as well as the subsequent [19] and [26], are more restrictive than ours.

Definition 2.2 (Optimal transport cost)**.** The optimal transport cost from $\sigma \in \mathscr{P}(\mathcal{X})$ to $\alpha \in \mathscr{P}(\mathcal{Y})$ for the cost function c is defined by

$$\mathcal{T}_c(\sigma, \alpha) := \inf_{\substack{T:\mathcal{X}\to\mathcal{Y} \\ T_{\#}\sigma = \alpha}} \int_{\mathcal{X}} c(\mathbf{x}, T(\mathbf{x}))\, \mathrm{d}\sigma(\mathbf{x}). \tag{2.3}$$

If $\sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})$ and $\alpha \in \mathscr{P}_{\mathrm{c}}(\mathcal{Y})$, then there exists a unique map $T : \mathcal{X} \to \mathcal{Y}$ achieving the minimum in (2.3); see, for example, [41]. We call T the *optimal transport map* from σ to α for the cost c.

Definition 2.3 (Geostrophic energy)**.** We define the compressible geostrophic energy functional $E : \mathscr{P}(\mathcal{X}) \times \mathscr{P}_{\mathrm{c}}(\mathcal{Y}) \to \mathbb{R} \cup \{+\infty\}$ by

$$E(\sigma, \alpha) = \mathcal{T}_c(\sigma, \alpha) + F(\sigma), \tag{2.4}$$

where $F : \mathscr{P}(\mathcal{X}) \to \mathbb{R} \cup \{+\infty\}$ is defined by

$$F(\sigma) = \begin{cases} \kappa \displaystyle\int_{\mathcal{X}} \sigma(\mathbf{x})^\gamma \, \mathrm{d}\mathbf{x} & \text{if } \sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X}), \\ +\infty & \text{otherwise,} \end{cases}$$

where $\gamma \in (1, 2)$ and $\kappa > 0$ are constants.

Lemma 2.4 (Existence and uniqueness of minimisers of E)**.** *Given $\alpha \in \mathscr{P}_{\mathrm{c}}(\mathcal{Y})$, the functional on $\mathscr{P}(\mathcal{X})$ defined by $\sigma \mapsto E(\sigma, \alpha)$ is strictly convex and has a unique minimiser $\sigma_* \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})$.*

Proof. This is proved in [20, Theorem 4.1] for the special case $\alpha \in \mathscr{P}_{\mathrm{ac}}(\mathcal{Y})$. The functional $\mathscr{P}(\mathcal{X}) \ni \sigma \mapsto \mathcal{T}_c(\sigma, \alpha)$ is continuous and convex [41, Propositions 7.4 & 7.17], and F is lower semi-continuous and strictly convex [41, Proposition 7.7 & Section 7.2]. Therefore E is lower semi-continuous and strictly convex. Since $\mathcal{X}$ is compact, $\mathscr{P}(\mathcal{X})$ is compact with respect to weak convergence of measures. Therefore $E(\cdot, \alpha)$ has a minimiser σ_*, which belongs to $\mathscr{P}_{\mathrm{ac}}(\mathcal{X})$ because $E(\cdot, \alpha)$ takes the value $+\infty$ on $\mathscr{P}(\mathcal{X}) \setminus \mathscr{P}_{\mathrm{ac}}(\mathcal{X})$. Since E is strictly convex, its minimiser is unique. □

Definition 2.5 (Optimal source measure and transport map)**.** Given $\alpha \in \mathscr{P}_{\mathrm{c}}(\mathcal{Y})$, we define the map $\sigma_* : \mathscr{P}_{\mathrm{c}}(\mathcal{Y}) \to \mathscr{P}_{\mathrm{ac}}(\mathcal{X})$ by

$$\sigma_*[\alpha] = \operatorname*{arg\,min}_{\sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})} E(\sigma, \alpha). \tag{2.5}$$

We denote by $T_\alpha : \mathcal{X} \to \mathcal{Y}$ the optimal transport map from $\sigma_*[\alpha]$ to α for the cost c.

The formulation of the compressible SG equations given in (1.1) and (1.4) is formal because the inverse of T_{α_t} may not be well defined. For example, if α_t is discrete, then T_{α_t} sends a set of positive Lebesgue measure to each point in the support of α_t, violating injectivity. We therefore introduce a suitable weak formulation of (1.1).

To derive the weak formulation, we assume provisionally that $T_{\alpha_t}^{-1}$ exists and that all functions appearing in (1.1) are sufficiently regular. Multiplying (1.1) by $\varphi \in \mathscr{D}(\mathcal{Y} \times \mathbb{R})$ and integrating by parts over the domain $\mathcal{Y} \times [0, \tau]$ yields

$$\int_0^\tau \int_{\mathcal{Y}} \big(\partial_t \varphi_t(\mathbf{y}) + J\big(\mathbf{y} - T_{\alpha_t}^{-1}(\mathbf{y})\big) \cdot \nabla \varphi_t(\mathbf{y})\big) \,\mathrm{d}\alpha_t(\mathbf{y})\mathrm{d}t = \int_{\mathcal{Y}} \varphi_\tau(\mathbf{y}) \,\mathrm{d}\alpha_\tau(\mathbf{y}) - \int_{\mathcal{Y}} \varphi_0(\mathbf{y}) \,\mathrm{d}\alpha_0(\mathbf{y}),$$

where we have used the notation $\varphi_t(\mathbf{y}) := \varphi(\mathbf{y}, t)$. By using the push-forward constraint $(T_{\alpha_t})_{\#}\sigma_*[\alpha_t] = \alpha_t$ to rewrite the term involving $T_{\alpha_t}^{-1}$, we arrive at the following weak formulation of the compressible SG equations.

Definition 2.6 (Weak solution)**.** We say that $\alpha \in \mathcal{C}([0, \tau]; \mathscr{P}_{\mathrm{c}}(\mathcal{Y}))$ is a *weak solution* of the compressible SG equation (1.1) with initial condition $\alpha_0 = \overline{\alpha} \in \mathscr{P}_{\mathrm{c}}(\mathcal{Y})$ if for all $\varphi \in \mathscr{D}(\mathcal{Y} \times \mathbb{R})$,

$$\int_0^\tau \int_{\mathcal{Y}} \left[\partial_t \varphi_t(\mathbf{y}) + (J\mathbf{y}) \cdot \nabla \varphi_t(\mathbf{y})\right] \mathrm{d}\alpha_t(\mathbf{y}) \,\mathrm{d}t - \int_0^\tau \int_{\mathcal{X}} (J\mathbf{x}) \cdot \nabla \varphi_t(T_{\alpha_t}(\mathbf{x})) \,\mathrm{d}\sigma_*[\alpha_t](\mathbf{x}) \,\mathrm{d}t$$
$$= \int_{\mathcal{Y}} \varphi_\tau(\mathbf{y}) \,\mathrm{d}\alpha_\tau(\mathbf{y}) - \int_{\mathcal{Y}} \varphi_0(\mathbf{y}) \,\mathrm{d}\overline{\alpha}(\mathbf{y}). \quad (2.6)$$

Remark 2.7 (Equivalent weak formulation)**.** Following [7, 12], the nonlinear term in (2.6) can be rewritten as a term that is *linear* in the optimal transport plan:

$$\int_0^\tau \int_{\mathcal{X}} (J\mathbf{x}) \cdot \nabla \varphi_t(T_{\alpha_t}(\mathbf{x})) \,\mathrm{d}\sigma_*[\alpha_t](\mathbf{x}) \,\mathrm{d}t = \int_0^\tau \int_{\mathcal{X}\times\mathcal{Y}} (J\mathbf{x}) \cdot \nabla \varphi_t(\mathbf{y}) \,\mathrm{d}\gamma[\alpha_t](\mathbf{x}, \mathbf{y}) \,\mathrm{d}t, \quad (2.7)$$

where $\gamma[\alpha_t] = (\mathrm{id}_{\mathcal{X}} \times T_{\alpha_t})_{\#}\sigma_*[\alpha_t]$ is the optimal plan for transporting $\sigma_*[\alpha_t]$ to α_t.

2.3. **Semi-discrete optimal transport.** In this section we summarise the definitions and results from semi-discrete optimal transport theory used throughout this paper. Here $N \in \mathbb{N}$ is any natural number.

- The set D^N of *seed vectors* $\mathbf{z}$ is given by
$$D^N := \left\{\mathbf{z} = \left(\mathbf{z}^1, \ldots, \mathbf{z}^N\right) \in \mathcal{Y}^N : \mathbf{z}^i \neq \mathbf{z}^j \text{ whenever } i \neq j\right\}.$$
- We define
$$\Delta^N := \left\{\mathbf{m} = (m^1, \ldots, m^N) \in (0,1)^N : \sum_{i=1}^N m^i = 1\right\}.$$
- The class of discrete probability measures with N masses is
$$\mathscr{P}^N(\mathcal{Y}) := \left\{\sum_{i=1}^N m^i \delta_{\mathbf{z}^i} : \mathbf{z} \in D^N,\ \mathbf{m} \in \Delta^N\right\}. \quad (2.8)$$
- We say that a discrete probability measure $\displaystyle\sum_{i=1}^N m^i \delta_{\mathbf{z}^i} \in \mathscr{P}^N(\mathcal{Y})$ is *well-prepared* if
$$\mathbf{z} \in D_0^N \text{ where } D_0^N := \left\{\mathbf{z} \in D^N \,:\, z_3^i \neq z_3^j \ \forall\, i, j \in \{1, \ldots, N\},\ i \neq j\right\}. \quad (2.9)$$
In other words, a discrete measure is well-prepared if the seeds $\mathbf{z}^i$ lie in distinct horizontal planes. It will be crucial to assume that the initial data is well-prepared to prove existence and uniqueness for the ODE (1.6); see Proposition 4.7. Solutions of the ODE (1.6) satisfy $z_3^i(t) \neq z_3^j(t)$ for all $i \neq j$ and all t, provided that $z_3^i(0) \neq z_3^j(0)$ (because $\dot{z}_3^k = 0$ for all k). Consequently the particles $\mathbf{z}^i$ cannot collide, which would preclude global existence of solutions because the vector field in the ODE (1.6) is not defined if the seeds are not distinct (the centroid map is not defined).

Definition 2.8 (c-Laguerre tessellation)**.** Given $(\boldsymbol{w}, \mathbf{z}) \in \mathbb{R}^N \times D^N$, the *$c$-Laguerre tessellation of $\mathcal{X}$ generated by* $(\boldsymbol{w}, \mathbf{z})$ is the collection of *Laguerre cells* $\{L_c^i(\boldsymbol{w}, \mathbf{z})\}_{i=1}^N$ defined by

$$L_c^i(\boldsymbol{w}, \mathbf{z}) := \left\{\mathbf{x} \in \mathcal{X} : c(\mathbf{x}, \mathbf{z}^i) - w^i \leq c(\mathbf{x}, \mathbf{z}^j) - w^j \ \forall\, j \in \{1, \ldots, N\}\right\}, \quad i \in \{1, \ldots, N\}.$$



(A) 5 Laguerre cells



(B) 10 Laguerre cells



(C) 25 Laguerre cells



(D) 50 Laguerre cells

FIGURE 2. c-Laguerre tessellations (see Definition 2.8) in the (x_1, x_3)-plane for the cost function $c = c_{2\mathrm{d}}$ (see (2.10)). The colours distinguish the cells. For each plot, $\mathcal{X} = [0,1]^2$, $f_{\mathrm{cor}} = 1$, $g = 1$, the seeds $\mathbf{z}^i$ were sampled uniformly from $\mathcal{X}$, and the weights w^i were chosen so that the cells have equal area (by maximising the dual function as in (2.13)).

The c-Laguerre cells form a tessellation of $\mathcal{X}$ in the sense that $\bigcup_{i=1}^N L_c^i(\boldsymbol{w}, \mathbf{z}) = \mathcal{X}$ and $\mathcal{L}^3\big(L_c^i(\boldsymbol{w}, \mathbf{z}) \cap L_c^j(\boldsymbol{w}, \mathbf{z})\big) = 0$ if $i \neq j$ (by [40, Proposition 37]). Intuitively, changing the *weight vector* $\boldsymbol{w} = (w^1, \dots, w^N)$ shifts the boundaries of the Laguerre cells, thus controlling their volumes. It is easy to check that if $\mathbf{z} \in D_0^N$, then each Laguerre cell $L_c^i(\boldsymbol{w}, \mathbf{z})$ is the intersection of $\mathcal{X}$ with $N-1$ paraboloids of the form

$$\left\{\mathbf{x} \in \mathbb{R}^3 : x_3 \leq -\frac{f_{\mathrm{cor}}^2}{2g}\left((x_1 - a_1)^2 + (x_2 - a_2)^2\right) + b\right\},$$

where $a_1, a_2 \in \mathbb{R}$ depend on $\mathbf{z}$ and $b \in \mathbb{R}$ depends on $(\boldsymbol{w}, \mathbf{z})$. In Figure 2 we plot c-Laguerre tessellations in two dimensions for the following 2D version of the cost (2.2):

$$c_{2\mathrm{d}}((x_1, x_3), (y_1, y_3)) = \frac{1}{y_3}\left(\frac{f_{\mathrm{cor}}^2}{2}(x_1 - y_1)^2 + g x_3\right). \tag{2.10}$$

We recall the following well-known result in semi-discrete optimal transport theory, which we will use frequently throughout the paper. For a proof see for example [40, Section 4].

Theorem 2.9 (The optimal transport map for semi-discrete optimal transport)**.** *Let $\sigma \in \mathcal{P}_{\mathrm{ac}}(\mathcal{X})$ and $\alpha^N = \sum_{i=1}^N m^i \delta_{\mathbf{z}^i} \in \mathscr{P}^N(\mathcal{Y})$. There there exists $\boldsymbol{w} \in \mathbb{R}^N$ satisfying*

$$\sigma(L_c^i(\boldsymbol{w}, \mathbf{z})) = m^i \quad \forall\, i \in \{1, \dots, N\}. \tag{2.11}$$

Given $\boldsymbol{w}$ satisfying (2.11)*, define $T : \mathcal{X} \to \mathcal{Y}$ by*

$$T = \sum_{i=1}^N \mathbf{z}^i \mathbb{1}_{L_c^i(\boldsymbol{w}, \mathbf{z})}. \tag{2.12}$$

In other words, T is the piecewise constant function given by $T(\mathbf{x}) = \mathbf{z}^i$ if $\mathbf{x} \in L_c^i(\boldsymbol{w}, \mathbf{z})$. Then $T : \mathcal{X} \to \mathcal{Y}$ is the unique optimal transport map from σ to α^N for the cost c. Moreover, the following Kantorovich duality holds:

$$\mathcal{T}_c(\sigma, \alpha^N) = \max_{\widetilde{\boldsymbol{w}} \in \mathbb{R}^N} \sum_{i=1}^N \left(m^i \widetilde{w}^i + \int_{L_c^i(\widetilde{\boldsymbol{w}}, \mathbf{z})} \big(c(\mathbf{x}, \mathbf{z}^i) - \widetilde{w}^i\big) \, \mathrm{d}\sigma(\mathbf{x}) \right). \tag{2.13}$$

Furthermore, $\boldsymbol{w} \in \mathbb{R}^N$ satisfies (2.11) *if and only if it maximises the dual function in* (2.13)*.*

Note that the solution $\boldsymbol{w} \in \mathbb{R}^N$ of (2.11) is not unique because $L_c^i(\boldsymbol{w} + \boldsymbol{a}, \mathbf{z}) = L_c^i(\boldsymbol{w}, \mathbf{z})$ for any constant vector $\boldsymbol{a} = (a, \dots, a) \in \mathbb{R}^N$. Also note that the map T defined in (2.12) is well-defined $\mathcal{L}^3$-almost everywhere because the intersection of any two c-Laguerre cells has zero Lebesgue measure by [40, Proposition 37] and the twist condition on c.

3. The dual problem

In this section we derive a dual formulation of the minimisation problem (2.5) for the case where α is a discrete measure. Additionally, we prove necessary and sufficient optimality conditions that characterise the optimal source measure $\sigma_*[\alpha]$ in terms of the solution of the dual problem. The main results are contained in Theorem 3.4.

Given $(\boldsymbol{w}, \mathbf{z}) \in \mathbb{R}^N \times D^N$, recall that the c-transform of $\boldsymbol{w}$ is the function $\boldsymbol{w}^c(\,\cdot\,; \mathbf{z}) : \mathcal{X} \to \mathbb{R}$ defined by

$$\boldsymbol{w}^c(\mathbf{x}; \mathbf{z}) = \min_{i \in \{1, \dots, N\}} \big\{ c(\mathbf{x}, \mathbf{z}^i) - w^i \big\}.$$

In particular, $\boldsymbol{w}^c(\mathbf{x}; \mathbf{z}) = c(\mathbf{x}, \mathbf{z}^i) - w^i$ if and only if $\mathbf{x} \in L_c^i(\boldsymbol{w}, \mathbf{z})$.

Definition 3.1 (Dual functional)**.** Given $\mathbf{m} \in \Delta^N$, define the dual functional $\mathcal{G} : \mathbb{R}^N \times D^N \to \mathbb{R}$ by

$$\mathcal{G}(\boldsymbol{w}, \mathbf{z}) = \sum_{i=1}^N m^i w^i - \int_{\mathcal{X}} f^*(-\boldsymbol{w}^c(\mathbf{x}; \mathbf{z})) \, \mathrm{d}\mathbf{x} = \sum_{i=1}^N \left(m^i w^i - \int_{L_c^i(\boldsymbol{w}, \mathbf{z})} f^*(w^i - c(\mathbf{x}, \mathbf{z}^i)) \, \mathrm{d}\mathbf{x} \right),$$

where $f^* : \mathbb{R} \to \mathbb{R}$ is the Legendre-Fenchel transform of the function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(s) = \begin{cases} \kappa s^\gamma & \text{if } s \geq 0, \\ +\infty & \text{if } s < 0, \end{cases}$$

where $\gamma \in (1, 2)$ and $\kappa > 0$.

Remark 3.2 (Properties of f)**.** The function f is strictly convex on $[0, \infty)$, lower semi-continuous, and its Legendre-Fenchel transform is given by

$$f^*(t) \coloneqq \sup_{s \in \mathbb{R}} (st - f(s)) = \begin{cases} \frac{1}{\gamma'} (\kappa\gamma)^{1-\gamma'} t^{\gamma'} & \text{if } t > 0, \\ 0 & \text{if } t \leq 0, \end{cases}$$

where $\gamma' \in (2, \infty)$ satisfies $\frac{1}{\gamma} + \frac{1}{\gamma'} = 1$. Note that $f^* \in \mathcal{C}^2(\mathbb{R})$.

Remark 3.3 (Derivation of the dual functional)**.** We give a short derivation of the dual functional before stating a rigorous duality theorem below (see Theorem 3.4). Here we identify vectors $\boldsymbol{w} \in \mathbb{R}^N$ with real-valued functions on the discrete set $\{\mathbf{z}^i\}_{i=1}^N$ via $\boldsymbol{w}(\mathbf{z}^i) = w^i$. Applying the standard Kantorovich Duality Theorem from optimal transport theory gives

$$\begin{aligned}
\inf_{\sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})} E(\sigma, \alpha^N) &= \inf_{\sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})} \Big(\mathcal{T}_c(\sigma, \alpha^N) + F(\sigma) \Big) \\
&= \inf_{\sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})} \left(\sup_{\boldsymbol{w} \in \mathbb{R}^N} \left(\int_{\mathcal{X}} \boldsymbol{w}^c \, \mathrm{d}\sigma + \int_{\mathcal{Y}} \boldsymbol{w} \, \mathrm{d}\alpha^N \right) + F(\sigma) \right) \\
&\geq \sup_{\boldsymbol{w} \in \mathbb{R}^N} \inf_{\sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})} \left(\int_{\mathcal{X}} \boldsymbol{w}^c \, \mathrm{d}\sigma + \int_{\mathcal{Y}} \boldsymbol{w} \, \mathrm{d}\alpha^N + F(\sigma) \right) \\
&= \sup_{\boldsymbol{w} \in \mathbb{R}^N} \left(\int_{\mathcal{Y}} \boldsymbol{w} \, \mathrm{d}\alpha^N - \sup_{\sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})} \left(\int_{\mathcal{X}} (-\boldsymbol{w}^c) \, \mathrm{d}\sigma - F(\sigma) \right) \right) \\
&= \sup_{\boldsymbol{w} \in \mathbb{R}^N} \left(\sum_{i=1}^N m^i w^i - \int_{\mathcal{X}} f^*(-\boldsymbol{w}^c) \, \mathrm{d}\mathbf{x} \right) \\
&= \sup_{\boldsymbol{w} \in \mathbb{R}^N} \mathcal{G}(\boldsymbol{w}, \mathbf{z}).
\end{aligned}$$

In Theorem 3.4 we prove that the inf on the left-hand side is in fact a min, the sup on the right-hand side is in fact a max, and the inequality is an equality.

