

LAGRANGIAN SOLUTIONS FOR THE SEMI-GEOSTROPHIC SHALLOW WATER SYSTEM IN PHYSICAL SPACE WITH GENERAL INITIAL DATA

MIKHAIL FELDMAN* AND ADRIAN TUDORASCU†

(Dedicated to Nina Nikolaevna Ural'tseva on her 80th birthday.)

Abstract. In order to accommodate general initial data, an appropriately relaxed notion of renormalized Lagrangian solutions for the Semi-Geostrophic Shallow Water system in physical space is introduced. This is shown to be consistent with previous notions, generalizing them. A weak stability result is obtained first, followed by a general existence result whose proof employs said stability and approximating solutions with regular initial data. The renormalization property ensures the return from physical to dual space and ultimately enables us to achieve the desired results.

1. Introduction. The Semi-Geostrophic equations are a model of large scale atmosphere/ocean flows, where “large-scale” means that the flow is rotation-dominated. Several versions of the Semi-Geostrophic model were studied recently, including equations for 3-dimensional incompressible flow between rigid boundaries (which we refer to as SG below), the Semi-Geostrophic Shallow Water (or SGSW) system for vertically-averaged 3-d incompressible flow with a free surface, and fully compressible flow. Physically interesting solutions for both SG and SGSW must satisfy the convexity principle, introduced by Cullen and Purser [12], which requires that the modified pressure function be convex; all solutions discussed here satisfy this condition. Works by J.-D. Benamou and Y. Brenier [6], and Cullen and Gangbo [10], and Cullen and Maroofi [11] prove that the equations for these models can be solved in “dual variables” (change of variables introduced by B.J. Hoskins [18]). However, the low regularity of solutions in dual variables makes it difficult to express these solutions in the original “physical” variables.

Cullen and Feldman in [9] defined weak Lagrangian solutions in physical variables, replacing the velocity by its corresponding flow map, and proved existence of such solutions for both SG and SGSW. They require some additional conditions/restrictions, which are essentially some form of strict convexity and boundedness of the initial data. Earlier work by L. Ambrosio and W. Gangbo [4] proved existence of solutions for SG in dual space for general convex initial data; however, the extra conditions cited above were still required for SGSW, even for dual space existence. In any case, insofar as the physical space is concerned, an explicit example considered in [14] and [15], shows that Lagrangian solutions in the sense of [9] do not exist for arbitrary convex data.

*Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706 (feldman@math.wic.edu)

†Department of Mathematics, West Virginia University, Morgantown, WV 26506 (adriant@math.wvu.edu)

Recently, in [16], we defined an appropriately weak version of Lagrangian solutions in physical space for SG, replacing the flow map by the corresponding transport plan, and proved existence of such solutions for all convex initial data, i.e. in the most general class of physically admissible solutions. In this paper, we extend the approach of [16] to the SGSW case. Specifically, we introduce and prove existence of a relaxed version of the weak Lagrangian solutions constructed in [9] for the SGSW system in a fixed two-dimensional domain.

Throughout the entire paper, $\Omega \subset \mathbb{R}^2$ is a given open, bounded set with C^1 boundary, and $T \in (0, \infty)$ is fixed. SGSW models the motion of a fluid rapidly rotating around the vertical axis x_3 , contained within the evolving 3-dimensional region $\mathcal{D}(t)$ which has the structure:

$$\mathcal{D}(t) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : (x_1, x_2) \in \Omega, 0 \leq x_3 \leq h(t, x_1, x_2)\},$$

where the region Ω of the (x_1, x_2) -plane is given and fixed, but the height h above the reference level is unknown and can evolve in time. The pressure on the top boundary of the fluid is a given constant p_0 , and

$$p(t, x_1, x_2, x_3) = [h(t, x_1, x_2) - x_3] + p_0.$$

The Shallow Water approximation amounts to the assumption that the horizontal components of the velocity are independent of x_3 . Then, the equations describing the motion of fluid in $\mathcal{D}(t)$ are written as a problem in the 2-dimensional domain Ω , for the unknown height function $h(t, x)$ and the horizontal components of velocity $\mathbf{u}(t, x)$, defined on $[0, T) \times \Omega$, where $\mathbf{u} = (u_1, u_2)$ and $x = (x_1, x_2)$. A version of the 2D Semi-Geostrophic Shallow Water system [10] is

$$\begin{aligned} D_t X &= J(X - x), \\ \partial_t h + \nabla \cdot (h\mathbf{u}) &= 0 \\ X &= \nabla P, \quad P = h + \frac{1}{2}|\text{Id}_\Omega|^2 \quad \text{in } [0, T) \times \Omega; \\ \mathbf{u} \cdot \nu &= 0 \quad \text{on } [0, T) \times \partial\Omega, \\ P(0, \cdot) &= P_0 \quad \text{in } \Omega, \end{aligned} \tag{1.1}$$

where ∇ stands for the spatial gradient, $D_t := \partial_t + \mathbf{u} \cdot \nabla$, and

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

One looks for solutions $(P, \mathbf{u})$ (where P is an artificial quantity known as “modified pressure” [9], [10]) such that $P : [0, T) \times \Omega \rightarrow \mathbb{R}$, $\mathbf{u} : [0, T) \times \Omega \rightarrow \mathbb{R}^2$, which satisfy the Cullen-Purser stability condition (see, e.g., [10]). This amounts to imposing that $P_t(\cdot) := P(t, \cdot)$ be convex for all $t \in [0, T)$. The notation $P_t(x)$ for $P(t, x)$ will be often used, especially when the emphasis is on P as a function of x for some fixed t .

Henceforth, we shall assume that $h_0 \geq 0$ in Ω . Moreover, we assume without loss of generality that

$$\int_{\Omega} h_0(x) dx = 1.$$

Then we have

$$P_0(x) \geq |x|^2/2 \text{ in } \Omega, \text{ and } \|P_0\|_{L^1(\Omega)} = 1 + \frac{1}{2} \int_{\Omega} |x|^2 dx.$$