Theorem 3.4 (Duality Theorem)**.** *Let $\mathbf{z} \in D^N$, $\mathbf{m} \in \Delta^N$ and define*

$$\alpha^N := \sum_{i=1}^N m^i \delta_{\mathbf{z}^i}.$$

Assume that $z_3^i \neq z_3^j$ for all $i, j \in \{1, \dots, N\}$ with $i \neq j$. Then the map $\sigma \mapsto E(\sigma, \alpha^N)$ is strictly convex and has a unique minimiser, and the map $\boldsymbol{w} \mapsto \mathcal{G}(\boldsymbol{w}, \mathbf{z})$ is concave and has a unique maximiser. Moreover,

$$\min_{\sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})} E(\sigma, \alpha^N) = \max_{\boldsymbol{w} \in \mathbb{R}^N} \mathcal{G}(\boldsymbol{w}, \mathbf{z}). \tag{3.1}$$

Furthermore, $\sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})$ minimises $E(\cdot, \alpha^N)$ and $\boldsymbol{w} \in \mathbb{R}^N$ maximises $\mathcal{G}(\cdot, \mathbf{z})$ if and only if

$$m^i = \int_{L_c^i(\boldsymbol{w}, \mathbf{z})} (f^*)'(w^i - c(\mathbf{x}, \mathbf{z}^i)) \, \mathrm{d}\mathbf{x} \quad \forall \, i \in \{1, \dots, N\}, \tag{3.2}$$

$$\sigma(\mathbf{x}) = (f^*)'(-\boldsymbol{w}^c(\mathbf{x}; \mathbf{z})) \quad \textit{for } \mathcal{L}^3\textit{-almost every } \mathbf{x} \in \mathcal{X}. \tag{3.3}$$

Theorem 3.4 is analogous to the duality result given in [42, Proposition 12] but with the cost c rather than the quadratic cost. Other closely related results include [38, Proposition 5.3] and [9, Theorems 3.1 and 3.2]. In addition, the Euler-Lagrange equation (3.3) can be derived from [41, Proposition 7.20].

Remark 3.5 (Optimal transport map)**.** If $\sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})$ minimises $E(\cdot, \alpha^N)$ and $\boldsymbol{w} \in \mathbb{R}^N$ maximises $\mathcal{G}(\cdot, \mathbf{z})$, then the optimality conditions (3.2) and (3.3) imply that the mass constraint (2.11) holds. In particular, this means that the map T given by (2.12) is the optimal transport map from σ to α^N for the cost c.

We begin by establishing the regularity of $\mathcal{G}$; see Proposition 3.8. To do this we use the following two lemmas, which will also be used in Section 4 to prove that the centroid map is continuously differentiable.

Lemma 3.6 (Regularity of integrals over Laguerre cells)**.** *Let $U \subset \mathbb{R}^N \times D_0^N$ be the open set*

$$U = \left\{ (\boldsymbol{w}, \mathbf{z}) \in \mathbb{R}^N \times D_0^N \, : \, \mathcal{L}^3(L_c^i(\boldsymbol{w}, \mathbf{z})) > 0 \; \forall \, i \in \{1, \dots, N\} \right\}.$$

Define $\boldsymbol{\Psi} = (\Psi^1, \dots, \Psi^N) : U \to \mathbb{R}^N$ *by*

$$\Psi^i(\boldsymbol{w}, \mathbf{z}) := \int_{L_c^i(\boldsymbol{w},\mathbf{z})} \zeta(\mathbf{x}, \boldsymbol{w}, \mathbf{z}) \, \mathrm{d}\mathbf{x},$$

where $\zeta : \mathcal{X} \times \mathbb{R}^N \times D^N \to \mathbb{R}$ *satisfies the following:*

- $\zeta(\cdot, \boldsymbol{w}, \mathbf{z}) \in \mathcal{C}(\mathcal{X})$ *for all* $(\boldsymbol{w}, \mathbf{z}) \in \mathbb{R}^N \times D^N$*;*
- *for each compact set* $K \subset \mathbb{R}^N \times D^N$, $\zeta(\mathbf{x}, \cdot, \cdot) \in \mathcal{C}^{0,1}(K)$ *for all* $\mathbf{x} \in \mathcal{X}$ *and there exists a constant* $L(K) > 0$ *such that* $\sup_{\mathbf{x}\in\mathcal{X}} \operatorname{Lip}(\zeta(\mathbf{x}, \cdot, \cdot)|_K) \le L(K)$*;*
- *for each* $(\boldsymbol{w}_0, \mathbf{z}_0) \in \mathbb{R}^N \times D^N$*, there exists an open set* $S(\boldsymbol{w}_0, \mathbf{z}_0) \subseteq \mathcal{X}$ *with* $\mathcal{L}^3(\mathcal{X} \setminus S(\boldsymbol{w}_0, \mathbf{z}_0)) = 0$ *such that the partial derivatives* $\partial\zeta/\partial\boldsymbol{w}(\mathbf{x}, \cdot, \cdot)$, $\partial\zeta/\partial\mathbf{z}(\mathbf{x}, \cdot, \cdot)$ *exist and are continuous at* $(\boldsymbol{w}_0, \mathbf{z}_0)$ *for all* $\mathbf{x} \in S(\boldsymbol{w}_0, \mathbf{z}_0)$*.*

Then $\boldsymbol{\Psi}$ *is continuously differentiable. Moreover, for all* $i, j \in \{1, \dots, N\}$, $i \neq j$,

$$\frac{\partial\Psi^i}{\partial w^j}(\boldsymbol{w}, \mathbf{z}) = \int_{L_c^i(\boldsymbol{w},\mathbf{z})} \frac{\partial\zeta}{\partial w^j}(\mathbf{x}, \boldsymbol{w}, \mathbf{z}) \, \mathrm{d}\mathbf{x} - \int_{L_c^i(\boldsymbol{w},\mathbf{z}) \cap L_c^j(\boldsymbol{w},\mathbf{z})} \frac{\zeta(\mathbf{x}, \boldsymbol{w}, \mathbf{z})}{\|\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^i) - \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^j)\|} \, \mathrm{d}\mathcal{H}^2(\mathbf{x}),$$

$$\frac{\partial\Psi^i}{\partial \mathbf{z}^j}(\boldsymbol{w}, \mathbf{z}) = \int_{L_c^i(\boldsymbol{w},\mathbf{z})} \frac{\partial\zeta}{\partial \mathbf{z}^j}(\mathbf{x}, \boldsymbol{w}, \mathbf{z}) \, \mathrm{d}\mathbf{x} + \int_{L_c^i(\boldsymbol{w},\mathbf{z}) \cap L_c^j(\boldsymbol{w},\mathbf{z})} \frac{\nabla_{\mathbf{y}} c(\mathbf{x}, \mathbf{z}^j) \, \zeta(\mathbf{x}, \boldsymbol{w}, \mathbf{z})}{\|\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^i) - \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^j)\|} \, \mathrm{d}\mathcal{H}^2(\mathbf{x}),$$

and

$$\frac{\partial\Psi^i}{\partial w^i}(\boldsymbol{w}, \mathbf{z}) = \int_{L_c^i(\boldsymbol{w},\mathbf{z})} \frac{\partial\zeta}{\partial w^i}(\mathbf{x}, \boldsymbol{w}, \mathbf{z}) \, \mathrm{d}\mathbf{x} + \sum_{j\neq i} \int_{L_c^i(\boldsymbol{w},\mathbf{z}) \cap L_c^j(\boldsymbol{w},\mathbf{z})} \frac{\zeta(\mathbf{x}, \boldsymbol{w}, \mathbf{z})}{\|\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^i) - \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^j)\|} \, \mathrm{d}\mathcal{H}^2(\mathbf{x}),$$

$$\frac{\partial\Psi^i}{\partial \mathbf{z}^i}(\boldsymbol{w}, \mathbf{z}) = \int_{L_c^i(\boldsymbol{w},\mathbf{z})} \frac{\partial\zeta}{\partial \mathbf{z}^i}(\mathbf{x}, \boldsymbol{w}, \mathbf{z}) \, \mathrm{d}\mathbf{x} - \sum_{j\neq i} \int_{L_c^i(\boldsymbol{w},\mathbf{z}) \cap L_c^j(\boldsymbol{w},\mathbf{z})} \frac{\nabla_{\mathbf{y}} c(\mathbf{x}, \mathbf{z}^j) \, \zeta(\mathbf{x}, \boldsymbol{w}, \mathbf{z})}{\|\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^i) - \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^j)\|} \, \mathrm{d}\mathcal{H}^2(\mathbf{x}).$$

Proof. The proof is rather technical and is given in Appendix B. □

Note that the set U in the statement of Lemma 3.6 is open because the map $(\boldsymbol{w}, \mathbf{z}) \mapsto \mathcal{L}^3(L_c^i(\boldsymbol{w}, \mathbf{z}))$ is continuous. (Continuity of this map with respect to $\boldsymbol{w}$ is proved for example in [40, Proposition 38(vii)]. The proof of continuity with respect to $(\boldsymbol{w}, \mathbf{z})$ is identical.) The constraint that the seeds in U satisfy $|z_3^i - z_3^j| > 0$ is not necessary, but it is sufficient for our purposes and it slightly simplifies the proof; see Remark B.2.

Lemma 3.7. *Define* $\zeta : \mathcal{X} \times \mathbb{R}^N \times D^N \to \mathbb{R}$ *by*

$$\zeta(\mathbf{x}, \boldsymbol{w}, \mathbf{z}) := (f^*)'(-\boldsymbol{w}^c(\mathbf{x}; \mathbf{z})).$$

Then ζ *satisfies the assumptions of Lemma 3.6.*

Proof. Since $(f^*)'$ is non-decreasing, we have that

$$\zeta(\mathbf{x}, \boldsymbol{w}, \mathbf{z}) = (f^*)'\Big(- \min_{i\in\{1,\dots,N\}} (c(\mathbf{x}, \mathbf{z}^i) - w^i)\Big) = \max_{i\in\{1,\dots,N\}} (f^*)'(w^i - c(\mathbf{x}, \mathbf{z}^i)).$$

For each $i \in \{1, \dots, N\}$, define $\zeta^i : \mathcal{X} \times \mathbb{R}^N \times D^N \to \mathbb{R}$ by

$$\zeta^i(\mathbf{x}, \boldsymbol{w}, \mathbf{z}) = (f^*)'(w^i - c(\mathbf{x}, \mathbf{z}^i)). \tag{3.4}$$

Then ζ^i is continuously differentiable because $(f^*)'$ and c are continuously differentiable.

Fix $(\boldsymbol{w}, \mathbf{z}) \in \mathbb{R}^N \times D^N$ and consider the function $\zeta(\cdot, \boldsymbol{w}, \mathbf{z}) = \max_i \zeta^i(\cdot, \boldsymbol{w}, \mathbf{z})$. For each $i \in \{1, \dots, N\}$ the function $\zeta^i(\cdot, \boldsymbol{w}, \mathbf{z})$ is globally Lipschitz on $\mathcal{X}$ because ζ^i is continuously differentiable and $\mathcal{X}$ is compact. Therefore $\zeta(\cdot, \boldsymbol{w}, \mathbf{z})$ is the pointwise maximum of a finite family of Lipschitz functions, hence Lipschitz. In particular, it is continuous.

Now we consider the functions $\zeta(\mathbf{x}, \cdot, \cdot) = \max_i \zeta^i(\mathbf{x}, \cdot, \cdot)$ indexed by $\mathbf{x} \in \mathcal{X}$. Let $K \subset \mathbb{R}^N \times D^N$ be compact. As above, $\zeta(\mathbf{x}, \cdot, \cdot)|_K$ is the pointwise maximum of a finite family of of Lipschitz functions, hence Lipschitz. Moreover, the Lipschitz constant of $\zeta^i(\mathbf{x}, \cdot, \cdot)|_K$ can be bounded from above by a constant $L^i(K)$ independent of $\mathbf{x}$ since $\partial\zeta^i/\partial\boldsymbol{w}$, $\partial\zeta^i/\partial\mathbf{z}$ are continuous, and $\mathcal{X}$ and K are compact. Therefore $\sup_{\mathbf{x}\in\mathcal{X}} \operatorname{Lip}(\zeta(\mathbf{x}, \cdot, \cdot)|_K) \le \max_i L^i(K) =: L(K)$, as required.

Let $(\boldsymbol{w}_0, \mathbf{z}_0) \in \mathbb{R}^N \times D^N$ and define $S(\boldsymbol{w}_0, \mathbf{z}_0) = \bigcup_{j=1}^N \operatorname{int}(L_c^j(\boldsymbol{w}_0, \mathbf{z}_0))$. Since the cost c is twisted, $\mathcal{L}^3(\mathcal{X} \setminus S(\boldsymbol{w}_0, \mathbf{z}_0)) = 0$ by [40, Proposition 37]. For all $\mathbf{x} \in S(\boldsymbol{w}_0, \mathbf{z}_0)$, we have

$$\frac{\partial \zeta}{\partial w^i}(\mathbf{x}, \boldsymbol{w}_0, \mathbf{z}_0) = \begin{cases} (f^*)''(w_0^i - c(\mathbf{x}, \mathbf{z}_0^i)) & \text{if } \mathbf{x} \in \operatorname{int}(L_c^i(\boldsymbol{w}_0, \mathbf{z}_0)), \\ 0 & \text{if } \mathbf{x} \in \operatorname{int}(L_c^j(\boldsymbol{w}_0, \mathbf{z}_0)) \text{ for } j \neq i, \end{cases}$$

$$\frac{\partial \zeta}{\partial \mathbf{z}^i}(\mathbf{x}, \boldsymbol{w}_0, \mathbf{z}_0) = \begin{cases} -\nabla_{\mathbf{y}} c(\mathbf{x}, \mathbf{z}_0^i)(f^*)''(w_0^i - c(\mathbf{x}, \mathbf{z}_0^i)) & \text{if } \mathbf{x} \in \operatorname{int}(L_c^i(\boldsymbol{w}_0, \mathbf{z}_0)), \\ 0 & \text{if } \mathbf{x} \in \operatorname{int}(L_c^j(\boldsymbol{w}_0, \mathbf{z}_0)) \text{ for } j \neq i. \end{cases}$$

(Note that the partial derivatives do not exist if $\mathbf{x} \in L_c^i(\boldsymbol{w}_0, \mathbf{z}_0) \cap L_c^j(\boldsymbol{w}_0, \mathbf{z}_0)$ for $j \neq i$.) In particular, for all $\mathbf{x} \in S(\boldsymbol{w}_0, \mathbf{z}_0)$, the partial derivatives $\partial\zeta/\partial \boldsymbol{w}(\mathbf{x}, \cdot, \cdot)$, $\partial\zeta/\partial \mathbf{z}(\mathbf{x}, \cdot, \cdot)$ exist at $(\boldsymbol{w}_0, \mathbf{z}_0)$. Now we prove that the partial derivatives are continuous. Let $\mathbf{x} \in S(\boldsymbol{w}_0, \mathbf{z}_0)$. Then there exists $i \in \{1, \dots, N\}$ such that $\mathbf{x} \in \operatorname{int}(L_c^i(\boldsymbol{w}_0, \mathbf{z}_0))$, so

$$c(\mathbf{x}, \mathbf{z}_0^i) - w_0^i < c(\mathbf{x}, \mathbf{z}_0^j) - w_0^j \quad \forall\, j \in \{1, \dots, N\} \setminus \{i\}.$$

Since the inequality is strict and c is continuous, $\mathbf{x} \in \operatorname{int}(L_c^i(\boldsymbol{w}, \mathbf{z}))$ for all $(\boldsymbol{w}, \mathbf{z})$ sufficiently close to $(\boldsymbol{w}_0, \mathbf{z}_0)$, and

$$\frac{\partial \zeta}{\partial w^i}(\mathbf{x}, \boldsymbol{w}, \mathbf{z}) - \frac{\partial \zeta}{\partial w^i}(\mathbf{x}, \boldsymbol{w}_0, \mathbf{z}_0) = \begin{cases} (f^*)''(w^i - c(\mathbf{x}, \mathbf{z}^i)) - (f^*)''(w_0^i - c(\mathbf{x}, \mathbf{z}_0^i)) & \text{if } \mathbf{x} \in \operatorname{int}(L_c^i(\boldsymbol{w}_0, \mathbf{z}_0)), \\ 0 & \text{if } \mathbf{x} \in \operatorname{int}(L_c^j(\boldsymbol{w}_0, \mathbf{z}_0)) \text{ for } j \neq i. \end{cases}$$

Since $(f^*)''$ and c are continuous, it follows that $\partial\zeta/\partial \boldsymbol{w}(\mathbf{x}, \cdot, \cdot)$ is continuous at $(\boldsymbol{w}_0, \mathbf{z}_0)$. Similarly, it can be shown that $\partial\zeta/\partial \mathbf{z}(\mathbf{x}, \cdot, \cdot)$ is continuous at $(\boldsymbol{w}_0, \mathbf{z}_0)$. This completes the proof. □

Proposition 3.8 (Regularity of $\mathcal{G}$)**.** *Let $\mathbf{z} \in D^N$ and $\boldsymbol{w} \in \mathbb{R}^N$. Then $\mathcal{G}(\cdot, \mathbf{z}) \in \mathcal{C}^1(\mathbb{R}^N)$, $\mathcal{G}(\boldsymbol{w}, \cdot) \in \mathcal{C}^1(D^N)$, and*

$$\frac{\partial \mathcal{G}}{\partial w^i}(\boldsymbol{w}, \mathbf{z}) = m^i - \int_{L_c^i(\boldsymbol{w}, \mathbf{z})} (f^*)'(w^i - c(\mathbf{x}, \mathbf{z}^i))\, \mathrm{d}\mathbf{x}, \tag{3.5}$$

$$\frac{\partial \mathcal{G}}{\partial \mathbf{z}^i}(\boldsymbol{w}, \mathbf{z}) = \int_{L_c^i(\boldsymbol{w}, \mathbf{z})} \nabla_{\mathbf{y}} c(\mathbf{x}, \mathbf{z}^i)\, (f^*)'(w^i - c(\mathbf{x}, \mathbf{z}^i))\, \mathrm{d}\mathbf{x}, \tag{3.6}$$

for all $i \in \{1, \dots, N\}$. Moreover, if $\mathbf{z} \in D^N$ satisfies $z_3^i \neq z_3^j$ for all $i, j \in \{1, \dots, N\}$, $i \neq j$, and $\boldsymbol{w} \in \mathbb{R}^N$ satisfies $\mathcal{L}^3(L_c^i(\boldsymbol{w}, \mathbf{z})) > 0$ for all $i \in \{1, \dots, N\}$, then $\mathcal{G}$ is twice continuously differentiable at $(\boldsymbol{w}, \mathbf{z})$, and the second-order partial derivatives of $\mathcal{G}$ with respect to $\boldsymbol{w}$ are

$$\begin{cases} \dfrac{\partial^2 \mathcal{G}}{\partial w^i \partial w^j}(\boldsymbol{w}, \mathbf{z}) & = \displaystyle\int_{L_c^i(\boldsymbol{w}, \mathbf{z}) \cap L_c^j(\boldsymbol{w}, \mathbf{z})} \frac{(f^*)'(w^i - c(\mathbf{x}, \mathbf{z}^i))}{\|\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^i) - \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^j)\|}\, \mathrm{d}\mathcal{H}^2(\mathbf{x}) \quad \text{for } j \neq i, \\ \dfrac{\partial^2 \mathcal{G}}{\partial w^i \partial w^i}(\boldsymbol{w}, \mathbf{z}) & = -\displaystyle\int_{L_c^i(\boldsymbol{w}, \mathbf{z})} (f^*)''(w^i - c(\mathbf{x}, \mathbf{z}^i))\, \mathrm{d}\mathbf{x} - \sum_{j \neq i} \frac{\partial^2 \mathcal{G}}{\partial w^i \partial w^j}(\boldsymbol{w}, \mathbf{z}). \end{cases} \tag{3.7}$$

If in addition

$$\int_{L_c^i(\boldsymbol{w}, \mathbf{z})} (f^*)'(w^i - c(\mathbf{x}, \mathbf{z}^i))\, \mathrm{d}\mathbf{x} > 0 \tag{3.8}$$

for all $i \in \{1, \dots, N\}$, then the Hessian matrix $D^2_{\boldsymbol{w}\boldsymbol{w}} \mathcal{G}(\boldsymbol{w}, \mathbf{z})$ is negative definite.

Proof. Fix $\mathbf{z}_0 \in D^N$. We begin by showing that $\mathcal{G}(\cdot, \mathbf{z}_0) \in \mathcal{C}^1(\mathbb{R}^N)$. Let $\boldsymbol{w}_0 \in \mathbb{R}^N$ and, as in the proof of Lemma 3.7, define the set

$$S(\boldsymbol{w}_0, \mathbf{z}_0) \coloneqq \bigcup_{j=1}^N \operatorname{int}(L_c^j(\boldsymbol{w}_0, \mathbf{z}_0)).$$

For any $\mathbf{x} \in S(\boldsymbol{w}_0, \mathbf{z}_0)$ there exists $i \in \{1, \dots, N\}$ and $\varepsilon > 0$ such that, for all $\boldsymbol{w} \in B_\varepsilon(\boldsymbol{w}_0)$,

$$c(\mathbf{x}, \mathbf{z}_0^i) - w^i < c(\mathbf{x}, \mathbf{z}_0^j) - w^j \quad \forall\, j \in \{1, \dots, N\}.$$

In particular, $\mathbf{x} \in L_c^i(\boldsymbol{w}, \mathbf{z}_0)$ for all $\boldsymbol{w} \in B_\varepsilon(\boldsymbol{w}_0)$, i.e.,

$$-\boldsymbol{w}^c(\mathbf{x}; \mathbf{z}) = w^i - c(\mathbf{x}, \mathbf{z}_0^i) \quad \forall\, \boldsymbol{w} \in B_\varepsilon(\boldsymbol{w}_0).$$

It follows that

$$\forall\, \mathbf{x} \in \operatorname{int}(L_c^i(\boldsymbol{w}_0, \mathbf{z}_0)), \quad \frac{\partial}{\partial w^j} f^*(-\boldsymbol{w}^c(\mathbf{x}; \mathbf{z}))\Big|_{\boldsymbol{w}=\boldsymbol{w}_0} = \begin{cases} (f^*)'(w_0^i - c(\mathbf{x}, \mathbf{z}_0^i)) & \text{if } j = i, \\ 0 & \text{otherwise.} \end{cases} \tag{3.9}$$

By continuity of $(f^*)'$,

$$\lim_{\boldsymbol{w} \to \boldsymbol{w}_0} (f^*)'(w^j - c(\mathbf{x}, \mathbf{z}_0^j)) \mathbb{1}_{L_c^j(\boldsymbol{w}, \mathbf{z}_0)}(\mathbf{x}) = (f^*)'(w_0^j - c(\mathbf{x}, \mathbf{z}_0^j)) \mathbb{1}_{L_c^j(\boldsymbol{w}_0, \mathbf{z}_0)}(\mathbf{x}) \tag{3.10}$$

for all $j \in \{1, \dots, N\}$.