Denoting $\alpha_t := \nabla P_{t\#} h_t$, where the right-hand side denotes the push-forward of the measure $h_t(x)dx$ by the map $\nabla P_t : \Omega \rightarrow \mathbb{R}^2$, we can formally rewrite (1.1) as the so-called SGSW in dual variables (see [10])

$$\partial_t \alpha + \nabla \cdot (U \alpha) = 0 \quad \text{in } [0, T) \times \mathbb{R}^2; \quad (1.2)$$

$$\nabla P(t, \cdot)_{\#} h(t, \cdot) = \alpha(t, \cdot) \quad \text{for any } t \in [0, T); \quad (1.3)$$

$$P(t, x) = h(t, x) + \frac{1}{2}|x|^2, \text{ with } P(t, \cdot) \text{ convex for all } t \in [0, T); \quad (1.4)$$

$$U(t, X) = J[X - \bar{\gamma}(t, X)], \quad (1.5)$$

$$\alpha(0, X) = \alpha_0(X) \quad \text{for a.e. } X \in \mathbb{R}^2, \quad (1.6)$$

where $\bar{\gamma}(t, X)$ (or $\bar{\gamma}_t(X)$ as alternate notation) denotes the barycentric projection onto α_t of the optimal transfer plan between α_t and h_t . It is a map defined by (see [4], [5])

$$\int_{\mathbb{R}^2} \xi(X) \cdot \bar{\gamma}(X) \alpha(dX) = \iint_{\mathbb{R}^2 \times \Omega} \xi(X) \cdot y \gamma(dX, dy) \quad (1.7)$$

for all continuous $\xi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ of at most quadratic growth, where γ is the (unique) optimal plan between α and h . Since h_t is absolutely continuous with respect to $\mathcal{L}^2$, we deduce

$$\int_{\mathbb{R}^2} \xi(X) \cdot \bar{\gamma}_t(X) \alpha_t(dX) = \int_{\Omega} \xi(\nabla P_t(y)) \cdot y h_t(y) dy. \quad (1.8)$$

If $\alpha_t \ll \mathcal{L}^2$ for a.e. $t \in (0, T)$, the dual velocity field $U(t, X) = J[X - \nabla P^*(t, X)]$ may be utilized instead of the one from the above display, where $P^*(t, \cdot)$ is the Legendre transform of $P(t, \cdot)$ over Ω , defined by

$$P^*(t, X) = \sup_{x \in \Omega} \{x \cdot X - P(t, x)\} \quad \text{for } (t, X) \in [0, T) \times \mathbb{R}^2.$$

A detailed explanation of why ∇P^* should be replaced by $\bar{\gamma}$ for the Semi-Geostrophic system in the general case can be found in the introduction of [15]. A similar argument can be used for SGSW.

The dual space solutions obtained in [4, 10] are not known to be regular enough to be translated into Eulerian solutions of the problem in physical

space. For both SG and SGSW, Cullen and Feldman [9] showed existence of physical space Lagrangian solutions in the case $\alpha_0 \in L^p(E)$ for some $p > 1$ ($E = \mathbb{R}^3$ for SG, while $E = \mathbb{R}^2$ for SGSW) by using the regular flow map of the dual space velocity, which exists by results of Ambrosio [1], to pull the dual space solutions back to physical space. Faria et al. [14] extended that to the case $p \geq 1$. An example with α_0 being a point mass, considered in [14] and [15], shows that the flow map in physical space is not stable to small perturbations in the case of singular α_0 (as opposed to the L^1 -case). However, as we show in [16], this can be resolved by replacing the flow map with the corresponding transport plan in the definition of weak Lagrangian solutions, and requiring that a renormalization property be satisfied (this property is satisfied in the regular case for solutions in [9]). In [16] we obtained weak stability, and existence for the renormalized relaxed Lagrangian solutions for SG physical space for any convex initial data.

Remark 1.1. *Recently, existence of Eulerian solutions for SG, for a class of initial data, was obtained by L. Ambrosio, M. Colombo, G. De Philippis, A. Figalli [2, 3]. Conditions include the requirement that the support of α_0 in the dual space is the whole space (where $\alpha_0 = \nabla P_{0\#} \mathcal{L}^3|_{\Omega}$ for SG case). No such result is available for SGSW. Furthermore, existence of Eulerian solutions in physical space for more general initial data, when the support of α_0 in the dual space may have a non-empty boundary, is presently not known for either SG or SGSW.*

In the present paper we extend the approach in [16] to the SGSW system, while overcoming some technical difficulties particular to the SGSW case. We propose a notion of solution that generalizes the weak Lagrangian solution in physical space introduced in [9]. Our approach involves a relaxation of the nonlinear interaction between the physical space maps ∇P_t and F_t (the Lagrangian path of the particle originating at $x \in \Omega$ at time $t = 0$; see [9]) by replacing the maps $F_t : \Omega \rightarrow \Omega$ such that $F_{\#} h_0 = h_t$ (streamlines in physical space, or, put differently, the flow of the physical velocity $\mathbf{u}$; see [9] or Definition 3.1 below) by the corresponding transport plans, which are measures on $\Omega \times \Omega$: it is, as discussed earlier, perhaps not so surprising that this approach is successful, given the connection between SG/SGSW and Optimal Transport.

In Section 2 we establish a connection between (1.3)–(1.4) and a variational principle first used by Cullen & Gangbo in [10]. The equivalence between (1.3)–(1.4) and said principle is essential to the stability of the objects involved in the subsequent analysis. We overcome some technical difficulties created by the fact that the dual space measures may be singular in our case; this is why the results in [4, 10] do not apply directly.

In Section 3 we introduce, using the intuition gained from Cullen & Feldman's weak Lagrangian solutions in [9], the notion of relaxed renormalized Lagrangian solutions in physical space; then show that, under additional regularity, a relaxed solution determines a classical solution for the SGSW system

(1.1), and that the new notion of solutions is compatible/consistent with the previous notion of weak Lagrangian solutions in [9]. We also prove in this section that a relaxed Lagrangian solution in physical space gives rise to a dual-space solution, which is used in the following sections to prove stability and existence of solutions in physical space, and to show their time-continuity in an appropriate sense.

Section 4 then moves to the dual-space problem (1.2)–(1.6), its main result being a weak stability result for dual-space solutions. This will be useful in Section 5, where the stability result in dual space will apply to obtain the corresponding result for relaxed Lagrangian solutions in physical space. We also note in the same section that, in light of the earlier existence result by Ambrosio & Gangbo [4] in case of regular initial data, we obtain (by employing said stability) a general existence result in dual space by approximation.

Finally, in Section 5 we prove first a weak stability, then an existence result for relaxed Lagrangian solutions in physical space. In order to get existence, the stability result applies after we approximate the initial data by appropriately regular measures and construct our approximating relaxed Lagrangian solutions on the blueprint of the weak Lagrangian solutions by Cullen & Feldman [9] (see Definition 3.1). Since these solutions give rise to weak distributional solutions of the problem in dual space, the weak stability result in dual space from Section 4 can be used here. We conclude Section 5 (and the paper) by proving some weak time-continuity of our relaxed Lagrangian solutions.

2. A variational principle for the physical-space to dual-space passage. In this section we record some existence and stability properties for pairs $(\mu, h) \in \mathcal{P}_2(\mathbb{R}^2) \times \mathcal{P}^{ac}(\Omega)$ such that $P = h + |\text{Id}_\Omega|^2/2$ is the restriction of a convex function to Ω and $\nabla P_\# h = \mu$. Recall here that $\mathcal{P}_2(\mathbb{R}^d)$ denotes the metric space [19] of Borel probabilities on $\mathbb{R}^d$ with finite second moments; it is endowed with the quadratic Wasserstein metric W_2 (see [19]). We also use $\mathcal{P}^{ac}(\Omega)$ to denote the set of Borel probabilities on Ω which are absolutely continuous with respect to the Lebesgue measure.

Proposition 2.1. *For any $\mu \in \mathcal{P}_2(\mathbb{R}^2)$ define $J[\mu] : \mathcal{P}(\Omega) \cap L^2(\Omega) \rightarrow \mathbb{R}$ by*

$$J[\mu](h) := W_2^2(\mu, h) + \|h\|_{L^2(\Omega)}^2.$$

Then:

- (i) $J[\mu]$ has a unique minimizer over $\mathcal{P}(\Omega) \cap L^2(\Omega)$. Let us denote this minimizer by $\mathcal{H}[\mu]$.
- (ii) If $\{\mu_n\}_n \subset \mathcal{P}_2(\mathbb{R}^2)$ converges to μ in the $\mathcal{P}_2(\mathbb{R}^2)$ topology, then $\{\mathcal{H}[\mu_n]\}_n$ converges to $\mathcal{H}[\mu]$ strongly in $L^2(\Omega)$.
- (iii) Assume Ω is connected, and $\mu \ll \mathcal{L}^2$ with $d\mu/d\mathcal{L}^2$ compactly supported. Then there exists a convex function P taking only finite values on an open ball

containing $\bar{\Omega}$ such that

$$P = \frac{1}{2}|\text{Id}_\Omega|^2 + \mathcal{H}[\mu] \text{ in } \Omega, \text{ and } \nabla P_\# \mathcal{H}[\mu] = \mu. \quad (2.1)$$

(iv) Suppose $\{\mu_n\}_n$ is a sequence as in (ii), and such that μ_n satisfies the assumptions from (iii) for all integers $n \geq 1$. Let $\{P_n\}_n$ be a sequence of convex potentials, each P_n corresponding to μ_n as in (iii). Then $\{P_n\}_n$ converges in $L^2(\Omega)$ to some P that satisfies (2.1) for the limiting μ .

Proof. (i) Note that $\mathcal{P}(\Omega) \cap L^2(\Omega)$ is closed with respect to the $L^2_{\text{weak}}(\Omega)$ -topology. So, if we take a minimizing sequence (as $J[\mu]$ clearly has a nonnegative infimum) $\{h_n\}_n$, this will converge (possibly, up to a subsequence) weakly in $L^2(\Omega)$ (and in the Wasserstein metric, since all continuous functions on Ω of at most quadratic growth are automatically in $L^2(\Omega)$; indeed, it is well-known that convergence with respect to the Wasserstein distance W_2 is equivalent to convergence as measures against test functions of at most quadratic growth [19]) to some $\tilde{h} \in \mathcal{P}(\Omega) \cap L^2(\Omega)$. Thus,

$$\liminf_{n \rightarrow \infty} J[\mu](h_n) \geq J[\mu](\tilde{h}),$$

i.e. $\tilde{h}$ is a minimizer. Uniqueness follows from the convexity of $h \mapsto W_2^2(\mu, h)$ and the strict convexity of $h \mapsto \|h\|_{L^2(\Omega)}^2$ with respect to the linear vector space structure of $L^2(\Omega)$.

(ii) We prove that $\mu_n \rightarrow \mu$ in $\mathcal{P}_2(\mathbb{R}^2)$ implies

$$J[\mu_n] \xrightarrow[n \rightarrow \infty]{\Gamma} J[\mu],$$

where the Γ -convergence holds with respect to the weak topology of $L^2(\Omega)$. Indeed, if $h_n \rightharpoonup h$ weakly in $L^2(\Omega)$, then h_n converges narrowly (i.e. with respect to continuous and bounded test functions) to h which, since Ω is bounded, is equivalent to $W_2(h_n, h) \rightarrow 0$. So,

$$\liminf_{n \rightarrow \infty} J[\mu_n](h_n) = W_2^2(\mu, h) + \liminf_{n \rightarrow \infty} \|h_n\|_{L^2(\Omega)}^2 \geq J[\mu](h).$$

Obviously, $J[\mu_n](h)$ converges to $J[\mu](h)$ for any $h \in \mathcal{P}(\Omega) \cap L^2(\Omega)$. Thus, the Γ -convergence claim above is proved. Therefore, the sequence $\{\mathcal{H}[\mu_n]\}_n$ consisting of the minimizers of $J[\mu_n]$ converges weakly in $L^2(\Omega)$ (and, consequently, in the metric W_2 as well) to the minimizer $\mathcal{H}[\mu]$ of $J[\mu]$. Since, by minimality, we have

$$W_2^2(\mu_n, h[\mu]) + \|\mathcal{H}[\mu]\|_{L^2(\Omega)}^2 \geq J[\mu_n](\mathcal{H}[\mu_n]),$$

we use $W_2(\mu_n, \mu) + W_2(\mathcal{H}[\mu_n], \mathcal{H}[\mu]) \rightarrow 0$ to infer that

$$\|\mathcal{H}[\mu]\|_{L^2(\Omega)} \geq \limsup_{n \rightarrow \infty} \|\mathcal{H}[\mu_n]\|_{L^2(\Omega)},$$

i.e. $\|\mathcal{H}[\mu_n] - \mathcal{H}[\mu]\|_{L^2(\Omega)} \rightarrow 0$ as $n \rightarrow \infty$.

(iii) This comes directly from Proposition 3.4 [10].

(iv) From (ii) and (iii) we obtain a convex function P defined on the convex envelope of Ω such that $\|P_n - P\|_{L^2(\Omega)} \rightarrow 0$ and $\mathcal{H}[\mu] = P - |\text{Id}_\Omega|^2/2$ in Ω . Thus, $\nabla P_n \rightarrow \nabla P$ pointwise $\mathcal{L}^2$ -a.e. in Ω (see, e.g., [13]). Use $\|\mathcal{H}[\mu_n] - \mathcal{H}[\mu]\|_{L^2(\Omega)} \rightarrow 0$ and Dominated Convergence to get

$$\int_{\Omega} \xi(\nabla P_n(x)) \mathcal{H}[\mu_n](x) dx \xrightarrow{n \rightarrow \infty} \int_{\Omega} \xi(\nabla P(x)) \mathcal{H}[\mu](x) dx$$

for all $\xi \in C_b(\mathbb{R}^2)$, i.e. $\nabla P_{n\#} \mathcal{H}[\mu_n]$ converges narrowly to $\nabla P_{\#} \mathcal{H}[\mu]$. We finish the proof by recalling that $\nabla P_{n\#} \mathcal{H}[\mu_n] = \mu_n$ and $W_2(\mu_n, \mu) \rightarrow 0$. $\square$

As an immediate consequence, we obtain the following result, which generalizes its counterpart in [10] to the case of a general probability measure with finite second moment.