As in the proof of Lemma 3.7, the map $\boldsymbol{w} \mapsto f^*(-\boldsymbol{w}^c(\mathbf{x}; \mathbf{z}))$ is locally Lipschitz continuous for all $\mathbf{x} \in \mathcal{X}$, and its Lipschitz constant can be bounded uniformly, independently of $\mathbf{x} \in \mathcal{X}$. Moreover, for all $\mathbf{x} \in S(\boldsymbol{w}_0, \mathbf{z}_0)$ this map is differentiable at $\boldsymbol{w}_0$ with derivative given by (3.9). Then, since $\mathcal{X}$ is compact and $\mathcal{X} \setminus S(\boldsymbol{w}_0, \mathbf{z}_0)$ is Lebesgue negligible, by the Dominated Convergence Theorem

$$\begin{aligned} \frac{\partial \mathcal{G}}{\partial w^j}(\boldsymbol{w}_0, \mathbf{z}_0) &= \frac{\partial}{\partial w^j} \left(\sum_{i=1}^N m^i w^i - \int_{\mathcal{X}} f^*(-\boldsymbol{w}^c(\mathbf{x}; \mathbf{z}))\, \mathrm{d}\mathbf{x} \right)\Bigg|_{\boldsymbol{w}=\boldsymbol{w}_0} \\ &= m^j - \int_{\mathcal{X}} \frac{\partial}{\partial w^j} f^*(-\boldsymbol{w}^c(\mathbf{x}; \mathbf{z}))\Big|_{\boldsymbol{w}=\boldsymbol{w}_0} \mathrm{d}\mathbf{x} \\ &= m^j - \int_{L_c^j(\boldsymbol{w}_0, \mathbf{z}_0)} (f^*)'(w_0^j - c(\mathbf{x}, \mathbf{z}_0^j))\, \mathrm{d}\mathbf{x}. \end{aligned}$$

This establishes (3.5). Similarly, using (3.10) and applying the Dominated Convergence Theorem proves continuity of the partial derivatives of $\mathcal{G}(\cdot, \mathbf{z}_0)$ at $\boldsymbol{w}_0$. Since $\boldsymbol{w}_0$ was arbitrary, it follows that $\mathcal{G}(\cdot, \mathbf{z}_0) \in \mathcal{C}^1(\mathbb{R}^N)$. Analogously, it can be proved that (3.6) holds and that $\mathcal{G}(\boldsymbol{w}_0, \cdot) \in \mathcal{C}^1(D^N)$ for all $\boldsymbol{w}_0 \in \mathbb{R}^N$; we omit the proof for brevity.

Now suppose that $\mathbf{z} \in D^N$ satisfies $z_3^i \neq z_3^j$ for all $i, j \in \{1, \dots, N\}$, $i \neq j$, and $\boldsymbol{w} \in \mathbb{R}^N$ satisfies $\mathcal{L}^3(L_c^i(\boldsymbol{w}, \mathbf{z})) > 0$ for all $i \in \{1, \dots, N\}$. By Lemmas 3.6 and 3.7, and equations (3.5) and (3.6), $\mathcal{G}$ is twice continuously differentiable in a neighbourhood of $(\boldsymbol{w}, \mathbf{z})$ and its second-order partial derivatives with respect to $\boldsymbol{w}$ are given by (3.7).

In addition, suppose that the positivity constraint (3.8) holds. Then

$$\int_{L_c^i(\boldsymbol{w}, \mathbf{z})} (f^*)''(w^i - c(\mathbf{x}, \mathbf{z}^i))\, \mathrm{d}\mathbf{x} > 0$$

for all $i \in \{1, \dots, N\}$. Using this inequality and the non-negativity of $(f^*)'$ and $(f^*)''$ gives

$$\left| \frac{\partial^2 \mathcal{G}}{\partial w^i \partial w^i}(\boldsymbol{w}, \mathbf{z}) \right| - \sum_{j \neq i} \left| \frac{\partial^2 \mathcal{G}}{\partial w^i \partial w^j}(\boldsymbol{w}, \mathbf{z}) \right| = \int_{L_c^i(\boldsymbol{w}, \mathbf{z})} (f^*)''(w^i - c(\mathbf{x}, \mathbf{z}^i))\, \mathrm{d}\mathbf{x} > 0.$$

Altogether, the Hessian $D^2_{\boldsymbol{w}\boldsymbol{w}} \mathcal{G}(\boldsymbol{w}, \mathbf{z})$ is strictly diagonally dominant, symmetric, and has negative diagonal entries, which implies that it is negative definite, as claimed. □

Now we prove the duality theorem, Theorem 3.4.

Proof of Theorem 3.4. By Lemma 2.4, $E(\cdot, \alpha^N)$ is strictly convex and has a unique minimiser over $\mathscr{P}_{\mathrm{ac}}(\mathcal{X})$.

Now we show that the maximum of $\mathcal{G}(\cdot, \mathbf{z})$ over $\mathbb{R}^N$ is attained. Since $\mathcal{G}(\cdot, \mathbf{z})$ is continuous, it suffices to show that $\lim_{\|\boldsymbol{w}\| \to \infty} \mathcal{G}(\boldsymbol{w}, \mathbf{z}) = -\infty$. Since $\mathcal{X}$ is compact and c is continuous, there exists a constant $M(\mathbf{z}) > 0$ such that

$$\max_{j \in \{1, \dots, N\}} \max_{\mathbf{x} \in \mathcal{X}} |c(\mathbf{x}, \mathbf{z}^j)| \leq M(\mathbf{z}).$$

Since f^* is non-decreasing,

$$f^*(-\boldsymbol{w}^c(\mathbf{x};\mathbf{z})) = f^*\left(\max_{j\in\{1,\dots,N\}}\left(w^j - c(\mathbf{x},\mathbf{z}^j)\right)\right) \geq f^*\left(\max_j w^j - M(\mathbf{z})\right).$$

For each $\boldsymbol{w}\in\mathbb{R}^N$, write $\boldsymbol{w} = \boldsymbol{w}_+ - \boldsymbol{w}_-$, where $\boldsymbol{w}_+, \boldsymbol{w}_-$ are the positive and negative parts of $\boldsymbol{w}$, namely, $w_+^i = \max\{w^i, 0\}$, $w_-^i = -\min\{w^i, 0\}$ for all $i\in\{1,\dots,N\}$. Since $\sum_{i=1}^N m^i = 1$,

$$\begin{aligned}
\mathcal{G}(\boldsymbol{w}) &\leq \sum_{i=1}^N m^i(w_+^i - w_-^i) - \mathcal{L}^3(\mathcal{X}) f^*\left(\max_j w^j - M(\mathbf{z})\right)\\
&\leq \|\boldsymbol{w}_+\|_\infty - \min_i m^i \,\|\boldsymbol{w}_-\|_\infty - \mathcal{L}^3(\mathcal{X}) f^*\left(\max_j w^j - M(\mathbf{z})\right)\\
&= -\min_i m^i \,\|\boldsymbol{w}_-\|_\infty + \begin{cases} 0 & \text{if } \boldsymbol{w}_+ = 0,\\ \|\boldsymbol{w}_+\|_\infty - \mathcal{L}^3(\mathcal{X}) f^*\left(\|\boldsymbol{w}_+\|_\infty - M(\mathbf{z})\right) & \text{if } \boldsymbol{w}_+ \neq 0.\end{cases}
\end{aligned}$$

The first term tends to $-\infty$ as $\|\boldsymbol{w}_-\|_\infty \to \infty$, and the second term tends to $-\infty$ as $\|\boldsymbol{w}_+\|_\infty \to \infty$ because f^* grows superlinearly at $+\infty$. Therefore $\lim_{\|\boldsymbol{w}\|\to\infty} \mathcal{G}(\boldsymbol{w},\mathbf{z}) = -\infty$, as claimed.

Next we show that $\mathcal{G}(\cdot,\mathbf{z})$ is concave. Since f^* is non-decreasing,

$$\mathcal{G}(\boldsymbol{w},\mathbf{z}) = \sum_{i=1}^N m^i w^i - \int_{\mathcal{X}} f^*(-\boldsymbol{w}^c(\mathbf{x};\mathbf{z}))\,\mathrm{d}\mathbf{x} = \sum_{i=1}^N m^i w^i - \int_{\mathcal{X}} \max_{j\in\{1,\dots,N\}} f^*(w^j - c(\mathbf{x},\mathbf{z}^j))\,\mathrm{d}\mathbf{x}.$$

Note that the map $\boldsymbol{w}\mapsto \max_{j\in\{1,\dots,N\}} f^*(w^j - c(\mathbf{x},\mathbf{z}^j))$ is convex because f^* is convex and the pointwise maximum over a family of convex functions is convex. Therefore it follows by linearity of integration that $\mathcal{G}(\cdot,\mathbf{z})$ is concave.

Now we prove that the maximiser of $\mathcal{G}(\cdot,\mathbf{z})$ is unique. Let $\boldsymbol{w}$ be a maximiser of $\mathcal{G}(\cdot,\mathbf{z})$. Then $\nabla_{\boldsymbol{w}}\mathcal{G}(\boldsymbol{w},\mathbf{z}) = 0$ and hence (3.8) holds (by (3.5) and because $m^i > 0$ for all i). Then by Proposition 3.8 the Hessian of $\mathcal{G}(\cdot,\mathbf{z})$ at $\boldsymbol{w}$ is negative definite, so $\boldsymbol{w}$ is an isolated maximiser. Since $\mathcal{G}(\cdot,\mathbf{z})$ is concave, its set of maximisers must be convex, hence connected. Therefore, $\boldsymbol{w}$ is the unique maximiser of $\mathcal{G}(\cdot,\mathbf{z})$.

Next we prove weak duality. Recall the Fenchel-Young inequality:

$$ab \leq f(a) + f^*(b) \quad \forall\, a,b\in\mathbb{R}, \tag{3.11}$$

$$ab = f(a) + f^*(b) \iff a\in\partial f^*(b) = \{(f^*)'(b)\} \iff b\in\partial f(a). \tag{3.12}$$

By the Kantorovich Duality Theorem (see (2.13)) and the Fenchel-Young inequality (3.11), for all $\sigma\in\mathscr{P}_{\mathrm{ac}}(\mathcal{X})$, $\boldsymbol{w}\in\mathbb{R}^N$,

$$\begin{aligned}
E(\sigma,\alpha^N) &= \mathcal{T}_c(\sigma,\alpha^N) + \int_{\mathcal{X}} f(\sigma(\mathbf{x}))\,\mathrm{d}\mathbf{x}\\
&= \max_{\widetilde{\boldsymbol{w}}\in\mathbb{R}^N}\left(\sum_{i=1}^N m^i\widetilde{w}^i + \int_{\mathcal{X}} \widetilde{\boldsymbol{w}}^c(\mathbf{x};\mathbf{z})\sigma(\mathbf{x})\,\mathrm{d}\mathbf{x}\right) + \int_{\mathcal{X}} f(\sigma(\mathbf{x}))\,\mathrm{d}\mathbf{x}\\
&\geq \sum_{i=1}^N m^i w^i + \int_{\mathcal{X}} \left(\boldsymbol{w}^c(\mathbf{x};\mathbf{z})\sigma(\mathbf{x}) + f(\sigma(\mathbf{x}))\right)\mathrm{d}\mathbf{x}\\
&\geq \sum_{i=1}^N m^i w^i - \int_{\mathcal{X}} f^*(-\boldsymbol{w}^c(\mathbf{x};\mathbf{z}))\,\mathrm{d}\mathbf{x} = \mathcal{G}(\boldsymbol{w},\mathbf{z}).
\end{aligned} \tag{3.13}$$

In particular, we have the weak duality

$$\min_{\sigma\in\mathscr{P}_{\mathrm{ac}}(\mathcal{X})} E(\sigma,\alpha^N) \geq \max_{\boldsymbol{w}\in\mathbb{R}^N} \mathcal{G}(\boldsymbol{w},\mathbf{z}). \tag{3.14}$$

It follows that if $E(\sigma,\alpha^N) = \mathcal{G}(\boldsymbol{w},\mathbf{z})$, then $\boldsymbol{w}$ is a maximiser of $\mathcal{G}(\cdot,\mathbf{z})$ and σ is a minimiser of $E(\cdot,\alpha^N)$.

Next we show that (3.2) and (3.3) are sufficient for optimality. Suppose that $(\sigma, \boldsymbol{w}) \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X}) \times \mathbb{R}^N$ satisfy (3.2) and (3.3). Then

$$\int_{L_c^i(\boldsymbol{w},\mathbf{z})} \sigma(\mathbf{x})\,\mathrm{d}\mathbf{x} = m^i \quad \forall\, i \in \{1,\dots,N\}.$$

By Theorem 2.9, $\boldsymbol{w}$ is an optimal Kantorovich potential for transporting σ to α^N and

$$\mathcal{T}_c(\sigma,\alpha^N) = \int_{\mathcal{X}} \boldsymbol{w}^c(\mathbf{x};\mathbf{z})\,\mathrm{d}\sigma(\mathbf{x}) + \sum_{i=1}^N m^i w^i. \tag{3.15}$$

By (3.3),

$$\sigma(\mathbf{x}) = (f^*)'(-\boldsymbol{w}^c(\mathbf{x};\mathbf{z})) \in \partial f^*(-\boldsymbol{w}^c(\mathbf{x};\mathbf{z})) \quad \forall\, \mathbf{x} \in \mathcal{X}.$$

Then by the Fenchel-Young inequality (3.12),

$$-\boldsymbol{w}^c(\mathbf{x};\mathbf{z})\,\sigma(\mathbf{x}) = f(\sigma(\mathbf{x})) + f^*(-\boldsymbol{w}^c(\mathbf{x};\mathbf{z})) \quad \forall\, \mathbf{x} \in \mathcal{X}. \tag{3.16}$$

Combining (3.15) and (3.16) gives

$$E(\sigma,\alpha^N) = \sum_{i=1}^N m^i w^i - \int_{\mathcal{X}} f^*(-\boldsymbol{w}^c(\mathbf{x};\mathbf{z}))\,\mathrm{d}\mathbf{x} = \mathcal{G}(\boldsymbol{w},\mathbf{z}).$$

Therefore, by (3.14), $\boldsymbol{w}$ is a maximiser of $\mathcal{G}(\cdot,\mathbf{z})$ and σ is a minimiser of $E(\cdot,\alpha^N)$.

We now show that (3.2) and (3.3) are necessary for optimality. Let $\sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})$ be the unique minimiser of $E(\cdot,\alpha^N)$ and let $\boldsymbol{w} \in \mathbb{R}^N$ be the unique maximiser of $\mathcal{G}(\cdot,\mathbf{z})$. Then $\nabla_{\boldsymbol{w}}\mathcal{G}(\boldsymbol{w},\mathbf{z}) = 0$ and hence, by (3.5), $\boldsymbol{w}$ satisfies (3.2). Define $\widetilde{\sigma} : \mathcal{X} \to \mathbb{R}$ by

$$\widetilde{\sigma}(\mathbf{x}) = (f^*)'(-\boldsymbol{w}^c(\mathbf{x};\mathbf{z})). \tag{3.17}$$

The pair $(\widetilde{\sigma},\boldsymbol{w}) \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X}) \times \mathbb{R}^N$ satisfies (3.2) and (3.3) by construction. It remains to show that $\widetilde{\sigma}$ is a probability measure and

$$\widetilde{\sigma} = \sigma \quad \mathcal{L}^3\text{-almost everywhere.} \tag{3.18}$$

Since $\sum_{i=1}^N m^i = 1$ and (3.2) holds, it follows that $\widetilde{\sigma}$ is a probability measure. Evaluating $E(\widetilde{\sigma},\alpha^N)$ as in the previous paragraph reveals that $E(\widetilde{\sigma},\alpha^N) = \mathcal{G}(\boldsymbol{w},\mathbf{z})$. As argued above, this holds if and only if $\widetilde{\sigma}$ and $\boldsymbol{w}$ are optimal. But since the minimiser of $E(\cdot,\alpha^N)$ is unique, (3.18) holds.

To conclude, we have shown that (3.2) and (3.3) are necessary and sufficient for optimality. The argument above for sufficiency implies that $E(\sigma,\alpha^N) = \mathcal{G}(\boldsymbol{w},\mathbf{z})$ for the unique pair of optimisers $(\sigma,\boldsymbol{w})$, so strong duality (3.1) holds. □

An immediate consequence of Theorem 3.4 and Lemma 3.7 is the following.

Corollary 3.9 (Regularity of the optimal source measure for discrete target measures)**.** *Let $\alpha^N \in \mathscr{P}^N(\mathcal{Y})$ be a discrete measure and $\sigma_*[\alpha^N] \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})$ be the optimal source measure given in Definition 2.5. Then $\sigma_*[\alpha^N]$ is Lipschitz continuous.*

4. Discrete solutions of the compressible SG equations

The goal of this section is to prove Theorem 1.1, namely, that for well-prepared initial data, discrete solutions of the compressible semi-geostrophic equations exist, are unique, and are energy conserving.

We first define the centroid map $\mathbf{C}$ that appears in the ODE (1.6) and show that it is continuously differentiable.

Definition 4.1 (Optimal weight map)**.** Given $\mathbf{m} \in \Delta^N$, define the map $\boldsymbol{w}_* : D_0^N \to \mathbb{R}^N$ by

$$\boldsymbol{w}_*(\mathbf{z}) := \operatorname*{arg\,max}_{\boldsymbol{w}\in\mathbb{R}^N} \mathcal{G}(\boldsymbol{w},\mathbf{z}). \tag{4.1}$$

Note that $\boldsymbol{w}_*$ is well defined by Theorem 3.4.

Definition 4.2 (Centroid map)**.** Given $\mathbf{m} \in \Delta^N$, define the centroid map $\mathbf{C} : D_0^N \to (\mathbb{R}^3)^N$ by $\mathbf{C}(\mathbf{z}) \coloneqq (\mathbf{C}^1(\mathbf{z}), \dots, \mathbf{C}^N(\mathbf{z}))$, where

$$\mathbf{C}^i(\mathbf{z}) \coloneqq \frac{1}{m^i} \int_{L_c^i(\boldsymbol{w}_*(\mathbf{z}),\mathbf{z})} \mathbf{x} \, \mathrm{d}\sigma_*[\alpha^N](\mathbf{x}), \tag{4.2}$$

where

$$\alpha^N = \sum_{i=1}^N m^i \delta_{\mathbf{z}^i}.$$

That is, $\mathbf{C}^i(\mathbf{z})$ is the centroid (or barycentre) of the set of points $T_{\alpha^N}^{-1}(\{\mathbf{z}^i\})$ transported to $\mathbf{z}^i$ in the optimal transport from $\sigma_*[\alpha^N]$ to α^N for the cost c.

To prove the existence of solutions of the ODE (1.6), we show that the transport velocity $W : D_0^N \to \mathbb{R}^{3N}$ is continuously differentiable, where W is defined by

$$W(\mathbf{z}) = J^N(\mathbf{z} - \mathbf{C}(\mathbf{z})),$$

where J^N is the block diagonal matrix $J^N := \mathrm{diag}(J, \dots, J) \in \mathbb{R}^{3N \times 3N}$.

Lemma 4.3 (Regularity of $\boldsymbol{w}_*$)**.** *The optimal weight map $\boldsymbol{w}_*$ is continuously differentiable.*

Proof. By definition of $\boldsymbol{w}_*$, $\nabla_{\boldsymbol{w}} \mathcal{G}(\boldsymbol{w}_*(\mathbf{z}), \mathbf{z}) = 0$ for all $\mathbf{z} \in D^N$. In order to show that $\boldsymbol{w}_*$ is continuously differentiable, we will apply the Implicit Function Theorem to the function $\nabla_{\boldsymbol{w}} \mathcal{G}$. Fix any point $\mathbf{z}_0 \in D_0^N$ and define $\boldsymbol{w}_0 = \boldsymbol{w}_*(\mathbf{z}_0)$. By Proposition 3.8 and equation (3.2), $\mathcal{G}$ is twice continuously differentiable in a neighbourhood of $(\boldsymbol{w}_0, \mathbf{z}_0)$ and $D^2_{\boldsymbol{w}\boldsymbol{w}} \mathcal{G}(\boldsymbol{w}_0, \mathbf{z}_0)$ is negative definite, hence invertible. By the Implicit Function Theorem, there exists an open set $U \subset D_0^N$ containing $\mathbf{z}_0$ and a unique, continuously differentiable function $\widetilde{\boldsymbol{w}} : U \to \mathbb{R}^N$ such that $\widetilde{\boldsymbol{w}}(\mathbf{z}_0) = \boldsymbol{w}_0$ and $\nabla_{\boldsymbol{w}} \mathcal{G}(\widetilde{\boldsymbol{w}}(\mathbf{z}), \mathbf{z}) = 0$ for all $\mathbf{z} \in U$. By uniqueness, $\widetilde{\boldsymbol{w}}$ coincides with $\boldsymbol{w}_*$ on U. Therefore $\boldsymbol{w}_*$ is continuously differentiable in a neighbourhood of $\mathbf{z}_0$, as required. □

Lemma 4.4 (Regularity of $\mathbf{C}$)**.** *The centroid map $\mathbf{C}$ is continuously differentiable.*

Proof. Recalling the expression for $\sigma_*[\alpha^N]$ derived in Theorem 3.4, we have

$$\mathbf{C}^i(\mathbf{z}) = \frac{1}{m^i} \int_{L_c^i(\boldsymbol{w}_*(\mathbf{z}),\mathbf{z})} \mathbf{x} \, (f^*)'(-(\boldsymbol{w}_*(\mathbf{z}))^c(\mathbf{x};\mathbf{z})) \, \mathrm{d}\mathbf{x}.$$

Define $\boldsymbol{\Psi} = (\Psi^1, \dots, \Psi^N) : \mathbb{R}^N \times D_0^N \to (\mathbb{R}^3)^N$ by

$$\Psi^i(\boldsymbol{w}, \mathbf{z}) = \frac{1}{m^i} \int_{L_c^i(\boldsymbol{w},\mathbf{z})} \mathbf{x} \, (f^*)'(-\boldsymbol{w}^c(\mathbf{x};\mathbf{z})) \, \mathrm{d}\mathbf{x}.$$

Then $\mathbf{C} = \Psi \circ (\boldsymbol{w}_*, \mathrm{id})$. For $j \in \{1, 2, 3\}$, define the functions $\zeta_j : \mathcal{X} \times \mathbb{R}^N \times D^N \to \mathbb{R}$ by $\zeta_j(\mathbf{x}, \boldsymbol{w}, \mathbf{z}) = x_j \, (f^*)'(-\boldsymbol{w}^c(\mathbf{x};\mathbf{z}))$. It follows from Lemma 3.7 that ζ_j satisfy the assumptions of Lemma 3.6, so Ψ is continuously differentiable by Lemma 3.6. Therefore $\mathbf{C}$ is continuously differentiable by Lemma 4.3 and the Chain Rule. □

We now show that solutions of the ODE (1.6) correspond to weak solutions of the PDE (1.1) with discrete initial data. The following is essentially a consequence of [8, Lemma 4.2].