Corollary 2.2. *For any $\mu \in \mathcal{P}_2(\mathbb{R}^2)$ there exists a convex $P : \mathbb{R}^2 \rightarrow \mathbb{R} \cup \{\infty\}$ such that $P \in L^2(\Omega)$, $\nabla P_{\#} \mathcal{H}[\mu] = \mu$, and $\mathcal{H}[\mu] = P - |\text{Id}_\Omega|^2/2$ in Ω .*

Proof. First, we can use Lemma 7.1.10 in [5] to approximate μ with respect to the quadratic Wasserstein distance by smooth, positive probability densities lying in $\mathcal{P}_2(\mathbb{R}^2)$ (just take an everywhere positive convolution kernel in the proof of said lemma). For all integers $k \geq 1$ consider $\xi_k \in C_c^\infty(B(0, k+1))$ such that $0 \leq \xi_k \leq 1$ in $\mathbb{R}^2$ and $\xi_k \equiv 1$ in $B(0, k)$. By Dominated Convergence,

$$\lambda_k := \int_{\mathbb{R}^2} \xi_k(x) \rho(x) dx \rightarrow \int_{\mathbb{R}^2} \rho(x) dx = 1 \text{ as } k \rightarrow \infty.$$

If we place $\rho_k := \xi_k \rho / \lambda_k$, we use $\lambda_k \rightarrow 1$ and Dominated Convergence again to infer that

$$\int_{\mathbb{R}^2} \rho_k(x) \zeta(x) dx \rightarrow \int_{\mathbb{R}^2} \rho(x) \zeta(x) dx \text{ as } k \rightarrow \infty$$

for all $\zeta \in C(\mathbb{R}^2)$ of at most quadratic growth. Thus, ρ_k is a sequence of smooth, compactly supported probability densities converging to ρ in the quadratic Wasserstein distance (once again, we use that fact that narrow convergence and convergence of the p -moments is equivalent to convergence in the p -Wasserstein distance; see, e.g., [19]). In conclusion, we can approximate μ in the metric space $(\mathcal{P}_2(\mathbb{R}^2), W_2)$ by smooth, compactly supported densities. We conclude by Proposition 2.1 (iv). $\square$

Since our initial data will be satisfying

$$\begin{aligned} P_0 : \mathbb{R}^2 \rightarrow \mathbb{R} \cup \{\infty\} \text{ is convex, such that } P_0|_{\Omega} \in L^2(\Omega), \\ P_0|_{\Omega} - \frac{1}{2} |\text{Id}_\Omega|^2 =: h_0 \in \mathcal{P}(\Omega), \text{ and } \nabla P_0|_{\Omega} \in L^2(h_0; \mathbb{R}^2), \end{aligned} \quad (2.2)$$

we would like to set $\alpha_0 := \nabla P_{0\#} h_0$, then approximate α_0 by $\{\alpha_0^n\}_n$ as in the proof of Corollary 2.2, then get the corresponding $L^2(\Omega)$ approximations

$P_0^n = \mathcal{H}[\alpha_0^n] + |\text{Id}_\Omega|^2/2$ of P_0 by means of Proposition 2.1 (iv). However, in order to apply Proposition 2.1 (iv) we need to know that $h_0 = \mathcal{H}[\alpha_0]$, i.e. h_0 minimizes $J[\alpha_0]$ over $\mathcal{P}(\Omega) \cap L^2(\Omega)$. It is not guaranteed *a priori* that (2.2) implies $h_0 = \mathcal{H}[\alpha_0]$ for $\alpha_0 := \nabla P_{0\#} h_0$. This is the converse to Proposition 2.1 (iii), and it is proved in [10]. However, the proof in [10] relies heavily on the assumption that $\alpha_0 \ll \mathcal{L}^2$ (as it starts with the assumption that for any $\rho \in \mathcal{P}^{ac}(\Omega)$ there exists a Borel map pushing α_0 forward to ρ ; then it uses the fact that ∇P_0^* pushes α_0 forward to h_0), so we need to adapt the proof to the general setting. From now on we will drop the notation $P|_\Omega$ for the restriction to Ω of a function P defined on $\mathbb{R}^2$; when we, for example, write $P \in L^2(\Omega)$, it is clear we mean that.

Proposition 2.3. *If P satisfies (2.2) (for $P_0 = P$), then $h = \mathcal{H}[\nabla P_{\#} h]$.*

Proof. Pick an arbitrary $\rho \in \mathcal{P}^{ac}(\Omega)$ and let $\eta \in \mathcal{P}(\mathbb{R}^2 \times \Omega)$ be an admissible transport plan between $\alpha := \nabla P_{\#} h$ and ρ (i.e. a Borel probability with marginals α and ρ , respectively). Let $\gamma := (\nabla P \times \text{Id}_\Omega)_{\#} h$ be the unique (since $h \ll \mathcal{L}^2$) optimal transport plan between α and h , and consider its barycentric projection [4], [5] onto α , i.e. the α -measurable Borel map $\bar{\gamma}$ such that

$$\begin{aligned} \int_{\mathbb{R}^2} \xi(X) \cdot \bar{\gamma}(X) \alpha(dX) &= \iint_{\mathbb{R}^2 \times \Omega} \xi(X) \cdot y \gamma(dX, dy) \\ &= \int_{\Omega} \xi(\nabla P(y)) \cdot y h(y) dy. \end{aligned} \quad (2.3)$$

Since $\bar{\gamma}(X) \in \partial P^*(X)$ (here ∂ denotes the subgradient) for α -a.e. $X \in \mathbb{R}^2$ (see [15]), we have $X \in \partial P(\bar{\gamma}(X))$ for α -a.e. $X \in \mathbb{R}^2$, which implies (by the convexity of P)

$$\int_{\mathbb{R}^2} \int_{\Omega} [P(y) - P(\bar{\gamma}(X))] \eta(dX, dy) \geq \int_{\mathbb{R}^2} \int_{\Omega} X \cdot [y - \bar{\gamma}(X)] \eta(dX, dy).$$

Using the marginals, we rewrite the above to get

$$\begin{aligned} \int_{\Omega} P(y) \rho(y) dy - \int_{\mathbb{R}^2} P(\bar{\gamma}(X)) \alpha(dX) &\geq \\ \int_{\mathbb{R}^2} \int_{\Omega} X \cdot y \eta(dX, dy) - \int_{\mathbb{R}^2} X \cdot \bar{\gamma}(X) \alpha(dX), \end{aligned}$$

which, by (2.3) and $\nabla P_{\#} h = \alpha$, means

$$\begin{aligned} \int_{\Omega} P(y) \rho(y) dy - \int_{\Omega} P(\bar{\gamma}(\nabla P(y))) h(y) dy &\geq \\ \int_{\mathbb{R}^2} \int_{\Omega} X \cdot y \eta(dX, dy) - \int_{\Omega} \nabla P(y) \cdot y h(y) dy. \end{aligned}$$