Proposition 4.5. *Let $\overline{z} \in D^N$ and let $\overline{\alpha}^N \in \mathscr{P}^N(\mathcal{Y})$ be the corresponding discrete initial measure given by*

$$\overline{\alpha}^N = \sum_{i=1}^N m^i \delta_{\overline{\mathbf{z}}^i}. \tag{4.3}$$

Fix a final time $\tau > 0$ and let $\mathbf{z} \in \mathcal{C}^1([0,\tau]; \mathcal{Y}^N)$. Define $\alpha^N : [0,\tau] \to \mathscr{P}^N(\mathcal{Y})$ by

$$\alpha_t^N = \sum_{i=1}^N m^i \delta_{\mathbf{z}^i(t)} \quad \forall \, t \in [0,\tau]. \tag{4.4}$$

Then α^N is a weak solution of the compressible semi-geostrophic equations (in the sense of Definition 2.6) with initial data $\alpha_0^N = \overline{\alpha}^N$ if and only if $\mathbf{z}$ is a solution of the ODE (1.6) with initial data $\mathbf{z}(0) = \overline{\mathbf{z}}$.

Proof. Suppose that $\mathbf{z} \in C^1([0,\tau];\mathcal{Y}^N)$ is a solution of the ODE (1.6) and let α^N be given by (4.4). We show that α^N satisfies (2.6). Let $\varphi \in \mathscr{D}(\mathcal{Y} \times \mathbb{R})$. Recall that we use the notation $\varphi_t(y) = \varphi(y,t)$, i.e., the subscript t does not denote differentiation. Then

$$
\begin{aligned}
&\int_0^\tau \int_{\mathcal{Y}} \left[\partial_t \varphi_t(\mathbf{y}) + (J\mathbf{y}) \cdot \nabla \varphi_t(\mathbf{y})\right] \mathrm{d}\alpha_t^N(\mathbf{y})\,\mathrm{d}t - \int_0^\tau \int_{\mathcal{X}} (J\mathbf{x}) \cdot \nabla \varphi_t(T_{\alpha_t^N}(\mathbf{x}))\,\mathrm{d}\sigma_*[\alpha_t^N](\mathbf{x})\,\mathrm{d}t \\
&= \sum_{i=1}^N m^i \int_0^\tau \left[\partial_t \varphi_t(\mathbf{z}^i(t)) + (J\mathbf{z}^i(t)) \cdot \nabla \varphi_t(\mathbf{z}^i(t))\right] \mathrm{d}t \\
&\quad - \int_0^\tau \sum_{i=1}^N \int_{L_c^i(\boldsymbol{w}_*(\mathbf{z}(t)),\mathbf{z}(t))} (J\mathbf{x}) \cdot \nabla \varphi_t(\mathbf{z}^i(t))\,\mathrm{d}\sigma_*[\alpha_t^N](\mathbf{x})\,\mathrm{d}t \\
&= \sum_{i=1}^N m^i \int_0^\tau \left[\partial_t \varphi_t(\mathbf{z}^i(t)) + J\left(\mathbf{z}^i(t) - \mathbf{C}^i(\mathbf{z}(t))\right) \cdot \nabla \varphi_t(\mathbf{z}^i(t))\right] \mathrm{d}t \\
&= \sum_{i=1}^N m^i \int_0^\tau \left[\partial_t \varphi_t(\mathbf{z}^i(t)) + \dot{\mathbf{z}}^i(t) \cdot \nabla \varphi_t(\mathbf{z}^i(t))\right] \mathrm{d}t \\
&= \sum_{i=1}^N m^i \int_0^\tau \frac{d}{dt} \varphi_t(\mathbf{z}^i(t))\,\mathrm{d}t \\
&= \sum_{i=1}^N m^i \left(\varphi_\tau(\mathbf{z}^i(\tau)) - \varphi_0(\mathbf{z}^i(0))\right) \\
&= \int_{\mathcal{Y}} \varphi_\tau(\mathbf{y})\,\mathrm{d}\alpha_\tau^N(\mathbf{y}) - \int_{\mathcal{Y}} \varphi_0(\mathbf{y})\,\mathrm{d}\overline{\alpha}^N(\mathbf{y}),
\end{aligned}
$$

as required. From here, the proof follows exactly the proof of Lemma 4.2 given in [8]. □

Next we prove global existence and uniqueness of solutions of the ODE (1.6) using the following *a priori* estimates.

Lemma 4.6 (*A priori* estimates)**.** *Let $\mathbf{z} : [0,\tau] \to D^N$ be continuously differentiable and satisfy the ODE (1.6) on the interval $[0,\tau]$. Let $R = R(\mathcal{X}) > 0$ satisfy $\mathcal{X} \subset B_R(0)$. Then, for each $i \in \{1,\dots,N\}$ and $t \in [0,\tau]$,*

$$\|\mathbf{z}^i(t)\| \le \|\mathbf{z}^i(0)\| + f_{\mathrm{cor}} R\tau, \tag{4.5}$$

$$\|\dot{\mathbf{z}}^i(t)\| \le f_{\mathrm{cor}} \left(\|\mathbf{z}^i(0)\| + R(f_{\mathrm{cor}}\tau + 1)\right). \tag{4.6}$$

Proof. By (1.6), and since J is skew-symmetric with $\|J\|_2 = f_{\mathrm{cor}}$,

$$\frac{d}{dt}\|\mathbf{z}^i(t)\| = \frac{\mathbf{z}^i}{\|\mathbf{z}^i(t)\|} \cdot \dot{\mathbf{z}}^i(t) = -\frac{\mathbf{z}^i(t) \cdot J\,\mathbf{C}^i(\mathbf{z}(t))}{\|\mathbf{z}^i(t)\|} \le f_{\mathrm{cor}} \|\mathbf{C}^i(\mathbf{z}(t))\| \le f_{\mathrm{cor}} R.$$

Integrating gives (4.5). By (1.6) and (4.5),

$$\|\dot{\mathbf{z}}^i(t)\| = \|J(\mathbf{z}^i(t) - \mathbf{C}^i(\mathbf{z}(t)))\| \le f_{\mathrm{cor}} \left(\|\mathbf{z}^i(t)\| + \|\mathbf{C}^i(\mathbf{z}(t))\|\right) \le f_{\mathrm{cor}} \left(\|\mathbf{z}^i(0)\| + f_{\mathrm{cor}} R\tau + R\right),$$

as required. □

Proposition 4.7 (Existence and uniqueness of solutions of the ODE (1.6))**.** *Let the final time $\tau \in (0,\infty)$ be arbitrary and let the initial data $\overline{\mathbf{z}} \in D_0^N$. Then there exists a unique function $\mathbf{z} \in \mathcal{C}^2([0,\tau]; D_0^N)$ such that $\mathbf{z}$ satisfies the ODE (1.6) on the interval $[0,\tau]$ and $\mathbf{z}(0) = \overline{\mathbf{z}}$.*

Proof. The proof of Proposition 4.7 follows exactly the proof of Proposition 4.4 of [8]. Indeed, local existence and uniqueness follows from the Picard–Lindelöf Theorem, Lemma 4.4, and the assumption on the initial data $\overline{\mathbf{z}}$. Global existence then follows from the *a priori* estimates (Lemma 4.6) and the fact that the third row of the matrix J is zero, which implies that any solution $\mathbf{z}$ of the ODE satisfies $z_3^i(t) = z_3^i(0)$ for all t and i. In particular, for all t, $\mathbf{z}(t)$ belongs to the subset D_0^N of D^N where the centroid map $\mathbf{C}$ is continuously differentiable. □

Furthermore, we prove that the discrete solutions conserve energy.

Proposition 4.8 (Solutions are energy-conserving)**.** *Any solution* $\mathbf{z} \in \mathcal{C}^1([0,\tau]; D_0^N)$ *of the ODE* (1.6) *is energy-conserving in the sense that*

$$\frac{d}{dt} E(\sigma_*[\alpha_t^N], \alpha_t^N) = 0 \quad \forall\, t \in [0,\tau],$$

where

$$\alpha_t^N = \sum_{i=1}^N m^i \delta_{\mathbf{z}^i(t)}.$$

Proof. By the duality theorem (Theorem 3.4),

$$\begin{aligned}
\frac{d}{dt} E(\sigma_*[\alpha_t^N], \alpha_t^N) &= \frac{d}{dt} \mathcal{G}(\boldsymbol{w}_*(\mathbf{z}(t)), \mathbf{z}(t)) \\
&= \sum_{i=1}^N \left(\frac{\partial \mathcal{G}}{\partial w^i}(\boldsymbol{w}_*(\mathbf{z}(t)), \mathbf{z}(t)) \frac{d}{dt} w_*^i(\mathbf{z}(t)) + \frac{\partial \mathcal{G}}{\partial \mathbf{z}^i}(\boldsymbol{w}_*(\mathbf{z}(t)), \mathbf{z}(t)) \cdot \dot{\mathbf{z}}^i(t) \right) \\
&= \sum_{i=1}^N \frac{\partial \mathcal{G}}{\partial \mathbf{z}^i}(\boldsymbol{w}_*(\mathbf{z}(t)), \mathbf{z}(t)) \cdot \dot{\mathbf{z}}^i(t) \qquad (4.7)
\end{aligned}$$

because $\boldsymbol{w}_*(\mathbf{z}(t))$ is the maximiser (and hence a critical point) of $\mathcal{G}(\cdot, \mathbf{z}(t))$. For brevity, define

$$L_c^i(t) := L_c^i(\boldsymbol{w}_*(\mathbf{z}(t)), \mathbf{z}(t)),$$

and let $\mathbf{C}^i(t) = \mathbf{C}^i(\mathbf{z}(t))$. By (3.3) and (3.6),

$$\begin{aligned}
\frac{\partial \mathcal{G}}{\partial \mathbf{z}^i}(\boldsymbol{w}_*(\mathbf{z}(t)), \mathbf{z}(t)) \cdot \dot{\mathbf{z}}^i(t) &= \left(\int_{L_c^i(t)} \nabla_{\mathbf{y}} c(\mathbf{x}, \mathbf{z}^i(t))\, \sigma_*[\alpha_t^N](\mathbf{x})\, \mathrm{d}\mathbf{x} \right) \cdot \dot{\mathbf{z}}^i(t) \\
&= \frac{1}{z_3^i(t)} \left(\int_{L_c^i(t)} \begin{pmatrix} f_{\mathrm{cor}}^2(z_1^i(t) - x_1) \\ f_{\mathrm{cor}}^2(z_2^i(t) - x_2) \\ -c(\mathbf{x}, \mathbf{z}^i(t)) \end{pmatrix} \sigma_*[\alpha_t^N](\mathbf{x})\, \mathrm{d}\mathbf{x} \right) \cdot \dot{\mathbf{z}}^i(t) \\
&= \frac{m^i f_{\mathrm{cor}}^2}{z_3^i(t)} (\mathbf{z}^i - \mathbf{C}^i(t)) \cdot J(\mathbf{z}^i - \mathbf{C}^i(t)) \\
&= 0 \qquad (4.8)
\end{aligned}$$

where the final two equalities hold by (1.6), skew-symmetry of J, and the fact that the third row of J is zero. Combining (4.7) and (4.8) completes the proof. □

Finally, we prove Theorem 1.1.

Proof of Theorem 1.1. Let $\overline{\alpha}^N = \sum_{i=1}^N m^i \delta_{\overline{\mathbf{z}}^i} \in \mathscr{P}^N(\mathcal{Y})$ be well-prepared in the sense of (2.9). By Proposition 4.7 the ODE (1.6) has a unique solution $\mathbf{z} \in \mathcal{C}^2([0,\tau]; D_0^N)$ satisfying $\mathbf{z}(0) = \overline{\mathbf{z}}$. Define $\alpha^N : [0,\tau] \to \mathscr{P}^N(\mathcal{Y})$ by $\alpha_t^N = \sum_{i=1}^N m^i \delta_{\mathbf{z}^i(t)}$. By Proposition 4.5 α^N is weak solution of the compressible SG equations (in the sense of Definition 2.6) with $\alpha_0^N = \overline{\alpha}^N$. Furthermore, α^N is energy-conserving by Proposition 4.8. □

5. Existence of weak solutions

In this section we prove Theorem 1.2. The strategy is to approximate the initial measure $\overline{\alpha} \in \mathscr{P}_{\mathrm{c}}(\mathcal{Y})$ by a sequence of discrete measures $\overline{\alpha}^N \in \mathscr{P}^N(\mathcal{Y})$, apply Theorem 1.1 to obtain a weak solution of (1.1) for each initial data $\overline{\alpha}^N$, and then pass to the limit $N \to \infty$ to obtain a weak solution of (1.1) with initial data $\overline{\alpha}$. This is the same proof strategy as for the *incompressible* system given in [8]. Indeed, some of the preliminary lemmas, namely Lemmas 5.1 and 5.2, are identical and are stated here without proof. The first lemma asserts that any probability measure $\mu \in \mathscr{P}_{\mathrm{c}}(\mathcal{Y})$ can be approximated in the W_1 metric by a discrete measure $\mu^N \in \mathscr{P}^N(\mathcal{Y})$ with seeds in distinct horizontal planes.

Lemma 5.1 (Quantization, [8, Lemma 5.1])**.** *Let $\mu \in \mathscr{P}_{\mathrm{c}}(\mathcal{Y})$. There exists a compact set $K \subset \mathcal{Y}$ and a sequence of well-prepared discrete probability measures $\mu^N \in \mathscr{P}^N(\mathcal{Y})$ (in the sense of (2.9)) such that*

$$\operatorname{spt}(\mu^N) \subset K \quad \forall\, N \in \mathbb{N}, \qquad \lim_{N\to\infty} W_1(\mu^N, \mu) = 0.$$

Lemma 5.2 (Compactness, [8, Lemma 5.2])**.** *Let $\overline{\alpha} \in \mathscr{P}_{\mathrm{c}}(\mathcal{Y})$ and let $\overline{\alpha}^N \in \mathscr{P}^N(\mathcal{Y})$ given by Lemma 5.1 be a sequence of well-prepared discrete discrete probability measures converging to $\overline{\alpha}$ in W_1. Fix $\tau \in (0,\infty)$. Let $\alpha^N \in \mathcal{C}^{0,1}([0,\tau];\mathscr{P}^N(\mathcal{Y}))$ given by Theorem 1.1 be a discrete weak solution of (1.1) with initial data $\overline{\alpha}^N$. Then the sequence $(\alpha^N)_{N\in\mathbb{N}}$ has a uniformly convergent subsequence in $\mathcal{C}([0,\tau];\mathscr{P}_{\mathrm{c}}(\mathcal{Y}))$. To be precise, there exists a subsequence (which we do not relabel) and a Lipschitz map $\alpha \in \mathcal{C}^{0,1}([0,\tau];\mathscr{P}_{\mathrm{c}}(\mathcal{Y}))$ such that*

$$\lim_{N\to\infty} \sup_{t\in[0,\tau]} W_1\big(\alpha_t^N, \alpha_t\big) = 0. \tag{5.1}$$

Moreover, there exists a compact set $K \subset \mathcal{Y}$ such that $\operatorname{spt}(\alpha_t^N) \subset K$ and $\operatorname{spt}(\alpha_t) \subset K$ for all $t \in [0,\tau]$ and $N \in \mathbb{N}$.

The final task is to show that α in Lemma 5.2 satisfies the PDE (1.1). One of the ingredients of the proof is the following.

Lemma 5.3 (Uniform convergence of source measures)**.** *Given a compact set $K \subset \mathcal{Y}$ and a sequence $\beta^N \in \mathscr{P}^N(\mathcal{Y}) \cap \mathscr{P}(K)$ converging weakly to $\beta \in \mathscr{P}(K)$, the sequence of optimal source measures $\sigma_*[\beta^N]$ converges uniformly to $\sigma_*[\beta]$.*

Proof. Recall from Corollary 3.9 that $\sigma_*[\beta^N]$ is Lipschitz continuous for all $N \in \mathbb{N}$. Moreover, it follows from the proof of Lemma 3.7 that its $W^{1,\infty}$-norm can be bounded independently of N (by noting that $\max_{\mathbf{z}\in D^N\cap K^N} \|\zeta^i(\cdot, \boldsymbol{w}_*(\mathbf{z}), \mathbf{z})\|_{\mathcal{C}^1(\mathcal{X})}$ is independent of N, where ζ^i was defined in equation (3.4)). Therefore the sequence $\sigma_*[\beta^N]$ is uniformly bounded and equicontinuous. By the Ascoli-Arzelá Theorem, there exists $\widetilde{\sigma} \in \mathcal{C}(\mathcal{X}) \cap \mathscr{P}_{\mathrm{ac}}(\mathcal{X})$ such that $\sigma_*[\beta^N]$ converges uniformly in $\mathcal{X}$ to $\widetilde{\sigma}$ (up to a subsequence that we do not relabel).

Next we show that $\widetilde{\sigma} = \sigma_*[\beta]$. The continuity of $\mathcal{T}_c$ (see [41, Theorem 1.51]) and the uniform convergence of $\sigma_*[\beta^N]$ imply that

$$E\big(\sigma_*[\beta^N], \beta^N\big) \to E(\widetilde{\sigma}, \beta) \quad \text{and} \quad E(\sigma_*[\beta], \beta^N) \to E(\sigma_*[\beta], \beta) \quad \text{as} \quad N \to \infty.$$

Assume for a contradiction that $E(\widetilde{\sigma}, \beta) > E(\sigma_*[\beta], \beta)$. Then $E(\sigma_*[\beta^N], \beta^N) > E(\sigma_*[\beta], \beta^N)$ for N sufficiently large, which contradicts the fact that $\sigma_*[\beta^N]$ is the global minimiser of $E(\cdot, \beta^N)$. Hence $E(\widetilde{\sigma}, \beta) = E(\sigma_*[\beta], \beta)$ and so $\widetilde{\sigma} = \sigma_*[\beta]$ by the uniqueness of the minimiser of $E(\cdot, \beta)$. Finally, note that all convergent subsequences of $\sigma_*[\beta^N]$ converge to $\sigma_*[\beta]$, hence the whole sequence must converge. □

For the proof of Theorem 1.2, weak convergence of $\sigma_*[\alpha_t^N]$ is in fact sufficient. However, uniform convergence of $\sigma_*[\alpha_t^N]$, along with the fact that $\sigma_*[\alpha_t^N]$ is Lipschitz continuous uniformly in α_t^N (see Lemma 3.7), implies the following.

Corollary 5.4 (Regularity of the optimal source measure for general target measures)**.** *Let $\beta \in \mathscr{P}_{\mathrm{c}}(\mathcal{Y})$ and $\sigma_*[\beta] \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})$ be the optimal source measure. Then $\sigma_*[\beta]$ is Lipschitz continuous.*

Proof. This is an immediate consequence of Corollary 3.9 and Lemmas 3.7, 5.1 and 5.3. □

Finally we prove our main theorem.

Proof of Theorem 1.2. Let $\alpha \in \mathcal{C}^{0,1}([0,\tau];\mathscr{P}_{\mathrm{c}}(\mathcal{Y}))$ be the limit of the sequence $(\alpha^N)_{N\in\mathbb{N}}$ obtained in Lemma 5.2, and let $K \subset \mathcal{Y}$ be a compact set such that $\mathrm{spt}(\alpha_t^N) \subset K$ and $\mathrm{spt}(\alpha_t) \subset K$ for all $t \in [0,\tau]$ and $N \in \mathbb{N}$. By (2.6) and (2.7), for all $\varphi \in \mathscr{D}(\mathcal{Y}\times\mathbb{R})$,

$$\int_0^\tau \int_{\mathcal{Y}} [\partial_t\varphi_t(\mathbf{y}) + (J\mathbf{y})\cdot\nabla\varphi_t(\mathbf{y})]\,\mathrm{d}\alpha_t^N(\mathbf{y})\,\mathrm{d}t - \int_0^\tau \int_{\mathcal{X}\times\mathcal{Y}} (J\mathbf{x})\cdot\nabla\varphi_t(\mathbf{y})\,\mathrm{d}\gamma[\alpha_t^N](\mathbf{x},\mathbf{y})\,\mathrm{d}t$$
$$= \int_{\mathcal{Y}} \varphi_\tau(\mathbf{y})\,\mathrm{d}\alpha_\tau^N(\mathbf{y}) - \int_{\mathcal{Y}} \varphi_0(\mathbf{y})\,\mathrm{d}\overline{\alpha}^N(\mathbf{y}). \quad (5.2)$$

Recall the following standard estimate in optimal transport theory (see [41, Exercise 38] or [44, Theorem 1.14]):

$$\left|\int_K \psi\ \mathrm{d}\big(\alpha_t^N - \alpha_t\big)\right| \le \|\nabla\psi\|_{L^\infty(K)} W_1(\alpha_t^N,\alpha_t) \quad \forall\ \psi \in W^{1,\infty}(K).$$

Then

$$\begin{aligned}
&\left|\int_0^\tau \int_{\mathcal{Y}} [\partial_t\varphi_t(\mathbf{y}) + (J\mathbf{y})\cdot\nabla\varphi_t(\mathbf{y})]\ \mathrm{d}\big(\alpha_t^N - \alpha_t\big)(\mathbf{y})\,\mathrm{d}t\right| \\
&\le \int_0^\tau \max_{\mathbf{y}\in K} \big\{\|\nabla_{\mathbf{y}}[\partial_t\varphi_t(\mathbf{y}) + (J\mathbf{y})\cdot\nabla\varphi_t(\mathbf{y})]\|\big\}\, W_1(\alpha_t^N,\alpha_t)\,\mathrm{d}t \\
&\le \tau \max_{\substack{t\in[0,\tau]\\ \mathbf{y}\in K}} \big\{\|\nabla_{\mathbf{y}}[\partial_t\varphi_t(\mathbf{y}) + (J\mathbf{y})\cdot\nabla\varphi_t(\mathbf{y})]\|\, W_1(\alpha_t^N,\alpha_t)\big\} \to 0 \quad \text{as} \quad N\to\infty \qquad (5.3)
\end{aligned}$$

by uniform convergence of α^N (seee (5.1)). Pointwise-in-time convergence of α^N implies that

$$\lim_{N\to\infty} \int_{\mathcal{Y}} \varphi_\tau(\mathbf{y})\ \mathrm{d}\big(\alpha_\tau^N - \alpha_\tau\big)(\mathbf{y}) = 0 \quad \text{and} \quad \lim_{N\to\infty} \int_{\mathcal{Y}} \varphi_0(\mathbf{y})\ \mathrm{d}\big(\overline{\alpha}^N - \overline{\alpha}\big)(\mathbf{y}) = 0.$$

By Lemma 5.3, for all $t \in [0,\tau]$, $\sigma_*[\alpha_t^N] \rightharpoonup \sigma_*[\alpha_t]$ in $\mathscr{P}(\mathcal{X})$. Also, $\alpha_t^N \rightharpoonup \alpha_t$ in $\mathscr{P}(K)$. Therefore $\gamma[\alpha_t^N] \rightharpoonup \gamma[\alpha_t]$ in $\mathscr{P}(\mathcal{X}\times K)$, where $\gamma[\alpha_t]$ is the unique optimal transport plan for transporting $\sigma_*[\alpha_t]$ to α_t; see [45, Theorem 5.20] . Define $F, F_N \in L^1([0,\tau])$ by

$$F_N(t) = \int_{\mathcal{X}\times\mathcal{Y}} (J\mathbf{x})\cdot\nabla\varphi_t(\mathbf{y})\,\mathrm{d}\gamma[\alpha_t^N](\mathbf{x},\mathbf{y}), \quad F(t) = \int_{\mathcal{X}\times\mathcal{Y}} (J\mathbf{x})\cdot\nabla\varphi_t(\mathbf{y})\,\mathrm{d}\gamma[\alpha_t](\mathbf{x},\mathbf{y}).$$

Then $\lim_{N\to\infty} F_N(t) = F(t)$ for all $t \in [0,\tau]$ and

$$\|F_N\|_{L^\infty([0,\tau])} \le f_{\mathrm{cor}} \max_{\mathbf{x}\in\mathcal{X}} \|\mathbf{x}\| \max_{t\in[0,\tau]} \|\nabla\varphi_t\|_{L^\infty(K)}.$$

By the Lebesgue Dominated Convergence Theorem,

$$\lim_{N\to\infty} \int_0^\tau \int_{\mathcal{X}\times\mathcal{Y}} (J\mathbf{x})\cdot\nabla\varphi_t(\mathbf{y})\,\mathrm{d}\gamma[\alpha_t^N](\mathbf{x},\mathbf{y})\,\mathrm{d}t = \int_0^\tau \int_{\mathcal{X}\times\mathcal{Y}} (J\mathbf{x})\cdot\nabla\varphi_t(\mathbf{y})\,\mathrm{d}\gamma[\alpha_t](\mathbf{x},\mathbf{y})\,\mathrm{d}t. \quad (5.4)$$

Combining equations (5.2)–(5.4) shows that α is a weak solution of (1.1), as required. □

Remark 5.5 (Convergence of the transport maps)**.** It can be shown that $T_{\alpha_t^N} \to T_{\alpha_t}$ as $N\to\infty$ in L^p for $p \in (1,\infty)$. This can be combined with the uniform convergence of the source measures to pass to the limit in the nonlinear term

$$\int_0^\tau \int_{\mathcal{X}} (J\mathbf{x})\cdot\nabla\varphi_t(T_{\alpha_t^N}(\mathbf{x}))\,\mathrm{d}\sigma_*[\alpha_t^N](\mathbf{x})\,\mathrm{d}t$$

to give an alternative, though more involved, proof of Theorem 1.2.

Remark 5.6 (Conservation of energy)**.** Proposition 4.8 states that discrete weak solutions of (1.1) conserve the geostrophic energy. In fact, the same is true for the weak solutions with arbitrary initial data $\overline{\alpha} \in \mathscr{P}_{\mathrm{c}}(\mathcal{Y})$ that we construct in the proof of Theorem 1.2. We give a quick sketch of the proof here. Let $\alpha \in \mathcal{C}^{0,1}([0,\tau];\mathscr{P}_{\mathrm{c}}(\mathcal{Y}))$ be a weak solution of (1.1) with initial measure $\overline{\alpha}$, and let $\alpha^N \in \mathcal{C}^{0,1}([0,\tau];\mathscr{P}^N(\mathcal{Y}))$, $N \in \mathbb{N}$, be discrete weak solutions that uniformly approximate α in the sense of (1.7). We proved that α^N conserve energy in Proposition 4.8. We will use this to prove that α also conserves energy. For each $t \in [0,\tau]$ define $H_t, H_t^N : \mathscr{P}(\mathcal{X}) \to \mathbb{R} \cup \{+\infty\}$ by

$$H_t(\sigma) = E(\sigma, \alpha_t), \qquad H_t^N(\sigma) = E(\sigma, \alpha_t^N).$$

Then H_t^N Γ-converges to H_t with respect to weak convergence of measures. To see this, note that E is lower semi-continuous because the transport cost $\mathcal{T}_c$ is continuous [45, Theorem 5.20] and F is lower semi-continuous [41, Proposition 7.7]. Take any sequence σ^N in $\mathscr{P}(\mathcal{X})$ such that $\sigma^N \rightharpoonup \sigma$. Recall that $\alpha_t^N \rightharpoonup \alpha_t$. Then

$$H_t(\sigma) = E(\sigma, \alpha_t) \leq \liminf_{N\to\infty} E(\sigma^N, \alpha_t^N) = \liminf_{N\to\infty} H_t^N(\sigma^N),$$

which is the liminf inequality for Γ-convergence. Moreover, the constant recovery sequence $\sigma^N = \sigma$ satisfies

$$\lim_{N\to\infty} H_t^N(\sigma^N) = \lim_{N\to\infty} E(\sigma, \alpha_t^N) = E(\sigma, \alpha_t) = H_t(\sigma),$$

which is the limsup inequality for Γ-convergence. Therefore $H_t^N \xrightarrow{\Gamma} H_t$, as claimed. By a standard property of Γ-convergence (convergence of the minimum values of the energy) and Proposition 4.8,

$$\begin{aligned} E(\sigma_*[\alpha_t], \alpha_t) = \min H_t &= \lim_{N\to\infty} \min H_t^N = \lim_{N\to\infty} E(\sigma_*[\alpha_t^N], \alpha_t^N) = \lim_{N\to\infty} E(\sigma_*[\alpha_0^N], \alpha_0^N) \\ &= \lim_{N\to\infty} \min H_0^N = \min H_0 = E(\sigma_*[\alpha_0], \alpha_0), \end{aligned}$$

which proves that α is energy-conserving, as claimed. Note, however, that we have not proved that the energy is conserved by *all* weak solutions of (1.1), only by those that can be approximated by discrete weak solutions as in Theorem 1.2.

6. Explicit examples

In this section we present two explicit solutions of (1.1): a steady state solution and a discrete solution with a single seed.

6.1. **Steady state example.**

In this section we make the additional assumption that the fluid domain satisfies $\mathcal{X} \subset \mathcal{Y}$.

Proposition 6.1 (Time-independent solution)**.** *Given $\ell \in \mathbb{R}$, define $\sigma_\ell : \mathcal{X} \to \mathbb{R}$ by*

$$\sigma_\ell(\mathbf{x}) := (f^*)'(\ell - g \ln x_3).$$

Then there exists $\ell_* \in \mathbb{R}$ such that

$$\int_{\mathcal{X}} \sigma_{\ell_*}(\mathbf{x}) \, \mathrm{d}\mathbf{x} = 1.$$

Define $\alpha : [0,\tau] \to \mathscr{P}_{\mathrm{c}}(\mathcal{Y})$ by

$$\alpha_t = \sigma_{\ell_*} \mathcal{L}^3 \llcorner \mathcal{X} \quad \forall\, t \in [0,\tau].$$

Then α is a weak steady state solution of the compressible SG equation (1.1)*.*

Remark 6.2 (Formal derivation)**.** We give a formal derivation of the steady state α before giving a rigorous proof of Proposition 6.1. By Definition 2.6, $\alpha \in \mathcal{C}([0,\tau];\mathscr{P}_{\mathrm{c}}(\mathcal{Y}))$ is a steady solution of (1.1) if α_t is independent of t, $\alpha_t = \sigma_*[\alpha_t]$ (in particular, $\mathrm{spt}(\alpha_t) \subseteq \mathcal{X}$), and $T_{\alpha_t} = \mathrm{id}$. Then

$$\alpha_t = \operatorname*{arg\,min}_{\sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})} E(\sigma, \alpha_t)$$

for all $t \in [0, \tau]$. It follows from [41, Proposition 7.20] that α_t satisfies the Euler-Lagrange equations

$$\begin{cases} \dfrac{\delta E}{\delta \sigma}(\alpha_t, \alpha_t) = \ell & \text{on the support of } \alpha_t, \\ \dfrac{\delta E}{\delta \sigma}(\alpha_t, \alpha_t) \geq \ell & \text{otherwise,} \end{cases}$$

for some $\ell \in \mathbb{R}$, where $\delta E/\delta\sigma$ denotes the first variation of E with respect to σ. Formally we expect

$$\frac{\delta E}{\delta \sigma}(\alpha_t, \alpha_t) = \varphi + f'(\alpha_t),$$

where $\varphi : \mathcal{X} \to \mathbb{R}$ is an optimal c-concave Kantorovich potential for transporting α_t to itself with cost c; see [41, Section 7.2]. Then

$$\begin{cases} \varphi + f'(\alpha_t) = \ell & \text{on the support of } \alpha_t, \\ \varphi + f'(\alpha_t) \geq \ell & \text{otherwise.} \end{cases}$$

If $T_{\alpha_t} = \mathrm{id}$ is the optimal transport map from α_t to itself , then $\nabla\varphi(\mathbf{x}) = \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{x})$ for almost all $\mathbf{x}$ in the support of α_t [41, Proposition 1.15]. Integrating this expression gives $\varphi(\mathbf{x}) = g \ln x_3$, up to a constant. Then solving the Euler-Lagrange equations yields

$$\alpha_t(\mathbf{x}) = (f')^{-1}(\ell - \varphi(\mathbf{x})) = (f^*)'(\ell - g \ln x_3)$$

for all $\mathbf{x}$ in the support of α_t. This is the steady state given in Proposition 6.1. The constant ℓ is determined by the constraint that α_t is a probability measure.

Proof of Proposition 6.1. First we show that there exists $\ell_* \in \mathbb{R}$ such that $\int_{\mathcal{X}} \sigma_{\ell_*}(\mathbf{x})\,\mathrm{d}\mathbf{x} = 1$. Since $X \subset \mathcal{Y} = \mathbb{R}^2 \times (\delta, 1/\delta)$ and $(f^*)'$ is non-decreasing, then for $\ell \in \mathbb{R}$

$$(f^*)'(\ell - g \ln x_3) \geq (f^*)'(\ell - g \ln(1/\delta))$$

for all $\mathbf{x} \in \mathcal{X}$. Therefore

$$\int_{\mathcal{X}} \sigma_\ell(\mathbf{x})\,\mathrm{d}\mathbf{x} \geq (f^*)'(\ell - g \ln(1/\delta))|\mathcal{X}| \to \infty \quad \text{as} \quad \ell \to \infty.$$

Let $\ell_\delta = g \ln \delta$. Since $x_3 \geq \delta$ for all $\mathbf{x} \in \mathcal{X}$,

$$\int_{\mathcal{X}} \sigma_{\ell_\delta}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathcal{X}} (f^*)'(\ell_\delta - g \ln x_3)\,\mathrm{d}\mathbf{x} \leq \int_{\mathcal{X}} (f^*)'(\ell_\delta - g \ln \delta)\,\mathrm{d}\mathbf{x} = (f^*)'(0)|\mathcal{X}| = 0.$$

The map $\ell \mapsto \int_{\mathcal{X}} \sigma_\ell(\mathbf{x})\,\mathrm{d}\mathbf{x}$ is continuous. Therefore, by the Intermediate Value Theorem, there exists $\ell_* \in \mathbb{R}$ such that

$$\int_{\mathcal{X}} \sigma_{\ell_*}(\mathbf{x})\,\mathrm{d}\mathbf{x} = 1,$$

as claimed. For the remainer of this proof, we take $\ell = \ell_*$

Next we show that the identity map is the optimal transport map for transporting any measure $\sigma \in \mathscr{P}(\mathcal{X})$ to itself, and that an optimal Kantorovich potential pair is (φ, φ^c), where $\varphi : \mathcal{X} \to \mathbb{R}$ is given by $\varphi(\mathbf{x}) = g \ln x_3$, and $\varphi^c : \mathcal{Y} \to \mathbb{R}$ is given by

$$\varphi^c(\mathbf{y}) = \min_{\mathbf{x} \in \mathcal{X}} \left(c(\mathbf{x}, \mathbf{y}) - \varphi(\mathbf{x})\right) = \min_{\mathbf{x} \in \mathcal{X}} \left(c(\mathbf{x}, \mathbf{y}) - g \ln x_3\right).$$

First we compute an explicit expression for $\varphi^c(\mathbf{y})$ for the special case $\mathbf{y} \in \mathcal{X} \cap \mathcal{Y} = \mathcal{X}$. Given $\mathbf{y} \in \mathcal{X}$, define $\theta_{\mathbf{y}} : \mathcal{X} \to \mathbb{R}$ by $\theta_{\mathbf{y}}(\mathbf{x}) = c(\mathbf{x}, \mathbf{y}) - g \ln x_3$. Then

$$\nabla\theta_{\mathbf{y}}(\mathbf{x}) = \begin{pmatrix} \frac{f_{\mathrm{cor}}^2}{y_3}(x_1 - y_1) \\ \frac{f_{\mathrm{cor}}^2}{y_3}(x_2 - y_2) \\ \frac{g}{y_3} - \frac{g}{x_3} \end{pmatrix}, \qquad D^2\theta_{\mathbf{y}}(\mathbf{x}) = \begin{pmatrix} \frac{f_{\mathrm{cor}}^2}{y_3} & 0 & 0 \\ 0 & \frac{f_{\mathrm{cor}}^2}{y_3} & 0 \\ 0 & 0 & \frac{g}{x_3^2} \end{pmatrix}.$$

Therefore $D^2\theta_{\mathbf{y}}(\mathbf{x})$ is positive definite for all $\mathbf{x} \in \mathcal{X}$. Hence $\theta_{\mathbf{y}}$ is strictly convex, its minimiser is $\mathbf{y}$, and its minimum value is

$$\varphi^c(\mathbf{y}) = \min_{\mathbf{x}\in\mathcal{X}} \theta_{\mathbf{y}}(\mathbf{x}) = \theta_{\mathbf{y}}(\mathbf{y}) = c(\mathbf{y},\mathbf{y}) - g\ln y_3 = g - g\ln y_3.$$

Note that $\varphi(\mathbf{x}) + \varphi^c(\mathbf{x}) = g$ for all $\mathbf{x} \in \mathcal{X}$ and that σ is supported on $\mathcal{X} \subset \mathcal{Y}$. Then

$$\mathcal{T}_c(\sigma,\sigma) \le \int_{\mathcal{X}} c(\mathbf{x},\mathbf{x})\,\mathrm{d}\sigma(\mathbf{x}) = g = \int_{\mathcal{X}} \varphi\,\mathrm{d}\sigma + \int_{\mathcal{Y}} \varphi^c\,\mathrm{d}\sigma \le \mathcal{T}_c(\sigma,\sigma),$$

which proves that the identity map is optimal for transporting $\sigma \in \mathscr{P}(\mathcal{X})$ to itself, and that (φ,φ^c) is an optimal Kantorovich potential pair, as claimed.

Lastly, we show that $\sigma_\ell = \arg\min_{\sigma\in\mathscr{P}_{\mathrm{ac}}(\mathcal{X})} E(\sigma,\sigma_\ell)$. If $\mathbf{x} \in \mathcal{X}$ satisfies $\sigma_\ell(\mathbf{x}) = 0$, then $0 \ge \ell - g\ln x_3 = \ell - \varphi(\mathbf{x})$. Therefore

$$\varphi(\mathbf{x}) + f'(\sigma_\ell(\mathbf{x})) = \varphi(\mathbf{x}) \ge \ell.$$

On the other hand, if $\mathbf{x} \in \mathcal{X}$ satisfies $\sigma_\ell(\mathbf{x}) > 0$, then

$$\varphi(\mathbf{x}) + f'(\sigma_\ell(\mathbf{x})) = g\ln x_3 + f'((f^*)'(\ell - g\ln x_3)) = g\ln x_3 + \ell - g\ln x_3 = \ell$$

because $f'((f^*)'(t)) = t$ for all $t > 0$. In summary,

$$\varphi(\mathbf{x}) + f'(\sigma_\ell(\mathbf{x})) \ge \ell \quad \forall\, \mathbf{x} \in \mathcal{X}, \tag{6.1}$$

and

$$\big(\varphi(\mathbf{x}) + f'(\sigma_\ell(\mathbf{x}))\big)\sigma_\ell(\mathbf{x}) = \ell\sigma_\ell(\mathbf{x}) \quad \forall\, \mathbf{x} \in \mathcal{X}. \tag{6.2}$$

By the Kantorovich Duality Theorem and the convexity of f, for all $\sigma \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})$,

$$\begin{aligned} E(\sigma,\sigma_\ell) &= \mathcal{T}_c(\sigma,\sigma_\ell) + \int_{\mathcal{X}} f(\sigma)\,\mathrm{d}\mathcal{L}^3 \\ &\ge \int_{\mathcal{X}} \varphi\,\mathrm{d}\sigma + \int_{\mathcal{Y}} \varphi^c\,\mathrm{d}\sigma_\ell + \int_{\mathcal{X}} \big(f(\sigma_\ell) + f'(\sigma_\ell)(\sigma - \sigma_\ell)\big)\,\mathrm{d}\mathcal{L}^3 \\ &= E(\sigma_\ell,\sigma_\ell) + \int_{\mathcal{X}} \big(\varphi + f'(\sigma_\ell)\big)(\sigma - \sigma_\ell)\,\mathrm{d}\mathcal{L}^3 \end{aligned} \tag{6.3}$$

since φ is an optimal Kantorovich potential for the transport from σ_ℓ to itself. Therefore, by (6.1), (6.2) and (6.3),

$$E(\sigma,\sigma_\ell) \ge E(\sigma_\ell,\sigma_\ell) + \ell\int_{\mathcal{X}} (\sigma - \sigma_\ell)\,\mathrm{d}\mathcal{L}^3 = E(\sigma_\ell,\sigma_\ell),$$

and so $\sigma_\ell = \arg\min_{\sigma\in\mathscr{P}_{\mathrm{ac}}(\mathcal{X})} E(\sigma,\sigma_\ell)$, as claimed.

In conclusion, we have shown that $\sigma_*[\sigma_\ell] = \sigma_\ell$ and $T_{\sigma_\ell} = \mathrm{id}$. Therefore the time-independent map $[0,\tau] \ni t \mapsto \alpha_t = \sigma_\ell$ is a weak solution of (1.1), as required. □

6.2. **Single seed example.** When there is only one seed, it is possible to construct an explicit solution for the case $\gamma = 2$. This is because the single Laguerre cell is the entire domain. Note that, to be physically meaningful, $\gamma \in (1,2)$. However, for these values of γ, we are not able to derive such an explicit solution.