Next we multiply the above inequality by 2 and add the second moments of h

and α to both sides to obtain

$$2 \int_{\Omega} P(y) \rho(y) dy - 2 \int_{\Omega} P(\bar{\gamma}(\nabla P(y))) h(y) dy + \int_{\Omega} [|y|^2 + |\nabla P(y)|^2] h(y) dy \geq \\ 2 \int_{\mathbb{R}^2} \int_{\Omega} X \cdot y \eta(dX, dy) + W_2^2(\alpha, h),$$

then we subtract the second moments of α and ρ to get, using the identity $2P - |\text{Id}_{\Omega}|^2 = 2h$ in Ω ,

$$2 \int_{\Omega} h(y) \rho(y) dy - 2 \int_{\Omega} P(\bar{\gamma}(\nabla P(y))) h(y) dy + \int_{\Omega} |y|^2 h(y) dy \geq \\ - \int_{\mathbb{R}^2} \int_{\Omega} |X - y|^2 \eta(dX, dy) + W_2^2(\alpha, h).$$

We use the relation between P and h again to replace $|y|^2$ by $2P(y) - 2h(y)$, then add $\|h\|_{L^2(\Omega)}^2 - \|\rho\|_{L^2(\Omega)}^2$ to both sides to arrive to

$$- \int_{\Omega} [h(y) - \rho(y)]^2 dy + 2 \int_{\Omega} P(y) h(y) dy - 2 \int_{\Omega} P(\bar{\gamma}(\nabla P(y))) h(y) dy \geq \\ - \int_{\mathbb{R}^2} \int_{\Omega} |X - y|^2 \eta(dX, dy) - \|\rho\|_{L^2(\Omega)}^2 + J[\alpha](h). \quad (2.4)$$

We use the convexity of P again to see that

$$\int_{\Omega} [P(y) - P(\bar{\gamma}(\nabla P(y)))] h(y) dy \leq \int_{\Omega} \nabla P(y) \cdot [y - \bar{\gamma}(\nabla P(y))] h(y) dy.$$

But (2.3) and $\nabla P_{\#} h = \alpha$ reveal that the right hand side of the above display is zero. Thus, (2.4) implies

$$\int_{\mathbb{R}^2} \int_{\Omega} |X - y|^2 \eta(dX, dy) + \|\rho\|_{L^2(\Omega)}^2 \geq J[\alpha](h).$$

By taking the infimum over all admissible transport plans η in the left hand side, we get the desired conclusion. $\square$

Remark 2.4. From Proposition 2.1 and Proposition 2.3 we know that we can approximate the initial convex $P_0 = h_0 + |\text{Id}_{\Omega}|^2/2$ (such that h_0 is a probability density in Ω) in $L^2(\Omega)$ in terms of the unique (whereas a priori only the gradient is unique, the constant of integration is fixed too due to the mass constraint on h_0^n) convex P_0^n such that $\nabla P_{0\#}^n \mathcal{H}[\alpha_0^n] = \alpha_0^n$, where $\alpha_0^n \ll \mathcal{L}^2$ has a smooth, compactly supported density and approximates $\alpha_0 := \nabla P_{0\#} h_0$ in $\mathcal{P}_2(\mathbb{R}^2)$.

Next we prove that the minimizers are stable under narrow convergence as well. This result is crucial to proving stability of dual space solutions.

Proposition 2.5. Let $\{\mu_n\}_n \cup \{\mu\} \subset \mathcal{P}_2(\mathbb{R}^2)$ such that μ_n converges to μ narrowly and $\sup_n W_2(\mu_n, \mu) < \infty$. Then $\mathcal{H}[\mu_n]$ converges weakly in $L^2(\Omega)$

(consequently, $W_2(\mathcal{H}[\mu_n], \mathcal{H}[\mu]) \rightarrow 0$) and everywhere locally uniformly in Ω to $\mathcal{H}[\mu]$. Furthermore, if $P_n := \mathcal{H}[\mu_n] + |\text{Id}_\Omega|^2/2$, then for any $\omega \Subset \Omega$ there exists a real constant C depending only on ω and $\sup_n \|P_n\|_{L^2(\Omega)}$ such that $\|\nabla P_n\|_{L^\infty(\omega; \mathbb{R}^2)} \leq C$. Also, $\nabla P_n \rightarrow \nabla P$ a.e. in Ω for all $t \in [0, T]$.

Proof. Let $\chi := \mathcal{L}^2|_\Omega$, which, clearly, belongs to $\mathcal{P}^{ac}(\Omega) \subset \mathcal{P}_2(\mathbb{R}^2)$. Then let us unburden notation by letting $h_n := \mathcal{H}[\mu_n]$. By the minimizing property and the triangle inequality, we get

$$\begin{aligned} W_2^2(\mu_n, h_n) + \|h_n\|_{L^2(\Omega)}^2 &\leq W_2^2(\mu_n, \chi) + \mathcal{L}^2(\Omega) \\ &\leq 2[W_2^2(\mu_n, \mu) + W_2^2(\mu, \chi)] + \mathcal{L}^2(\Omega), \end{aligned}$$

which shows, according to the hypothesis, that the sequence $\{\|h_n\|_{L^2(\Omega)}\}_n$ is bounded. Let h be the weak $L^2(\Omega)$ limit of one of its subsequences (not relabeled). Our plan is to show that $h = \mathcal{H}[\mu]$. (Then, by the uniqueness of the limit, the conclusion will be that the entire sequence $\{\|h_n\|_{L^2(\Omega)}\}_n$ converges weakly in $L^2(\Omega)$ to h .) Let $P_n := h_n + |\text{Id}_\Omega|^2/2$ in Ω ; so, P_n converges weakly in $L^2(\Omega)$ to $P := h + |\text{Id}_\Omega|^2/2$. Since P_n is nonnegative and convex, we have that so is P_n^2 . From Jenssen's inequality applied to P_n^2 over any ball $B(x, r/2) \subset \Omega$ (with $r > 0$ sufficiently small) we derive

$$P_n^2(x) \leq \frac{4}{\pi r^2} \|P_n\|_{L^2(\Omega)}^2,$$

which means that P_n is uniformly bounded on the closure of $\Omega_r := \{x \in \Omega : \text{dist}(x, \partial\Omega) > r\}$. Convexity also gives (for all n and a.e. $x \in \bar{\Omega}_{2r}$)

$$\frac{2}{r} [P_n(x \pm r\mathbf{e}_i/2) - P_n(x)] \geq \pm \nabla P_n(x) \cdot \mathbf{e}_i$$

for $i = 1, 2$. Thus, $\{P_n\}_n$ is bounded in $W^{1,\infty}(\Omega_{2r})$ for all sufficiently small $r > 0$, i.e. $\{P_n\}_n$ is bounded in $W_{loc}^{1,\infty}(\Omega)$ (with bounds on the gradients as specified in the statement of the proposition). Thus, up to a subsequence, $\{P_n\}_n$ converges locally uniformly to (necessarily) its weak $L^2(\Omega)$ -limit P . By the uniqueness of the limit, the entire sequence $\{P_n\}_n$ does so. By convexity again, we also get that ∇P_n converges to ∇P a.e. in Ω (see [13]). By Dominated Convergence, this implies that $\zeta \circ \nabla P_n$ converges strongly in $L^p(\Omega)$ to $\zeta \circ \nabla P$ for all $\zeta \in C_b(\mathbb{R}^2)$ and all $1 \leq p < \infty$. Since $h_n \rightharpoonup h$ weakly in $L^2(\Omega)$, we get

$$\int_\Omega \zeta \circ \nabla P_n(x) h_n(x) dx \xrightarrow{n \rightarrow \infty} \int_\Omega \zeta \circ \nabla P(x) h(x) dx$$

for all $\zeta \in C_b(\mathbb{R}^2)$, i.e. $\mu_n = \nabla P_n \# h_n$ converges narrowly to $\nabla P \# h$. Therefore, $\mu = \nabla P \# h$, and since $P = h + |\text{Id}_\Omega|^2/2$, Proposition 2.3 yields $h = \mathcal{H}[\mu]$. $\square$