Proposition 6.3 (Elliptic orbit of a single seed)**.** *Fix $\gamma = 2$, $\kappa = \frac{1}{2}$ and $\tau > 0$. Let $\mathcal{X} = [-a,a]\times[-b,b]\times[0,h]$, where $a,b,h > 0$. Let $\overline{\mathbf{z}} \in \mathcal{Y}$ be the initial position of the particle. Define $\mathbf{z} : (-\infty,\infty) \to \mathcal{Y}$ to be the solution of the linear ODE*

$$\begin{aligned} \dot{\mathbf{z}}(t) &= f_{\mathrm{cor}} \begin{pmatrix} 0 & -A & 0 \\ B & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{z}(t), \\ \mathbf{z}(0) &= \overline{\mathbf{z}}, \end{aligned}$$

where

$$A = 1 - \frac{|\mathcal{X}| f_{\mathrm{cor}}^2}{3\overline{z}_3} b^2, \qquad B = 1 - \frac{|\mathcal{X}| f_{\mathrm{cor}}^2}{3\overline{z}_3} a^2, \qquad |\mathcal{X}| = 4abh.$$

Assume that $\overline{z}_3 > \frac{|\mathcal{X}| f_{\mathrm{cor}}^2}{3} \max\{a^2, b^2\}$ *so that* $A > 0$ *and* $B > 0$. *Define the ellipse* $\mathcal{E} \subset \mathcal{Y}$ *by*

$$\mathcal{E} \coloneqq \{\mathbf{z}(t) : t \in (-\infty, \infty)\} = \left\{ (y_1, y_2, \overline{z}_3) \in \mathbb{R}^2 \times \{\overline{z}_3\} \,:\, \frac{y_1^2}{A} + \frac{y_2^2}{B} = \frac{\overline{z}_1^2}{A} + \frac{\overline{z}_2^2}{B} \right\}.$$

Assume additionally that $\overline{z}_3$ *is sufficiently large so that*

$$c(\mathbf{x}, \mathbf{y}) - \frac{1}{|\mathcal{X}|} \int_{\mathcal{X}} c(\cdot, \mathbf{y}) \, \mathrm{d}\mathcal{L}^3 < \frac{1}{|\mathcal{X}|} \quad \forall\, (\mathbf{x}, \mathbf{y}) \in \mathcal{X} \times \mathcal{E}. \tag{6.4}$$

Define $\alpha : [0, \tau] \to \mathscr{P}_{\mathrm{c}}(\mathcal{Y})$ *by* $\alpha_t = \delta_{\mathbf{z}(t)}$. *Then the following hold:*

(1) The optimal weight map $w_* : D \to \mathbb{R}$ *(defined in* (4.1)*) satisfies the following:*

$$w_*(\mathbf{z}(t)) = \frac{1}{|\mathcal{X}|} \left(1 + \int_{\mathcal{X}} c(\mathbf{x}, \mathbf{z}(t)) \, \mathrm{d}\mathbf{x} \right) \quad \forall\, t \in (-\infty, \infty).$$

(2) The optimal source measure $\sigma_*[\alpha_t] \in \mathscr{P}_{\mathrm{ac}}(\mathcal{X})$ *is given by*

$$\sigma_*[\alpha_t](\mathbf{x}) = w_*(\mathbf{z}(t)) - c(\mathbf{x}, \mathbf{z}(t)).$$

(3) α *is a weak solution of the compressible SG equation* (1.1). *In particular, the trajectory of the seed is the ellipse* $\mathcal{E}$. *Moreover, if* $a = b$, *then the seed moves on a circle. If* $\overline{z}_1 = \overline{z}_2 = 0$, *then the seed is stationary.*

Proof. Since $\gamma = 2$ and $\kappa = \frac{1}{2}$,

$$f^*(t) = \begin{cases} \frac{1}{2} t^2 & \text{if } t > 0, \\ 0 & \text{if } t \le 0. \end{cases}$$

Define $w : (-\infty, \infty) \to \mathbb{R}$ by

$$w(t) = \frac{1}{|\mathcal{X}|} \left(1 + \int_{\mathcal{X}} c(\mathbf{x}, \mathbf{z}(t)) \, \mathrm{d}\mathbf{x} \right).$$

Then, for all $\mathbf{x} \in \mathcal{X}$, $t \in (-\infty, \infty)$,

$$w(t) - c(\mathbf{x}, \mathbf{z}(t)) = \frac{1}{|\mathcal{X}|} \left(1 + \int_{\mathcal{X}} c(\tilde{\mathbf{x}}, \mathbf{z}(t)) \, \mathrm{d}\tilde{\mathbf{x}} \right) - c(\mathbf{x}, \mathbf{z}(t)) > 0$$

by (6.4). Therefore

$$\frac{\partial \mathcal{G}}{\partial w}(w(t), \mathbf{z}(t)) = 1 - \int_{\mathcal{X}} (f^*)'(w(t) - c(\mathbf{x}, \mathbf{z}(t))) \, \mathrm{d}\mathbf{x} = 1 - \int_{\mathcal{X}} (w(t) - c(\mathbf{x}, \mathbf{z}(t))) \, \mathrm{d}\mathbf{x} = 0,$$

hence $w(t)$ maximises the concave function $\mathcal{G}(\cdot, \mathbf{z}(t))$. Since the maximiser of $\mathcal{G}(\cdot, \mathbf{z}(t))$ is unique by Theorem 3.4, it follows that $w_*(\mathbf{z}(t)) = w(t)$, as claimed. The expression for $\sigma_*[\alpha_t]$ then follows immediately from (3.3).

It remains to prove that α is a weak solution the compressible SG equation (1.1). By Proposition 4.5, it suffices to show that $\mathbf{z}$ satisfies the following ODE:

$$\begin{aligned} \dot{z}_1(t) &= f_{\mathrm{cor}}(-z_2(t) + C_2(\mathbf{z}(t))), \\ \dot{z}_2(t) &= f_{\mathrm{cor}}(z_1(t) - C_1(\mathbf{z}(t))), \\ \dot{z}_3(t) &= 0. \end{aligned}$$

In other words, we need to show that

$$C_1(\mathbf{z}(t)) = (1 - B) z_1(t), \qquad C_2(\mathbf{z}(t)) = (1 - A) z_2(t).$$

We will just verify the expression for C_1; the expression for C_2 is similar. By Definition 4.2 and the symmetry of the domain $\mathcal{X}$,

$$\begin{aligned} C_1(\mathbf{z}(t)) &= \int_{\mathcal{X}} x_1 \, \sigma_*[\alpha_t](\mathbf{x}) \, \mathrm{d}\mathbf{x} \\ &= \int_{\mathcal{X}} x_1 \, (w_*(\mathbf{z}(t)) - c(\mathbf{x}, \mathbf{z}(t))) \, \mathrm{d}\mathbf{x} \end{aligned}$$

$$
\begin{aligned}
&= -\frac{1}{\overline{z}_3}\int_{-a}^{a}\int_{-b}^{b}\int_{0}^{h} x_1\left(\frac{f_{\mathrm{cor}}^2}{2}(x_1 - z_1(t))^2 + \frac{f_{\mathrm{cor}}^2}{2}(x_2 - z_2(t))^2 + gx_3\right)\mathrm{d}x_3\,\mathrm{d}x_2\,\mathrm{d}x_1\\
&= \frac{f_{\mathrm{cor}}^2}{\overline{z}_3}\int_{-a}^{a}\int_{-b}^{b}\int_{0}^{h} x_1^2 z_1(t)\,\mathrm{d}x_3\,\mathrm{d}x_2\,\mathrm{d}x_1\\
&= \frac{4a^3bhf_{\mathrm{cor}}^2}{3\overline{z}_3}z_1(t)\\
&= (1-B)z_1(t),
\end{aligned}
$$

as required. □

Acknowledgements

We thank Mike Cullen for stimulating discussions related to this work. DPB would like to thank the UK Engineering and Physical Sciences Research Council (EPSRC) for financial support via the grant EP/V00204X/1. CE gratefully acknowledges the support of the EPSRC via the grant EP/W522570/1, and of the German Research Foundation (DFG) through the CRC 1456 'Mathematics of Experiment', project A03. TPL is supported by The Maxwell Institute Graduate School in Modelling, Analysis, and Computation, a Centre for Doctoral Training funded by the EPSRC (grant EP/S023291/1), the Scottish Funding Council, Heriot-Watt University and the University of Edinburgh. BP gratefully acknowledges the support of the EPSRC via the grant EP/P011543/1.

Appendix A. Assumptions on the fluid domain $\mathcal{X}$

Recall that $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$ is the diffeomorphism defined by

$$\Phi(\mathbf{x}) = \left(f_{\mathrm{cor}}^{-2}\,x_1,\, f_{\mathrm{cor}}^{-2}\,x_2,\, g^{-1}\left(x_3 - \tfrac{1}{2}f_{\mathrm{cor}}^{-2}\left(x_1^2 + x_2^2\right)\right)\right).$$

In this appendix we discuss the assumption that the fluid domain $\mathcal{X}$ satisfies

$$\Phi^{-1}(\mathcal{X}) \text{ is convex.} \tag{A.1}$$

In particular, we give examples of domains $\mathcal{X}$ satisfying (A.1), we prove that (A.1) is equivalent to $\mathcal{X}$ being c-convex [35, Definition 1.2], and we prove that, given any c-Laguerre tessellation $\{L_c^i(\boldsymbol{w}, \mathbf{z})\}_{i=1}^N$, then $\{\Phi^{-1}(L_c^i(\boldsymbol{w}, \mathbf{z}))\}_{i=1}^N$ is a classical Laguerre tessellation (with respect to the quadratic cost). In particular, $\Phi^{-1}(L_c^i(\boldsymbol{w}, \mathbf{z}))$ is a convex polyhedron for all i. Moreover, we will see that the cost c satisfies Loeper's condition [35, Definition 1.1].

Remark A.1 (Interpreting Φ in terms of c-exponential maps)**.** Let $\mathbf{y}_0 = (0, 0, 1)$. It is easy to check that Φ is the c-exponential map $\exp_{\mathbf{y}_0}^c$ [35, Remark 4.4]:

$$\Phi = \exp_{\mathbf{y}_0}^c := (-D_{\mathbf{y}}c(\cdot, \mathbf{y}_0))^{-1}.$$

In other words,

$$-D_{\mathbf{y}}c(\Phi(\mathbf{x}), \mathbf{y}_0) = \mathbf{x}.$$

Note that c-exponential maps play an important role in regularity theory for semi-discrete optimal transport [35].

Example A.2 (Domains satisfying assumption (A.1))**.** Let $(a_1, a_2, a_3) \in \mathbb{R}^3$. It can be shown that any paraboloid of the form

$$\mathcal{X}_1 = \left\{\mathbf{x} \in \mathbb{R}^3 : x_3 \leq -\frac{f_{\mathrm{cor}}^2}{2g}\left((x_1 - a_1)^2 + (x_2 - a_2)^2\right) + a_3\right\} \tag{A.2}$$

satisfies (A.1). Its inverse image $\Phi^{-1}(\mathcal{X}_1)$ is a half-space of the form

$$\Phi^{-1}(\mathcal{X}_1) = \{\mathbf{p} \in \mathbb{R}^3 : p_3 \leq b_1p_1 + b_2p_2 + c\},$$

where $b_1, b_2, c \in \mathbb{R}$. In particular, it is convex. Replacing the inequality $\leq$ in (A.2) by $\geq$ gives another example of a set satisfying (A.1). Another example is the half-space

$$\mathcal{X}_2 = \{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} \cdot \mathbf{n} \geq d\}$$

where $d \in \mathbb{R}$ and $\mathbf{n} \in \mathbb{R}^3$ with $n_3 \geq 0$. Its inverse image $\Phi^{-1}(\mathcal{X}_2)$ has the following form for some $\mathbf{b} \in \mathbb{R}^3$, $c \in \mathbb{R}$:

$$\Phi^{-1}(\mathcal{X}_2) = \left\{ \mathbf{p} \in \mathbb{R}^3 : \frac{n_3}{g f_{\text{cor}}^2}(p_1^2 + p_2^2) + \mathbf{b} \cdot \mathbf{p} + c \leq 0 \right\}. \tag{A.3}$$

This is convex because it has the form $\{\mathbf{p} \in \mathbb{R}^3 : F(\mathbf{p}) \leq 0\}$ with F convex. We can build other examples by using the fact that any intersection of sets satisfying (A.1) also satisfies (A.1). For example, "blister pack" sets of the form

$$\mathcal{X}_3 = \left\{ \mathbf{x} \in \mathbb{R}^3 : 0 \leq x_3 \leq -\frac{f_{\text{cor}}^2}{2g}\left((x_1 - a_1)^2 + (x_2 - a_2)^2\right) + a_3 \right\}$$

satisfy (A.1) for all $\mathbf{a} \in \mathbb{R}^3$. It can be seen from (A.3) that the half-space $\{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} \cdot \mathbf{n} \geq d\}$ does *not* satisfy (A.1) if $n_3 < 0$.

Remark A.3 (Rectangular domains do not satisfy (A.1))**.** Rectangular domains of the form $\mathcal{X} = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$ do not satisfy assumption (A.1). These domains are perhaps the most natural from a physical point of view. However, in this case an alternative, bespoke argument can be used to prove Corollary B.1, and hence Theorems 1.1, 1.2; see [37].

Recall the following from [35, Definitions 1.1 and 1.2]: The cost c satisfies *Loeper's condition* if, for each $\mathbf{y} \in \mathcal{Y}$, there exists a diffeomorphism $\Phi_{\mathbf{y}} : \mathbb{R}^3 \to \mathbb{R}^3$ such that the map

$$\mathbb{R}^3 \ni \mathbf{x} \mapsto c(\Phi_{\mathbf{y}}(\mathbf{x}), \mathbf{y}) - c(\Phi_{\mathbf{y}}(\mathbf{x}), \mathbf{z})$$

is quasi-convex for all $\mathbf{z} \in \mathcal{Y}$ (meaning that its sublevel sets are convex). If in addition the set $\Phi_{\mathbf{y}}^{-1}(\mathcal{X})$ is convex for each $\mathbf{y} \in \mathcal{Y}$, then we say that $\mathcal{X}$ is *c-convex*.

Lemma A.4 (Loeper's condition and c-convexity)**.** *Let $\mathbf{x} \in \mathcal{X}$, $\mathbf{y}, \mathbf{z} \in \mathcal{Y}$. Then*

$$c(\Phi(\mathbf{x}), \mathbf{y}) = -\frac{1}{y_3}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ -1 \end{pmatrix} + \frac{f_{\text{cor}}^2}{2y_3}\left(y_1^2 + y_2^2\right). \tag{A.4}$$

In particular, $u : \mathbb{R}^3 \to \mathbb{R}$ defined by

$$u(\mathbf{x}) = c(\Phi(\mathbf{x}), \mathbf{y}) - c(\Phi(\mathbf{x}), \mathbf{z})$$

is affine and the cost c satisfies Loeper's condition. Moreover, $\mathcal{X}$ satisfies assumption (A.1) *if and only if $\mathcal{X}$ is c-convex.*

Proof. Equation (A.4) is a direct computation. Since u is affine, then it is quasi-convex. Therefore, by definition, c satisfies Loeper's condition (with $\Phi_{\mathbf{y}} = \Phi$ independent of $\mathbf{y}$), and assumption (A.1) is equivalent to c-convexity of $\mathcal{X}$. □

To prove that c satisfies Loeper's condition, we could have also chosen the c-exponential maps $\Phi_{\mathbf{y}} = \exp_{\mathbf{y}}^c := (-D_{\mathbf{y}} c(\cdot, \mathbf{y}))^{-1}$ depending on $\mathbf{y}$. However, the fact that Φ is independent of $\mathbf{y}$ (and that u is affine) has important consequences, as we will now see. Let

$$D_2^N := \left\{ \mathbf{y} = (\mathbf{y}^1, \ldots, \mathbf{y}^N) \in (\mathbb{R}^3)^N : \mathbf{y}^i \neq \mathbf{y}^j \text{ whenever } i \neq j \right\}.$$

Recall the following definition:

Definition A.5 (Classical Laguerre tessellations)**.** Given $(\boldsymbol{\psi}, \mathbf{y}) \in \mathbb{R}^N \times D_2^N$, the (classical) *Laguerre tessellation* of $\Phi^{-1}(\mathcal{X})$ generated by $(\boldsymbol{\psi}, \mathbf{y})$ is the partition $\{L_2^i(\boldsymbol{\psi}, \mathbf{y})\}_{i=1}^N$ defined by

$$L_2^i(\boldsymbol{\psi}, \mathbf{y}) = \left\{ \mathbf{x} \in \Phi^{-1}(\mathcal{X}) : \|\mathbf{x} - \mathbf{y}^i\|^2 - \psi^i \leq \|\mathbf{x} - \mathbf{y}^j\|^2 - \psi^j \ \forall\, j \in \{1, \ldots, N\} \right\}.$$

The Laguerre cells $L_2^i(\boldsymbol{\psi}, \mathbf{y})$ are the intersection of $\Phi^{-1}(\mathcal{X})$ with convex polyhedra. Consequently, unlike c-Laguerre cells, they can be computed very efficiently. Classical Laguerre tessellations arise in semi-discrete optimal transport problems with *quadratic* cost. The following lemma asserts that c-Laguerre tessellations are just classical Laguerre tessellations in disguise. It also gives a practical way of computing them.

Lemma A.6 (Rewriting c-Laguerre tessellations as classical Laguerre tessellations)**.** *Define* $\widehat{\mathbf{y}} : \mathcal{Y}^N \to (\mathbb{R}^3)^N$ *by* $\widehat{\mathbf{y}}(\mathbf{z}) = (\widehat{\mathbf{y}}^1(\mathbf{z}), \dots, \widehat{\mathbf{y}}^N(\mathbf{z}))$*, where*

$$\widehat{\mathbf{y}}^i(\mathbf{z}) = \frac{1}{2z_3^i} \begin{pmatrix} z_1^i \\ z_2^i \\ -1 \end{pmatrix}.$$

Define $\widehat{\boldsymbol{\psi}} : \mathbb{R}^N \times \mathcal{Y}^N \to \mathbb{R}^N$ *by*

$$\widehat{\psi^i}(\boldsymbol{w}, \mathbf{z}) = w^i + \left(\frac{z_1^i}{2z_3^i}\right)^2 + \left(\frac{z_2^i}{2z_3^i}\right)^2 + \left(\frac{1}{2z_3^i}\right)^2 - \frac{f_{\mathrm{cor}}^2}{2z_3^i}\left(\left(z_1^i\right)^2 + \left(z_2^i\right)^2\right).$$

Then, for all $i \in \{1, \dots, N\}$ *and* $(\boldsymbol{w}, \mathbf{z}) \in \mathbb{R}^N \times D^N$*,*

$$\Phi^{-1}(L_c^i(\boldsymbol{w}, \mathbf{z})) = L_2^i(\widehat{\boldsymbol{\psi}}(\boldsymbol{w}, \mathbf{z}), \widehat{\mathbf{y}}(\mathbf{z})).$$

Proof. By (A.4), for all $\mathbf{x} \in \mathbb{R}^3$, $\mathbf{z} \in \mathcal{Y}$, $\boldsymbol{w} \in \mathbb{R}^N$,

$$\begin{aligned} c(\Phi(\mathbf{x}), \mathbf{z}^i) - w^i &= -\frac{1}{z_3^i} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \cdot \begin{pmatrix} z_1^i \\ z_2^i \\ -1 \end{pmatrix} + \frac{f_{\mathrm{cor}}^2}{2z_3^i}\left(\left(z_1^i\right)^2 + \left(z_2^i\right)^2\right) - w^i \\ &= \|\mathbf{x} - \widehat{\mathbf{y}}^i(\mathbf{z})\|^2 - \widehat{\psi^i}(\boldsymbol{w}, \mathbf{z}) - \|\mathbf{x}\|^2. \end{aligned} \tag{A.5}$$

Let $\mathbf{x} \in \Phi^{-1}(\mathcal{X})$. Then

$$\begin{aligned} \mathbf{x} \in \Phi^{-1}(L_c^i(\boldsymbol{w}, \mathbf{z})) \quad &\iff \quad c(\Phi(\mathbf{x}), \mathbf{z}^i) - w^i \le c(\Phi(\mathbf{x}), \mathbf{z}^j) - w^j \quad \forall\, j \\ &\iff \quad \|\mathbf{x} - \widehat{\mathbf{y}}^i(\mathbf{z})\|^2 - \widehat{\psi^i}(\boldsymbol{w}, \mathbf{z}) \le \|\mathbf{x} - \widehat{\mathbf{y}}^j(\mathbf{z})\|^2 - \widehat{\psi^j}(\boldsymbol{w}, \mathbf{z}) \quad \forall\, j \\ &\iff \quad \mathbf{x} \in L_2^i(\widehat{\boldsymbol{\psi}}(\boldsymbol{w}, \mathbf{z}), \widehat{\mathbf{y}}(\mathbf{z})), \end{aligned}$$

as required. □

Appendix B. Proof of Lemma 3.6

In this section we prove Lemma 3.6. The main ingredient of the proof is the following lemma.

Lemma B.1 (Regularity theory for semi-discrete optimal transport)**.** *Let* $\zeta : \mathcal{X} \to \mathbb{R}$ *be continuous. Let* $U \subset \mathbb{R}^N \times D_0^N$ *be the open set*

$$U = \left\{(\boldsymbol{w}, \mathbf{z}) \in \mathbb{R}^N \times D_0^N \,:\, \mathcal{L}^3(L_c^i(\boldsymbol{w}, \mathbf{z})) > 0 \;\forall\, i \in \{1, \dots, N\}\right\}.$$

Define $\boldsymbol{\Psi} = (\Psi^1, \dots, \Psi^N) : U \to \mathbb{R}^N$ *by*

$$\Psi^i(\boldsymbol{w}, \mathbf{z}) := \int_{L_c^i(\boldsymbol{w}, \mathbf{z})} \zeta(\mathbf{x})\, \mathrm{d}\mathbf{x}.$$

Then $\boldsymbol{\Psi}$ *is continuously differentiable. Moreover, for all* $j \in \{1, \dots, N\} \setminus \{i\}$*,*

$$\frac{\partial \Psi^i}{\partial w^j}(\boldsymbol{w}, \mathbf{z}) = -\int_{L_c^i(\boldsymbol{w}, \mathbf{z}) \cap L_c^j(\boldsymbol{w}, \mathbf{z})} \frac{\zeta(\mathbf{x})}{\|\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^i) - \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^j)\|}\, \mathrm{d}\mathcal{H}^2(\mathbf{x}), \tag{B.1}$$

$$\frac{\partial \Psi^i}{\partial \mathbf{z}^j}(\boldsymbol{w}, \mathbf{z}) = \int_{L_c^i(\boldsymbol{w}, \mathbf{z}) \cap L_c^j(\boldsymbol{w}, \mathbf{z})} \frac{\nabla_{\mathbf{y}} c(\mathbf{x}, \mathbf{z}^j)\, \zeta(\mathbf{x})}{\|\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^i) - \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^j)\|}\, \mathrm{d}\mathcal{H}^2(\mathbf{x}), \tag{B.2}$$

and

$$\frac{\partial \Psi^i}{\partial w^i}(\boldsymbol{w}, \mathbf{z}) = -\sum_{j \neq i} \frac{\partial \Psi^j}{\partial w^i}(\boldsymbol{w}, \mathbf{z}), \qquad \frac{\partial \Psi^i}{\partial \mathbf{z}^i}(\boldsymbol{w}, \mathbf{z}) = -\sum_{j \neq i} \frac{\partial \Psi^j}{\partial \mathbf{z}^i}(\boldsymbol{w}, \mathbf{z}). \tag{B.3}$$

This lemma could be proved using [22, Theorem 1], by proving that our transport cost c satisfies assumptions (Diff-2) and (Cont-2) of [22]. However, this is very involved. We give a shorter proof using Lemma A.6 to reduce to the case of the classical quadratic transport cost.

Proof. Note that $\det(D\Phi) = f_{\mathrm{cor}}^{-4} g^{-1} > 0$ is constant. By the change of variables formula and Lemma A.6,

$$\Psi^i(\boldsymbol{w},\mathbf{z}) = \int_{L_c^i(\boldsymbol{w},\mathbf{z})} \zeta \,\mathrm{d}\mathcal{L}^3 = \int_{\Phi^{-1}(L_c^i(\boldsymbol{w},\mathbf{z}))} \zeta\circ\Phi \det(D\Phi)\,\mathrm{d}\mathcal{L}^3 = \det(D\Phi)\int_{L_2^i(\widehat{\boldsymbol{\psi}}(\boldsymbol{w},\mathbf{z}),\widehat{\mathbf{y}}(\mathbf{z}))} \zeta\circ\Phi\,\mathrm{d}\mathcal{L}^3.$$

Let $U_2 \subset \mathbb{R}^N \times D_2^N$ be the open set

$$U_2 = \left\{(\boldsymbol{\psi},\mathbf{y}) \in \mathbb{R}^N \times D_2^N \,:\, \mathcal{L}^3(L_2^i(\boldsymbol{\psi},\mathbf{y})) > 0 \;\forall\, i \in \{1,\dots,N\}\right\}.$$

Define $\boldsymbol{\Psi}_2 = (\Psi_2^1,\dots,\Psi_2^N) : U_2 \to \mathbb{R}^N$ by

$$\Psi_2^i(\boldsymbol{\psi},\mathbf{y}) := \int_{L_2^i(\boldsymbol{\psi},\mathbf{y})} \zeta(\Phi(\mathbf{x}))\,\mathrm{d}\mathbf{x}.$$

Then

$$\boldsymbol{\Psi}(\boldsymbol{w},\mathbf{z}) = \det(D\Phi)\boldsymbol{\Psi}_2(\widehat{\boldsymbol{\psi}}(\boldsymbol{w},\mathbf{z}),\widehat{\mathbf{y}}(\mathbf{z})). \tag{B.4}$$

By assumption (A.1) and [22, Proposition 2], $\boldsymbol{\Psi}_2$ is continuously differentiable and, for all $j \in \{1,\dots,N\} \setminus \{i\}$,

$$\frac{\partial\Psi_2^i}{\partial\psi^j}(\boldsymbol{\psi},\mathbf{y}) = -\frac{1}{2\|\mathbf{y}^i-\mathbf{y}^j\|}\int_{L_2^i(\boldsymbol{\psi},\mathbf{y})\cap L_2^j(\boldsymbol{\psi},\mathbf{y})} \zeta(\Phi(\mathbf{x}))\,\mathrm{d}\mathcal{H}^2(\mathbf{x}), \tag{B.5}$$

$$\frac{\partial\Psi_2^i}{\partial\mathbf{y}^j}(\boldsymbol{\psi},\mathbf{y}) = \frac{1}{\|\mathbf{y}^i-\mathbf{y}^j\|}\int_{L_2^i(\boldsymbol{\psi},\mathbf{y})\cap L_2^j(\boldsymbol{\psi},\mathbf{y})} (\mathbf{y}^j-\mathbf{x})\zeta(\Phi(\mathbf{x}))\,\mathrm{d}\mathcal{H}^2(\mathbf{x}), \tag{B.6}$$

and

$$\frac{\partial\Psi_2^i}{\partial\psi^i}(\boldsymbol{\psi},\mathbf{y}) = -\sum_{j\neq i}\frac{\partial\Psi_2^j}{\partial\psi^i}(\boldsymbol{\psi},\mathbf{y}), \qquad \frac{\partial\Psi_2^i}{\partial\mathbf{y}^i}(\boldsymbol{\psi},\mathbf{y}) = -\sum_{j\neq i}\frac{\partial\Psi_2^j}{\partial\mathbf{y}^i}(\boldsymbol{\psi},\mathbf{y}). \tag{B.7}$$

Note that $(\widehat{\boldsymbol{\psi}}(\boldsymbol{w},\mathbf{z}),\widehat{\mathbf{y}}(\mathbf{z})) \in U_2$ for all $(\boldsymbol{w},\mathbf{z}) \in U$. Therefore, by the Chain Rule and the smoothness of $\widehat{\boldsymbol{\psi}}$ and $\widehat{\mathbf{y}}$, it follows that $\boldsymbol{\Psi}$ is also continuously differentiable.

Next we compute $\partial\Psi^i/\partial w^j$ for $i \neq j$. For brevity let

$$L_2^{ij} = L_2^i(\widehat{\boldsymbol{\psi}}(\boldsymbol{w},\mathbf{z}),\widehat{\mathbf{y}}(\mathbf{z})) \cap L_2^j(\widehat{\boldsymbol{\psi}}(\boldsymbol{w},\mathbf{z}),\widehat{\mathbf{y}}(\mathbf{z})).$$

Then

$$\begin{aligned}\frac{\partial\Psi^i}{\partial w^j}(\boldsymbol{w},\mathbf{z}) &= \det(D\Phi)\sum_{k=1}^N \frac{\partial\Psi_2^i}{\partial\psi^k}(\widehat{\boldsymbol{\psi}}(\boldsymbol{w},\mathbf{z}),\widehat{\mathbf{y}}(\mathbf{z}))\frac{\partial\widehat{\psi}^k}{\partial w^j}(\boldsymbol{w},\mathbf{z})\\ &= \det(D\Phi)\frac{\partial\Psi_2^i}{\partial\psi^j}(\widehat{\boldsymbol{\psi}}(\boldsymbol{w},\mathbf{z}),\widehat{\mathbf{y}}(\mathbf{z}))\\ &= -\frac{\det(D\Phi)}{2\|\widehat{\mathbf{y}}^i(\mathbf{z})-\widehat{\mathbf{y}}^j(\mathbf{z})\|}\int_{L_2^{ij}} \zeta\circ\Phi\,\mathrm{d}\mathcal{H}^2.\end{aligned} \tag{B.8}$$

By the generalised area formula (see, e.g., [3, Theorem 2.91] or [11, Equation (2.7)]),

$$\int_{L_2^{ij}} \zeta\circ\Phi\,\mathrm{d}\mathcal{H}^2 = \int_{\Phi(L_2^{ij})} \frac{\zeta}{J_\Phi\circ\Phi^{-1}}\,\mathrm{d}\mathcal{H}^2 = \int_{L_c^i(\boldsymbol{w},\mathbf{z})\cap L_c^j(\boldsymbol{w},\mathbf{z})} \frac{\zeta}{J_\Phi\circ\Phi^{-1}}\,\mathrm{d}\mathcal{H}^2, \tag{B.9}$$

where $J_\Phi : \mathbb{R}^3 \to \mathbb{R}$ is the Jacobian

$$J_\Phi = \|(D\Phi)^{-T}\mathbf{n}\|\det(D\Phi), \qquad \mathbf{n} = \frac{\widehat{\mathbf{y}}^j(\mathbf{z})-\widehat{\mathbf{y}}^i(\mathbf{z})}{\|\widehat{\mathbf{y}}^j(\mathbf{z})-\widehat{\mathbf{y}}^i(\mathbf{z})\|}.$$

Note that $\mathbf{n}$ is the outer unit normal to $L_2^i(\widehat{\boldsymbol{\psi}}(\boldsymbol{w},\mathbf{z}),\widehat{\mathbf{y}}(\mathbf{z}))$ on the face L_2^{ij}. We will prove below that $J_\Phi\circ\Phi^{-1}$ is uniformly bounded from below by a positive constant, hence the integrals in (B.9) are well defined and the generalised area formula is valid.

By equation (A.5),

$$c(\mathbf{x},\mathbf{z}^i) - \boldsymbol{w}^i = -2\Phi^{-1}(\mathbf{x})\cdot\widehat{\mathbf{y}}^i(\mathbf{z}) + \|\widehat{\mathbf{y}}^i(\mathbf{z})\|^2 - \widehat{\psi}^i(\boldsymbol{w},\mathbf{z}). \tag{B.10}$$

Differentiating this with respect to $\mathbf{x}$ gives

$$\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^i) = -2(D\Phi^{-1}(\mathbf{x}))^T \widehat{\mathbf{y}}^i(\mathbf{z}) = -2(D\Phi(\Phi^{-1}(\mathbf{x}))^{-T} \widehat{\mathbf{y}}^i(\mathbf{z}).$$

Therefore

$$\|\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^i) - \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^j)\| = \frac{2J_\Phi(\Phi^{-1}(\mathbf{x}))\|\widehat{\mathbf{y}}^j(\mathbf{z}) - \widehat{\mathbf{y}}^i(\mathbf{z})\|}{\det(D\Phi)}. \tag{B.11}$$

Combining equations (B.8), (B.9) and (B.11) proves (B.1), as required.

Next we prove a lower bound on $J_\Phi \circ \Phi^{-1}$. Let $\mathbf{z} \in D_0^N$ and $\eta = \min_{i \neq j} |z_3^i - z_3^j| > 0$. Then

$$\|\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^i) - \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^j)\| \geq |\partial_{x_3} c(\mathbf{x}, \mathbf{z}^i) - \partial_{x_3} c(\mathbf{x}, \mathbf{z}^j)| = g\frac{|z_3^i - z_3^j|}{z_3^i z_3^j} \geq g\eta\delta^2 \tag{B.12}$$

since $\mathbf{z}^i, \mathbf{z}^j \in \mathcal{Y} = \mathbb{R}^2 \times (\delta, 1/\delta)$. By (B.11), for all $\mathbf{x} \in \mathcal{X}$,

$$J_\Phi(\Phi^{-1}(\mathbf{x})) = \frac{\det(D\Phi)\|\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^i) - \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^j)\|}{2\|\widehat{\mathbf{y}}^j(\mathbf{z}) - \widehat{\mathbf{y}}^i(\mathbf{z})\|} \geq \frac{f_{\text{cor}}^{-4}\,\eta\,\delta^2}{2\|\widehat{\mathbf{y}}^j(\mathbf{z}) - \widehat{\mathbf{y}}^i(\mathbf{z})\|},$$

which is the required uniform lower bound.

Next we prove (B.2). Differentiating (B.10) with respect to $\mathbf{z}^i$ gives

$$\nabla_{\mathbf{y}} c(\mathbf{x}, \mathbf{z}^i) = 2\left(\frac{\partial \widehat{\mathbf{y}}^i}{\partial \mathbf{z}^i}(\mathbf{z})\right)^T (-\Phi^{-1}(\mathbf{x}) + \widehat{\mathbf{y}}^i(\mathbf{z})) - \frac{\partial \widehat{\psi}^i}{\partial \mathbf{z}^i}(\boldsymbol{w}, \mathbf{z}). \tag{B.13}$$

By (B.4), for all $i \neq j$,

$$\begin{aligned}
&\frac{\partial \Psi^i}{\partial \mathbf{z}^j}(\boldsymbol{w}, \mathbf{z}) = \\
&= \det(D\Phi)\left(\frac{\partial \Psi_2^i}{\partial \psi^j}(\widehat{\boldsymbol{\psi}}(\boldsymbol{w}, \mathbf{z}), \widehat{\mathbf{y}}(\mathbf{z}))\frac{\partial \widehat{\psi}^j}{\partial \mathbf{z}^j}(\boldsymbol{w}, \mathbf{z}) + \left(\frac{\partial \widehat{\mathbf{y}}^j}{\partial \mathbf{z}^j}(\mathbf{z})\right)^T \frac{\partial \Psi_2^i}{\partial \mathbf{y}^j}(\widehat{\boldsymbol{\psi}}(\boldsymbol{w}, \mathbf{z}), \widehat{\mathbf{y}}(\mathbf{z}))\right) \\
&= -\frac{\det(D\Phi)}{2\|\widehat{\mathbf{y}}^i(\mathbf{z}) - \widehat{\mathbf{y}}^j(\mathbf{z})\|}\int_{L_2^{ij}} \zeta(\Phi(\mathbf{x}))\,\mathrm{d}\mathcal{H}^2(\mathbf{x})\,\frac{\partial \widehat{\psi}^j}{\partial \mathbf{z}^j}(\boldsymbol{w}, \mathbf{z}) \\
&\quad + \left(\frac{\partial \widehat{\mathbf{y}}^j}{\partial \mathbf{z}^j}(\mathbf{z})\right)^T \frac{\det(D\Phi)}{\|\widehat{\mathbf{y}}^i(\mathbf{z}) - \widehat{\mathbf{y}}^j(\mathbf{z})\|}\int_{L_2^{ij}} (\widehat{\mathbf{y}}^j(\mathbf{z}) - \mathbf{x})\zeta(\Phi(\mathbf{x}))\,\mathrm{d}\mathcal{H}^2(\mathbf{x}) \\
&= -\frac{\det(D\Phi)}{2\|\widehat{\mathbf{y}}^i(\mathbf{z}) - \widehat{\mathbf{y}}^j(\mathbf{z})\|}\int_{\Phi(L_2^{ij})} \frac{\zeta(\mathbf{x})}{J_\Phi(\Phi^{-1}(\mathbf{x}))}\,\mathrm{d}\mathcal{H}^2(\mathbf{x})\,\frac{\partial \widehat{\psi}^j}{\partial \mathbf{z}^j}(\boldsymbol{w}, \mathbf{z}) \\
&\quad + \left(\frac{\partial \widehat{\mathbf{y}}^j}{\partial \mathbf{z}^j}(\mathbf{z})\right)^T \frac{\det(D\Phi)}{\|\widehat{\mathbf{y}}^i(\mathbf{z}) - \widehat{\mathbf{y}}^j(\mathbf{z})\|}\int_{\Phi(L_2^{ij})} \frac{(\widehat{\mathbf{y}}^j(\mathbf{z}) - \Phi^{-1}(\mathbf{x}))\zeta(\mathbf{x})}{J_\Phi(\Phi^{-1}(\mathbf{x}))}\,\mathrm{d}\mathcal{H}^2(\mathbf{x}) \\
&= \frac{\det(D\Phi)}{2\|\widehat{\mathbf{y}}^i(\mathbf{z}) - \widehat{\mathbf{y}}^j(\mathbf{z})\|}\int_{\Phi(L_2^{ij})} \frac{\zeta}{J_\Phi \circ \Phi^{-1}}\left(2\left(\frac{\partial \widehat{\mathbf{y}}^j}{\partial \mathbf{z}^j}(\mathbf{z})\right)^T (\widehat{\mathbf{y}}^j(\mathbf{z}) - \Phi^{-1}) - \frac{\partial \widehat{\psi}^j}{\partial \mathbf{z}^j}(\boldsymbol{w}, \mathbf{z})\right)\mathrm{d}\mathcal{H}^2.
\end{aligned}$$

Combining this with (B.11) and (B.13) proves (B.2), as desired.

Finally, (B.3) follows by differentiating

$$\sum_{j=1}^N \Psi^j(\boldsymbol{w}, \mathbf{z}) = \int_{\mathcal{X}} \zeta(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

with respect to w^i and $\mathbf{z}^i$. □

Remark B.2 (The well-preparedness assumption)**.** The estimate (B.12), and hence Lemma B.1, could be proved under the weaker assumption that the seeds are merely distinct, $\mathbf{z} \in D^N$, without the well-preparedness assumption $\mathbf{z} \in D_0^N$. This is done by keeping the x_1 and x_2 components of $\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^i) - \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}^j)$ in the inequality (B.12). It follows that Lemma 3.6 also holds without the well-preparedness assumption, i.e., we could replace D_0^N by D^N in the definition of U in both Lemma B.1 and Lemma 3.6. However, since the well-preparedness

assumption is crucial for the proof of Proposition 4.7 (to ensure that seeds do not collide), we keep it in Lemmas B.1 and 3.6 to simplify the proof of (B.12).

Finally we are in a position to prove Lemma 3.6.

Proof of Lemma 3.6. For each $i \in \{1, \dots, N\}$, define the auxiliary function $h^i : U \times U \to \mathbb{R}$ by

$$h^i(\boldsymbol{w}_1, \mathbf{z}_1, \boldsymbol{w}_2, \mathbf{z}_2) = \int_{L_c^i(\boldsymbol{w}_1, \mathbf{z}_1)} \zeta(\mathbf{x}, \boldsymbol{w}_2, \mathbf{z}_2)\, \mathrm{d}\mathbf{x}.$$

Then $\Psi^i(\boldsymbol{w}, \mathbf{z}) = h^i(\boldsymbol{w}, \mathbf{z}, \boldsymbol{w}, \mathbf{z})$. Fix an arbitrary point $(\boldsymbol{w}_0, \mathbf{z}_0) \in U$ and let $K \subset U$ be a compact neighbourhood of $(\boldsymbol{w}_0, \mathbf{z}_0)$ such that $|z_3^i - z_3^j| \geq \eta$ for all $i, j \in \{1, \dots, N\}$, $i \neq j$, for some $\eta > 0$. We will prove that h^i is continuously differentiable at $(\boldsymbol{w}_0, \mathbf{z}_0, \boldsymbol{w}_0, \mathbf{z}_0)$, and hence Ψ^i is continuously differentiable at $(\boldsymbol{w}_0, \mathbf{z}_0)$. To be precise, firstly we will prove that

$$\forall\, (\boldsymbol{w}_2, \mathbf{z}_2) \in U,\ \frac{\partial h^i}{\partial \boldsymbol{w}_1}(\cdot, \cdot, \boldsymbol{w}_2, \mathbf{z}_2) \text{ and } \frac{\partial h^i}{\partial \mathbf{z}_1}(\cdot, \cdot, \boldsymbol{w}_2, \mathbf{z}_2) \text{ are continuous in } U, \tag{B.14}$$

$$\left\{ \frac{\partial h^i}{\partial \boldsymbol{w}_1}(\boldsymbol{w}_1, \mathbf{z}_1, \cdot, \cdot) \right\}_{(\boldsymbol{w}_1, \mathbf{z}_1) \in K} \text{ is uniformly Lipschitz continuous in } K, \tag{B.15}$$

$$\left\{ \frac{\partial h^i}{\partial \mathbf{z}_1}(\boldsymbol{w}_1, \mathbf{z}_1, \cdot, \cdot) \right\}_{(\boldsymbol{w}_1, \mathbf{z}_1) \in K} \text{ is uniformly Lipschitz continuous in } K. \tag{B.16}$$

Here *uniformly Lipschitz* means that the Lipschitz constants are independent of $(\boldsymbol{w}_1, \mathbf{z}_1) \in K$. It follows from (B.14)–(B.16) that the partial derivatives $\partial h^i / \partial \boldsymbol{w}_1$ and $\partial h^i / \partial \mathbf{z}_1$ are continuous at $(\boldsymbol{w}_0, \mathbf{z}_0, \boldsymbol{w}_0, \mathbf{z}_0)$. Secondly we will prove that

$$\forall\, (\boldsymbol{w}_1, \mathbf{z}_1) \in U,\ \frac{\partial h^i}{\partial \boldsymbol{w}_2}(\boldsymbol{w}_1, \mathbf{z}_1, \cdot, \cdot) \text{ and } \frac{\partial h^i}{\partial \mathbf{z}_2}(\boldsymbol{w}_1, \mathbf{z}_1, \cdot, \cdot) \text{ are continuous in } U, \tag{B.17}$$

$$\left\{ \frac{\partial h^i}{\partial \boldsymbol{w}_2}(\cdot, \cdot, \boldsymbol{w}_2, \mathbf{z}_2) \right\}_{(\boldsymbol{w}_2, \mathbf{z}_2) \in K} \text{ and } \left\{ \frac{\partial h^i}{\partial \mathbf{z}_2}(\cdot, \cdot, \boldsymbol{w}_2, \mathbf{z}_2) \right\}_{(\boldsymbol{w}_2, \mathbf{z}_2) \in K} \text{ are equicontinuous in } K. \tag{B.18}$$

Therefore the partial derivatives $\partial h^i / \partial \boldsymbol{w}_2$ and $\partial h^i / \partial \mathbf{z}_2$ are also continuous at $(\boldsymbol{w}_0, \mathbf{z}_0, \boldsymbol{w}_0, \mathbf{z}_0)$.