3. Relaxed renormalized Lagrangian solutions in physical space.

The starting point will be the weak Lagrangian solutions in the physical space. Such solutions were introduced by Cullen & Feldman [9], and existence of

Lagrangian solutions for (1.1) was shown in [9] for any compactly supported $\alpha_0 = \nabla P_{0\#} h_0 \in L^q(\mathbb{R}^2)$ for some $q > 1$, on the basis of Ambrosio's theory on transport equations and flows of BV vector fields [1]. These results were extended to the case $q = 1$ in [14]. The restriction $\alpha_0 \ll \mathcal{L}^2$ amounts to requiring some *strict convexity* property of the potentials P_t , which is certainly more than what Cullen and Purser's stability condition imposes (which is simply the convexity of P_t). The definition of weak Lagrangian solutions in the physical space in [9] is the following:

Definition 3.1. *Let $P_0 \in L^\infty(\Omega)$ be convex such that $h_0 := P_0 - |\text{Id}_\Omega|^2/2 \in \mathcal{P}(\Omega)$, and let $p \in [1, \infty)$. Let $P : [0, T) \times \Omega \rightarrow \mathbb{R}$ such that*

$$P_t := P(t, \cdot) \text{ is convex for all } t \in [0, T), \quad P_t - |\text{Id}_\Omega|^2/2 =: h_t \in \mathcal{P}(\Omega)$$

and

$$P \in L^\infty([0, T); W^{1, \infty}(\Omega)) \cap C([0, T); W^{1, p}(\Omega)).$$

Let $F : [0, T) \times \Omega \rightarrow \Omega$ be a Borel map such that $F \in C([0, T); L^p(h_0; \mathbb{R}^2))$.

Then the pair (P, F) is called a weak Lagrangian solution of (1.1) in $[0, T) \times \Omega$ if the following hold:

- i. $F(0, x) = x$, $P(0, x) = P_0(x)$ for h_0 -a.e. $x \in \Omega$,
- ii. for any $t > 0$ the mapping $F_t = F(t, \cdot) : \Omega \rightarrow \Omega$ satisfies $F_{t\#} h_0 = h_t$;
- iii. There exists a Borel map $F^* : [0, T) \times \Omega \rightarrow \Omega$ such that for every $t \in (0, T)$ the map $F_t^* = F^*(t, \cdot) : \Omega \rightarrow \Omega$ satisfies: $F_{t\#}^* h_t = h_0$, and satisfies $F_t^* \circ F_t \equiv \text{Id}_\Omega$ h_0 -a.e. and $F_t \circ F_t^* \equiv \text{Id}_\Omega$ h_t -a.e. in Ω ;
- iv. The function

$$Z(t, x) = \nabla P(t, F(t, x)) \tag{3.1}$$

is a distributional solution of

$$\begin{aligned} \partial_t Z(t, x) &= J[Z(t, x) - F(t, x)] \text{ for all } t \in [0, T) \text{ and } h_0\text{-a.e. } x \in \Omega, \\ Z(0, x) &= \nabla P_0(x) \text{ for } h_0\text{-a.e. } x \in \Omega. \end{aligned} \tag{3.2}$$

Note that the sense in which (3.2) must be satisfied is

$$\int_0^T \int_\Omega [Z \cdot \partial_t \varphi + J(Z - F) \cdot \varphi] h_0 dx dt + \int_\Omega \nabla P_0(x) \cdot \varphi(0, x) h_0(x) dx = 0 \tag{3.3}$$

for any $\varphi \in C_c^1([0, T) \times \Omega; \mathbb{R}^2)$. This implies (as observed in [9] in a similar case) $t \mapsto Z(t, x) \in H^1(0, T; \mathbb{R}^2)$ for $\mathcal{L}^2$ -a.e. $x \in \Omega$, and the differential equation in (3.2) is satisfied in the $\mathcal{L}^2$ -a.e. in $(0, T) \times \Omega$. Thus, for $\xi \in C^1(\mathbb{R}^2) \cap \text{Lip}(\mathbb{R}^2)$ the map $t \mapsto \xi(Z(t, x))$ is absolutely continuous for $\mathcal{L}^2$ -a.e. $x \in \Omega$ and

$$\frac{d}{dt} \xi(Z(t, x)) h_0(x) = \nabla \xi(Z(t, x)) \cdot J[Z(t, x) - F(t, x)] h_0(x) \tag{3.4}$$

for a.e. $t \in [0, T]$. Consequently, a more general, “renormalized” version of (3.3) is available in the form

$$\begin{aligned} \int_0^T \int_{\Omega} \left\{ \xi(Z(t, x)) \partial_t \zeta(t, x) + \nabla \xi(Z(t, x)) \cdot J[F(t, x) \right. \\ \left. - Z(t, x)] \zeta(t, x) \right\} h_0(x) dx dt + \int_{\Omega} \xi(\nabla P_0(x)) \zeta(0, x) h_0(x) dx = 0, \end{aligned} \quad (3.5)$$

for any $\xi \in C^1(\mathbb{R}^2) \cap \text{Lip}(\mathbb{R}^2)$, $\zeta \in C_c^1([0, T] \times \Omega)$.

Remark 3.2. Note that the condition $P \in L^\infty([0, T]; W^{1,\infty}(\Omega))$ from Definition 3.1 makes the choice of test functions $\xi \in C_c^1(\mathbb{R}^2)$ equivalent to the choice $\xi \in C^1(\mathbb{R}^2) \cap \text{Lip}(\mathbb{R}^2)$ in (3.4), i.e. under the assumption $P \in L^\infty([0, T]; W^{1,\infty}(\Omega))$, (3.4) is equivalent to the differential equation in (3.2) regardless of which of the two classes we use.