Step 1: proof of (B.14)–(B.16). By Lemma B.1, the function $h^i(\cdot, \cdot, \boldsymbol{w}_2, \mathbf{z}_2)$ is continuously differentiable in U for all $(\boldsymbol{w}_2, \mathbf{z}_2) \in U$. (Here we have used the assumption in Lemma 3.6 that $\zeta(\cdot, \boldsymbol{w}_2, \mathbf{z}_2)$ is continuous.) This proves (B.14). Moreover, we read off from Lemma B.1 that, for all $j \in \{1, \dots, N\} \setminus \{i\}$,

$$\frac{\partial h^i}{\partial w_1^j}(\boldsymbol{w}_1, \mathbf{z}_1, \boldsymbol{w}_2, \mathbf{z}_2) = -\int_{L_c^i(\boldsymbol{w}_1, \mathbf{z}_1) \cap L_c^j(\boldsymbol{w}_1, \mathbf{z}_1)} \frac{\zeta(\mathbf{x}, \boldsymbol{w}_2, \mathbf{z}_2)}{\|\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}_1^i) - \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}_1^j)\|}\, \mathrm{d}\mathcal{H}^2(\mathbf{x}), \tag{B.19}$$

$$\frac{\partial h^i}{\partial \mathbf{z}_1^j}(\boldsymbol{w}_1, \mathbf{z}_1, \boldsymbol{w}_2, \mathbf{z}_2) = \int_{L_c^i(\boldsymbol{w}_1, \mathbf{z}_1) \cap L_c^j(\boldsymbol{w}_1, \mathbf{z}_1)} \frac{\nabla_{\mathbf{y}} c(\mathbf{x}, \mathbf{z}_1^j)\, \zeta(\mathbf{x}, \boldsymbol{w}_2, \mathbf{z}_2)}{\|\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}_1^i) - \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}_1^j)\|}\, \mathrm{d}\mathcal{H}^2(\mathbf{x}), \tag{B.20}$$

and

$$\frac{\partial h^i}{\partial w_1^i}(\boldsymbol{w}_1, \mathbf{z}_1, \boldsymbol{w}_2, \mathbf{z}_2) = -\sum_{j \neq i} \frac{\partial h^j}{\partial w_1^i}(\boldsymbol{w}_1, \mathbf{z}_1, \boldsymbol{w}_2, \mathbf{z}_2), \tag{B.21}$$

$$\frac{\partial h^i}{\partial \mathbf{z}_1^i}(\boldsymbol{w}_1, \mathbf{z}_1, \boldsymbol{w}_2, \mathbf{z}_2) = -\sum_{j \neq i} \frac{\partial h^j}{\partial \mathbf{z}_1^i}(\boldsymbol{w}_1, \mathbf{z}_1, \boldsymbol{w}_2, \mathbf{z}_2). \tag{B.22}$$

Next we prove (B.15). For all $\mathbf{x} \in \mathcal{X}$, $(\boldsymbol{w}_1, \mathbf{z}_1) \in K$,

$$\|\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}_1^i) - \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{z}_1^k)\| \geq g\eta\delta^2$$

by (B.12). For all $(\boldsymbol{w}_1, \mathbf{z}_1)$, $(\boldsymbol{w}_2, \mathbf{z}_2)$, $(\widetilde{\boldsymbol{w}}_2, \widetilde{\mathbf{z}}_2) \in K$ and all $i, j \in \{1, \dots, N\}$, $i \neq j$,

$$\left| \frac{\partial h^i}{\partial w_1^j}(\boldsymbol{w}_1, \mathbf{z}_1, \boldsymbol{w}_2, \mathbf{z}_2) - \frac{\partial h^i}{\partial w_1^j}(\boldsymbol{w}_1, \mathbf{z}_1, \widetilde{\boldsymbol{w}}_2, \widetilde{\mathbf{z}}_2) \right|$$

$$\leq \int_{L_c^i(\boldsymbol{w}_1,\mathbf{z}_1)\cap L_c^j(\boldsymbol{w}_1,\mathbf{z}_1)} \frac{|\zeta(\mathbf{x},\boldsymbol{w}_2,\mathbf{z}_2)-\zeta(\mathbf{x},\widetilde{\boldsymbol{w}}_2,\widetilde{\mathbf{z}}_2)|}{\|\nabla_{\mathbf{x}}c(\mathbf{x},\mathbf{z}_1^i)-\nabla_{\mathbf{x}}c(\mathbf{x},\mathbf{z}_1^j)\|}\,\mathrm{d}\mathcal{H}^2(\mathbf{x})$$
$$\leq \mathcal{H}^2(L_c^i(\boldsymbol{w}_1,\mathbf{z}_1)\cap L_c^j(\boldsymbol{w}_1,\mathbf{z}_1))\frac{L(K)}{g\eta\delta^2}\|(\boldsymbol{w}_2,\mathbf{z}_2)-(\widetilde{\boldsymbol{w}}_2,\widetilde{\mathbf{z}}_2)\|, \tag{B.23}$$

where $L(K)$ was defined in the statement of Lemma 3.6. Define

$$f^{ij}=\left\{\mathbf{x}\in\mathbb{R}^3 : c(\mathbf{x},\mathbf{z}_1^i)-w_1^i=c(\mathbf{x},\mathbf{z}_1^j)-w_1^j\right\}.$$

Note that $L_c^i(\boldsymbol{w}_1,\mathbf{z}_1)\cap L_c^j(\boldsymbol{w}_1,\mathbf{z}_1)\subseteq f^{ij}$. By [25, Theorem 2.8],

$$\begin{aligned}\mathcal{H}^2(L_c^i(\boldsymbol{w}_1,\mathbf{z}_1)\cap L_c^j(\boldsymbol{w}_1,\mathbf{z}_1)) &\leq \mathcal{H}^2(\mathcal{X}\cap f^{ij})\\ &\leq \mathrm{Lip}(\Phi|_{\mathcal{X}})^2\mathcal{H}^2\left(\Phi^{-1}(\mathcal{X}\cap f^{ij})\right)\\ &= \mathrm{Lip}(\Phi|_{\mathcal{X}})^2\mathcal{H}^2\left(\Phi^{-1}(\mathcal{X})\cap\Phi^{-1}(f^{ij})\right)\\ &\leq \mathrm{Lip}(\Phi|_{\mathcal{X}})^2\,\pi\left(\frac{\mathrm{diam}(\Phi^{-1}(\mathcal{X}))}{2}\right)^2,\end{aligned} \tag{B.24}$$

where in the last line we used the fact that $\Phi^{-1}(f^{ij})$ is a plane (by Lemma A.4) and the isodiametric inequality. Combining (B.23) and (B.24) proves that $\partial h^i/\partial w_1^j(\boldsymbol{w}_1,\mathbf{z}_1,\cdot,\cdot)$ is Lipschitz continuous in K with Lipschitz constant independent of $(\boldsymbol{w}_1,\mathbf{z}_1)\in K$. By (B.21), the same holds for $\partial h^i/\partial w_1^i(\boldsymbol{w}_1,\mathbf{z}_1,\cdot,\cdot)$. This proves (B.15). The proof of (B.16) is very similar.

Step 2: proof of (B.17), (B.18). First we show that the partial derivative $\partial h^i/\partial\boldsymbol{w}_2$ exists. Fix $(\boldsymbol{w}_1,\mathbf{z}_1,\boldsymbol{w}_2,\mathbf{z}_2)\in U\times U$ and let $k\in\{1,\dots,N\}$ and $\varepsilon\in(0,1)$. To show that h^i is differentiable with respect to $\boldsymbol{w}_2$, consider the difference quotient

$$\frac{h^i(\boldsymbol{w}_1,\mathbf{z}_1,\boldsymbol{w}_2+\varepsilon\mathbf{e}_k,\mathbf{z}_2)-h^i(\boldsymbol{w}_1,\mathbf{z}_1,\boldsymbol{w}_2,\mathbf{z}_2)}{\varepsilon}=\int_{L_c^i(\boldsymbol{w}_1,\mathbf{z}_1)}\frac{\zeta(\mathbf{x},\boldsymbol{w}_2+\varepsilon\mathbf{e}_k,\mathbf{z}_2)-\zeta(\mathbf{x},\boldsymbol{w}_2,\mathbf{z}_2)}{\varepsilon}\,\mathrm{d}\mathbf{x},$$

where $\{\mathbf{e}_j\}_{j=1}^N$ is the standard basis of $\mathbb{R}^N$. Note that

$$\lim_{\varepsilon\to0}\frac{\zeta(\mathbf{x},\boldsymbol{w}_2+\varepsilon\mathbf{e}_k,\mathbf{z}_2)-\zeta(\mathbf{x},\boldsymbol{w}_2,\mathbf{z}_2)}{\varepsilon}=\frac{\partial\zeta}{\partial\boldsymbol{w}}(\mathbf{x},\boldsymbol{w}_2,\mathbf{z}_2)$$

for all $\mathbf{x}\in S(\boldsymbol{w}_2,\mathbf{z}_2)$, in particular, for $\mathcal{L}^3$-almost every $\mathbf{x}\in L_c^i(\boldsymbol{w}_1,\mathbf{z}_1)$. Moreover,

$$\left|\frac{\zeta(\mathbf{x},\boldsymbol{w}_2+\varepsilon\mathbf{e}_k,\mathbf{z}_2)-\zeta(\mathbf{x},\boldsymbol{w}_2,\mathbf{z}_2)}{\varepsilon}\right|\leq L(\widetilde{K}),$$

where $\widetilde{K}=\{(\boldsymbol{w}_2+\rho\mathbf{e}_k,\mathbf{z}_2):\rho\in[0,1]\}$ and L was defined in the statement of Lemma 3.6. Therefore, by the Lebesgue Dominated Convergence Theorem,

$$\frac{\partial h^i}{\partial\boldsymbol{w}_2}(\boldsymbol{w}_1,\mathbf{z}_1,\boldsymbol{w}_2,\mathbf{z}_2)=\int_{L_c^i(\boldsymbol{w}_1,\mathbf{z}_1)}\frac{\partial\zeta}{\partial\boldsymbol{w}}(\mathbf{x},\boldsymbol{w}_2,\mathbf{z}_2)\,\mathrm{d}\mathbf{x}. \tag{B.25}$$

Similarly,

$$\frac{\partial h^i}{\partial\mathbf{z}_2}(\boldsymbol{w}_1,\mathbf{z}_1,\boldsymbol{w}_2,\mathbf{z}_2)=\int_{L_c^i(\boldsymbol{w}_1,\mathbf{z}_1)}\frac{\partial\zeta}{\partial\mathbf{z}}(\mathbf{x},\boldsymbol{w}_2,\mathbf{z}_2)\,\mathrm{d}\mathbf{x}. \tag{B.26}$$

Then (B.17) follows from the assumption that the partial derivatives $\partial\zeta/\partial\boldsymbol{w}(\mathbf{x},\cdot,\cdot)$, $\partial\zeta/\partial\mathbf{z}(\mathbf{x},\cdot,\cdot)$ are continuous at $(\boldsymbol{w}_2,\mathbf{z}_2)$ for all $\mathbf{x}\in S(\boldsymbol{w}_2,\mathbf{z}_2)$, along with another application of the Lebesgue Dominated Convergence Theorem.

Finally we prove (B.18). For all $(\boldsymbol{w}_1,\mathbf{z}_1)$, $(\widetilde{\boldsymbol{w}}_1,\widetilde{\mathbf{z}}_1)$, $(\boldsymbol{w}_2,\mathbf{z}_2)\in K$,

$$\begin{aligned}&\left\|\frac{\partial h^i}{\partial\boldsymbol{w}_2}(\widetilde{\boldsymbol{w}}_1,\widetilde{\mathbf{z}}_1,\boldsymbol{w}_2,\mathbf{z}_2)-\frac{\partial h^i}{\partial\boldsymbol{w}_2}(\boldsymbol{w}_1,\mathbf{z}_1,\boldsymbol{w}_2,\mathbf{z}_2)\right\|\\ &\quad=\left\|\int_{L_c^i(\widetilde{\boldsymbol{w}}_1,\widetilde{\mathbf{z}}_1)}\frac{\partial\zeta}{\partial\boldsymbol{w}}(\mathbf{x},\boldsymbol{w}_2,\mathbf{z}_2)\,\mathrm{d}\mathbf{x}-\int_{L_c^i(\boldsymbol{w}_1,\mathbf{z}_1)}\frac{\partial\zeta}{\partial\boldsymbol{w}}(\mathbf{x},\boldsymbol{w}_2,\mathbf{z}_2)\,\mathrm{d}\mathbf{x}\right\|\end{aligned}$$

$$\leq \max_{\mathbf{x}\in\mathcal{X}} \left\| \frac{\partial \zeta}{\partial \boldsymbol{w}}(\mathbf{x}, \boldsymbol{w}_2, \mathbf{z}_2) \right\| \left\| \chi_{L_c^i(\widetilde{\boldsymbol{w}}_1, \widetilde{\mathbf{z}}_1)} - \chi_{L_c^i(\boldsymbol{w}_1, \mathbf{z}_1)} \right\|_{L^1(\mathcal{X})}$$
$$\leq L(K) \left\| \chi_{L_c^i(\widetilde{\boldsymbol{w}}_1, \widetilde{\mathbf{z}}_1)} - \chi_{L_c^i(\boldsymbol{w}_1, \mathbf{z}_1)} \right\|_{L^1(\mathcal{X})}. \tag{B.27}$$

It can be read off from the proof of [40, Proposition 38(vii)] that

$$\|\chi_{L_c^i(\widetilde{\boldsymbol{w}}_1, \widetilde{\mathbf{z}}_1)} - \chi_{L_c^i(\boldsymbol{w}_1, \mathbf{z}_1)}\|_{L^1(\mathcal{X})} \to 0 \quad \text{as} \quad (\widetilde{\boldsymbol{w}}_1, \widetilde{\mathbf{z}}_1) \to (\boldsymbol{w}_1, \mathbf{z}_1).$$

Then, by (B.27), $\partial h^i / \partial \boldsymbol{w}_2(\cdot, \cdot, \boldsymbol{w}_2, \mathbf{z}_2)$ is continuous on the compact set K, hence uniformly continuous, and moreover its modulus of continuity is independent of $(\boldsymbol{w}_2, \mathbf{z}_2)$. The same holds for $\partial h^i / \partial \mathbf{z}_2(\cdot, \cdot, \boldsymbol{w}_2, \mathbf{z}_2)$. This proves (B.18).

Step 3. The expressions for the partial derivatives of Ψ^i in the statement of Lemma 3.6 can be read off immediately from (B.19)–(B.22), (B.25) and (B.26). □

References

[1] L. Ambrosio, M. Colombo, G. De Philippis, and A. Figalli. Existence of Eulerian solutions to the semi-geostrophic equations in physical space: The 2-dimensional periodic case. *Communications in Partial Differential Equations*, 37(12):2209–2227, 2012.

[2] L. Ambrosio, M. Colombo, G. De Philippis, and A. Figalli. A global existence result for the semigeostrophic equations in three dimensional convex domains. *Discrete & Continuous Dynamical Systems - A*, 34(4):1251–1268, 2014.

[3] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford, 2000.

[4] L. Ambrosio and W. Gangbo. Hamiltonian ODEs in the Wasserstein space of probability measures. *Communications on Pure and Applied Mathematics*, 61(1):18–53, 2008.

[5] J.-D. Benamou and Y. Brenier. Weak existence for the semi-geostrophic equations formulated as a coupled Monge–Ampère transport problem. *SIAM Journal on Applied Mathematics*, 58(5):1450–1461, 1998.

[6] J.-D. Benamou, C. J. Cotter, J. M. J. Francis, and H. Malamut. Entropic approximations of the semigeostrophic shallow water equations. *arXiv:2507.16415*, 2025.

[7] J.-D. Benamou, C. J. Cotter, and H. Malamut. Entropic optimal transport solutions of the semi-geostrophic equations. *Journal of Computational Physics*, 500:112745, 2024.

[8] D. P. Bourne, C. P. Egan, B. Pelloni, and M. Wilkinson. Semi-discrete optimal transport methods for the semi-geostrophic equations. *Calculus of Variations and Partial Differential Equations*, 61(1):39, 2022.

[9] D. P. Bourne, B. Schmitzer, and B. Wirth. Semi-discrete unbalanced optimal transport and quantization. *European Journal of Applied Mathematics*, page 1–43, 2025.

[10] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on Pure and Applied Mathematics*, 44(4):375–417, 1991.

[11] F. Cagnetti, M. G. Mora, and M. Morini. A second order minimality condition for the Mumford-Shah functional. *Calculus of Variations and Partial Differential Equations*, 33:37–74, 2008.

[12] G. Carlier and H. Malamut. Well-posedness and convergence of entropic approximation of semi-geostrophic equations. *Discrete and Continuous Dynamical Systems*, 45(12):4697–4711, 2025.

[13] J. Cheng, M. J. P. Cullen, and M. Feldman. Classical solutions to semi-geostrophic system with variable Coriolis parameter. *Archive for Rational Mechanics and Analysis*, 227:215–272, 2018.

[14] M. J. P. Cullen. The use of semi-geostrophic theory to diagnose the behaviour of an atmospheric GCM. *Fluids*, 3(4):72, 2018.

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

[16] M. J. P. Cullen and M. Feldman. Lagrangian solutions of semi-geostrophic equations in physical space. *SIAM Journal on Mathematical Analysis*, 37(5):1371–1395, 2006.

[17] 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(3):241–273, 2001.

[18] M. J. P. Cullen, W. Gangbo, and G. Pisante. The semi-geostrophic equations discretized in reference and dual variables. *Archive for Rational Mechanics and Analysis*, 185(2):341–363, 2007.

[19] M. J. P. Cullen, D. K. Gilbert, and B. Pelloni. Solutions of the fully compressible semi-geostrophic system. *Communications in Partial Differential Equations*, 39(4):591–625, 2014.

[20] M. J. P. Cullen and H. Maroofi. The fully compressible semi-geostrophic system from meteorology. *Archive for Rational Mechanics and Analysis*, 167(4):309–336, 2003.

[21] M. J. P. Cullen and R. J. Purser. An extended Lagrangian theory of semi-geostrophic frontogenesis. *Journal of the Atmospheric Sciences*, 41(9):1477–1497, 1984.

[22] F. de Gournay, J. Kahn, and L. Lebrat. Differentiation and regularity of semi-discrete optimal transport with respect to the parameters of the discrete measure. *Numerische Mathematik*, 141(2):429–453, 2019.

[23] C. P. Egan, D. P. Bourne, C. J. Cotter, M. J. P. Cullen, B. Pelloni, S. M. Roper, and M. Wilkinson. A new implementation of the geometric method for solving the Eady slice equations. *Journal of Computational Physics*, 469:111542, 2022.

[24] A. Eliassen. *The Quasi-static Equations of Motion with Pressure as Independent Variable*. Grøndahl, 1949.

[25] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions, Revised Edition*. Chapman & Hall, 2015.

[26] J. Faria. On the existence and weak stability of solutions to the compressible semi-geostrophic equations. *Journal of Mathematical Analysis and Applications*, 406(2):447–463, 2013.

[27] M. Feldman and A. Tudorascu. On Lagrangian solutions for the semi-geostrophic system with singular initial data. *SIAM Journal on Mathematical Analysis*, 45(3):1616–1640, 2013.

[28] M. Feldman and A. Tudorascu. On the semi-geostrophic system in physical space with general initial data. *Archive for Rational Mechanics and Analysis*, 218(1):527–551, 2015.

[29] M. Feldman and A. Tudorascu. Lagrangian solutions for the semi-geostrophic shallow water system in physical space with general initial data. *St. Petersburg Mathematical Journal*, 27(3):547–568, 2016.

[30] M. Feldman and A. Tudorascu. The semi-geostrophic system: Weak-strong uniqueness under uniform convexity. *Calculus of Variations and Partial Differential Equations*, 56(6):158, 2017.

[31] M. C. L. Filho and H. J. N. Lopes. Existence of a weak solution for the semi-geostrophic equation with integrable initial data. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 132(2):329–339, 2002.

[32] T. O. Gallouët and Q. Mérigot. A Lagrangian scheme à la Brenier for the incompressible Euler equations. *Foundations of Computational Mathematics*, 18(4):835–865, 2018.

[33] T. O. Gallouët, Q. Mérigot, and A. Natale. Convergence of a Lagrangian discretization for barotropic fluids and porous media flow. *SIAM Journal on Mathematical Analysis*, 54(3):2990–3018, 2022.

[34] B. J. Hoskins. The geostrophic momentum approximation and the semi-geostrophic equations. *Journal of the Atmospheric Sciences*, 32(2):233–242, 1975.

[35] J. Kitagawa, Q. Mérigot, and B. Thibert. Convergence of a Newton algorithm for semi-discrete optimal transport. *Journal of the European Mathematical Society*, 21(9):2603–2651, 2019.

[36] T. Lavier. A semi-discrete optimal transport scheme for the 3d incompressible semi-geostrophic equations. *arXiv:2411.00575*, 2024.

[37] T. Lavier. *Optimal Transport Techniques for Atmospheric Dynamics*. PhD thesis, Heriot-Watt University, 2025.

[38] H. Leclerc, Q. Mérigot, F. Santambrogio, and F. Stra. Lagrangian discretization of crowd motion and linear diffusion. *SIAM Journal on Numerical Analysis*, 58(4):2093–2118, 2020.
[39] G. Loeper. A fully nonlinear version of the incompressible Euler equations: The semi-geostrophic system. *SIAM Journal on Mathematical Analysis*, 38(3):795–823, 2006.
[40] Q. Mérigot and B. Thibert. Optimal transport: discretization and algorithms. In A. Bonito and R. H. Nochetto, editors, *Handbook of Numerical Analysis*, volume 22, pages 133–212. Elsevier, 2021.
[41] F. Santambrogio. *Optimal Transport for Applied Mathematicians*. Birkhäuser, 2015.
[42] C. Sarrazin. Lagrangian discretization of variational mean field games. *SIAM Journal on Control and Optimization*, 60(3):1365–1392, 2022.
[43] L. Silini. Local existence of smooth solutions for the semigeostrophic equations on curved domains. *SIAM Journal on Mathematical Analysis*, 55(6):6554–6579, 2023.
[44] C. Villani. *Topics in optimal transportation*. American Mathematical Society, 2003.
[45] C. Villani. *Optimal Transport: Old and NewA vertical slice frontogenesis test case for compressible nonhydrostatic dynamical cores of atmospheric models*. Springer, 2009.
[46] H. Yamazaki and C. J. Cotter. A vertical slice frontogenesis test case for compressible nonhydrostatic dynamical cores of atmospheric models. *arXiv:2501.09752*, 2025.