Remark 3.3. The use of arbitrary $\xi \in C^1(\mathbb{R}^2) \cap \text{Lip}(\mathbb{R}^2)$ instead of the obvious first choices $\xi_k(x) = x_k$ for $k = 1, 2$ is enabled by the fact that, as discussed above, equation (3.2) is satisfied not only in the sense of distributions, but pointwise a.e. as well. This amounts to solutions defined through this general choice of test functions having some sort of “renormalization” property. A consequence of this property is that it allows to prove that the relaxed solutions, defined below, give rise to distributional solutions in dual space. Using $C_c^1(\mathbb{R}^2)$ test functions instead will produce the same result (see Theorem 3.7). The only place where a significant difference (as a result of using one class over the other) is the proof of the physical space weak stability Theorem 5.1, which goes through only if we choose $C_c^1(\mathbb{R}^2)$. This is a choice we make (in spite of weakening the notion of solution even further) in order to accommodate the most general initial data for our main existence result.

Our goal is to define relaxed Lagrangian solutions in the physical space, replacing maps $F_t : \Omega \rightarrow \Omega$ by the corresponding transport plans, which are measures σ_t on $\Omega \times \Omega$. Thus we base the definition of relaxed Lagrangian solution on the pair (P, σ) , and try to capture all the features of (P, F) from Definition 3.1 (for which $\sigma_t = (\text{Id}_\Omega \times F(t, \cdot))_\# h_0$).

More precisely, assuming existence of Lagrangian solution (P, F) in the sense of Definition 3.1, we define the measure σ on $(0, T) \times \Omega \times \Omega$ by

$$\int_0^T \int_{\Omega} \int_{\Omega} \xi(t, x, y) \sigma(dt, dx, dy) = \int_0^T \int_{\Omega} \xi(t, x, F(t, x)) h_0(x) dx dt, \quad (3.6)$$

for all $\xi \in C_b((0, T) \times \Omega \times \Omega)$. We notice first that the property $F(t, \cdot)_\# h_0 = h_t$ shows that σ disintegrates as

$$\int_0^T \int_{\Omega} \int_{\Omega} \xi(t, x, y) \sigma(dt, dx, dy) = \int_0^T \left(\int_{\Omega} \int_{\Omega} \xi(t, x, y) \sigma_t(dx, dy) \right) dt \quad (3.7)$$

for all $\xi \in C_b((0, T) \times \Omega \times \Omega)$, where $[0, T] \ni t \mapsto \sigma_t = (\text{Id}_\Omega \times F_t)_\# h_0$ is a Borel family of measures on $\Omega \times \Omega$ such that

$$\text{proj}_x \sigma_t = h_0, \text{ and } \text{proj}_y \sigma_t = h_t \quad \text{for all } t \in [0, T]. \quad (3.8)$$

If σ is a measure on the product space $X \times Y$, by $\mu := \text{proj}_x \sigma$ (the x -marginal of σ) we denote the measure on X given by $\mu(A) = \sigma(A \times Y)$ for all Borel sets $A \in X$. The second condition from the above display is equivalent, in view of (3.7), to $\text{proj}_{(t,y)} \sigma = h$ (as a function of both t and y). Furthermore, (3.5) becomes, by (3.1),

$$\begin{aligned} & \int_0^T \iint_{\Omega \times \Omega} \left\{ \xi(\nabla P_t(y)) \partial_t \zeta(t, x) + \nabla \xi(\nabla P_t(y)) \cdot J[y \right. \\ & \quad \left. - \nabla P_t(y)] \zeta(t, x) \right\} \sigma(dt, dx, dy) + \int_{\Omega} \xi(\nabla P_0(x)) \zeta(0, x) h_0(x) dx = 0 \end{aligned} \quad (3.9)$$

for all $\xi \in C_c^1(\mathbb{R}^2)$ and all $\zeta \in C_c^1([0, T] \times \Omega)$. Let us, from now on, use the notation $\int_{\Omega^2}$ for $\int_{\Omega} \int_{\Omega}$.

We note that if P is as in Definition 3.1, then $P_t - |\text{Id}_{\Omega}|^2/2 =: h_t \in \mathcal{P}(\Omega)$; thus $h_t \in L^1(\Omega)$. Also $\nabla P_t \in L^\infty(\Omega)$ with ess sup norms uniformly bounded with respect to $t \in [0, T]$, which implies that $\nabla P_t \in L^2(h_t; \mathbb{R}^2)$ for a.e. $t \in [0, T]$ with $\|\nabla P_t\|_{L^2(h_t; \mathbb{R}^2)}$ bounded uniformly with respect to $t \in [0, T]$. In particular, we get that P and h from Definition 3.1 satisfy $\nabla P \in L^2(h; \mathbb{R}^2)$ (here, both P and h are regarded as functions of time-space on $[0, T] \times \Omega$; also, the ∇ notation always means the spatial gradient).

Lemma 3.4. *Let $P : [0, T] \times \mathbb{R}^2 \rightarrow \mathbb{R} \cup \{\infty\}$ be a Borel function such that $P(t, \cdot)$ is convex and $P(t, \cdot) - |\text{Id}_{\Omega}|^2/2 =: h(t, \cdot) \in \mathcal{P}(\Omega)$ for all $t \in [0, T]$. If $\nabla P \in L^2(h; \mathbb{R}^2)$ and σ is a nonnegative, finite Borel measure on $(0, T) \times \Omega \times \Omega$ satisfying (3.7) and (3.8), then the left-hand-side of the equation (3.9) is well-defined for all $\xi \in C^1(\mathbb{R}^2) \cap \text{Lip}(\mathbb{R}^2)$.*

Proof. Take any sequence $\{\zeta_n\}_n \subset C_c^\infty((0, T) \times \Omega; \mathbb{R}^2)$ which converges to ∇P in $L^2(h; \mathbb{R}^2)$ and $\mathcal{L}^3$ -a.e. in $(0, T) \times \Omega$. By (3.7) and (3.8),

$$\begin{aligned} & \int_0^T \int_{\Omega^2} |\xi \circ \zeta_m(t, y) - \xi \circ \zeta_n(t, y)|^2 \sigma(dt, dx, dy) = \\ & \quad \int_0^T \int_{\Omega} |\xi \circ \zeta_m(t, y) - \xi \circ \zeta_n(t, y)|^2 h_t(y) dy dt \end{aligned}$$

and ξ is Lipschitz, we deduce that $(0, T) \times \Omega^2 \ni (t, x, y) \mapsto \xi \circ \zeta_n(t, y)$ is a Cauchy sequence in $L^2(\sigma)$. Thus, it is convergent in this space, and, by a standard argument, we see that the limit of $(t, x, y) \mapsto \xi \circ \zeta_n(t, y)$ in $L^2(\sigma)$ will not depend on the choice of the sequence $\{\zeta_n\}_n$: indeed, if sequences $\{\zeta_n^{(1)}\}_n$ and $\{\zeta_n^{(2)}\}_n$ converge to ∇P in $L^2(h; \mathbb{R}^2)$ then the sequence $\{\bar{\zeta}_n\}_n$ defined by $\bar{\zeta}_{2n} = \zeta_n^{(1)}$ and $\bar{\zeta}_{2n+1} = \zeta_n^{(2)}$, also converges to ∇P in $L^2(h; \mathbb{R}^2)$, and thus $(0, T) \times \Omega^2 \ni (t, x, y) \mapsto \xi \circ \bar{\zeta}_n(t, y)$ is a Cauchy sequence in $L^2(\sigma)$, which proves the assertion. This gives unequivocal meaning to $(0, T) \times \Omega^2 \ni (t, x, y) \mapsto \xi \circ \nabla P_t(y)$ as an element of $L^2(\sigma)$.

Next, note that

$$\begin{aligned} & \int_0^T \int_{\Omega^2} |y|^2 |\nabla \xi \circ \zeta_m(t, y) - \nabla \xi \circ \zeta_n(t, y)|^2 \sigma(dt, dx, dy) \\ & \leq \sup_{y \in \Omega} |y|^2 \int_0^T \int_{\Omega} |\nabla \xi \circ \zeta_m(t, y) - \nabla \xi \circ \zeta_n(t, y)|^2 h_t(y) dy dt. \end{aligned} \quad (3.10)$$

But

$$\int_0^T \int_{\Omega} |\nabla \xi \circ \nabla P_t(y) - \nabla \xi \circ \zeta_n(t, y)|^2 h_t(y) dy dt \xrightarrow{n \rightarrow \infty} 0$$

due to the fact that the integrand converges to zero pointwise $\mathcal{L}^3$ -a.e. in $(0, T) \times \Omega$ and is bounded a.e. by $4\|\nabla \xi\|_{\infty}^2 h \in L^1((0, T) \times \Omega)$. Thus, $\nabla \xi \circ \zeta_n$ converges to $\nabla \xi \circ \nabla P$ in $L^2(h; \mathbb{R}^2)$. This implies $\{\nabla \xi \circ \zeta_n\}_n$ is Cauchy in $L^2(h; \mathbb{R}^2)$, which, by (3.10), shows that $(0, T) \times \Omega^2 \ni (t, x, y) \mapsto \{\nabla \xi \circ \zeta_n(t, y) \cdot y\}_n$ is Cauchy in $L^2(\sigma)$. Since, as observed above, its limit in $L^2(\sigma)$ will be independent of the chosen sequence $\{\zeta_n\}_n$, we now have a meaning for $(0, T) \times \Omega^2 \ni (t, x, y) \mapsto \nabla \xi \circ \nabla P_t(y) \cdot y$ as an element of $L^2(\sigma)$.

Finally, to show that $(0, T) \times \Omega^2 \ni (t, x, y) \mapsto (\nabla \xi \circ \nabla P_t(y)) \cdot \nabla P_t(y) \in L^2(\sigma)$, we first show that $(\nabla \xi \circ \zeta_n) \cdot \zeta_n$ converges to $(\nabla \xi \circ \nabla P) \cdot \nabla P$ in $L^2(h; \mathbb{R}^2)$. Indeed,

$$\begin{aligned} & \int_0^T \int_{\Omega} |(\nabla \xi \circ \zeta_n(t, y)) \cdot \zeta_n(t, y) - (\nabla \xi \circ \nabla P_t(y)) \cdot \nabla P_t(y)|^2 h_t(y) dy dt \\ & \leq 2 \int_0^T \int_{\Omega} |\nabla \xi \circ \zeta_n(y)|^2 |\zeta_n(t, y) - \nabla P_t(y)|^2 h_t(y) dy dt \\ & \quad + 2 \int_0^T \int_{\Omega} |\nabla P_t(y)|^2 |\nabla \xi(\zeta_n(t, y)) - \nabla \xi(\nabla P_t(y))|^2 h_t(y) dy dt \\ & =: I_1 + I_2. \end{aligned}$$

Now, $0 \leq I_1 \leq 2\|\nabla \xi\|_{L^\infty(\mathbb{R}^2)}^2 \|\zeta_n - \nabla P\|_{L^2(h; \mathbb{R}^2)}^2 \rightarrow 0$. Also, $I_2 \rightarrow 0$ by Dominated Convergence, since $|\nabla P|^2 h \in L^1((0, T) \times \Omega)$. It follows that the sequence $(\nabla \xi \circ \zeta_n) \cdot \zeta_n$ converges in $L^2(h; \mathbb{R}^2)$, thus it is Cauchy there, which leads to its being Cauchy in $L^2(\sigma)$ (again, by (3.8)). $\square$

We now have all the ingredients for the new definition.

Definition 3.5. *Let P_0 satisfy (2.2). Consider a Borel function*

$$P : [0, T) \times \mathbb{R}^2 \rightarrow \mathbb{R} \cup \{\infty\} \text{ such that } P(t, \cdot) \text{ is convex for all } t \in [0, T),$$

and a

$$\text{Borel family of probability measures } [0, T) \ni t \mapsto \sigma_t \in \mathcal{P}(\Omega \times \Omega).$$

Let σ be given by $d\sigma = d\sigma_t dt$ as in (3.7). We say that (P, σ) is a relaxed Lagrangian solution for the SGSW system (1.1) with initial data P_0 if

- i. $P_t - |\text{Id}_\Omega|^2/2 =: h_t \in \mathcal{P}(\Omega)$ for all $t \in [0, T]$;
- ii. $\nabla P \in L^2(h; \mathbb{R}^2)$ (as functions of both variables);
- iii. (3.8), (3.9) hold.

Let us now record a useful fact, which will be used to prove the next theorem, as well as other statements in the sequel.

Lemma 3.6. *Suppose $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ and γ is an optimal plan between μ and ν for the quadratic cost. Then*

$$W_2(\mu, \nu) \geq \|\bar{\gamma} - \text{Id}_{\mathbb{R}^d}\|_{L^2(\mu; \mathbb{R}^d)}, \quad (3.11)$$

where $\bar{\gamma}$ is the barycentric projection of γ onto μ (see (1.7)).

Proof. From (1.7) applied to μ and ν , and by using the properties

$$\text{proj}_x \gamma = \mu, \quad \text{proj}_y \gamma = \nu$$

(which hold since $\gamma(x, y)$ is a transportation plan between μ and ν), we get

$$\int_{\mathbb{R}^d} \xi(x) \cdot \bar{\gamma}(x) \mu(dx) \leq \left(\int_{\mathbb{R}^d} |\xi(x)|^2 \mu(dx) \right)^{1/2} \left(\int_{\mathbb{R}^d} |y|^2 \nu(dy) \right)^{1/2},$$

for all $\xi \in C_c(\mathbb{R}^d; \mathbb{R}^d)$, which implies that $\|\text{Id}_{\mathbb{R}^d}\|_{L^2(\nu; \mathbb{R}^d)} \geq \|\bar{\gamma}\|_{L^2(\mu; \mathbb{R}^d)}$. This leads to

$$\begin{aligned} W_2^2(\mu, \nu) &= \int_{\mathbb{R}^d} |x|^2 \mu(dx) + \int_{\mathbb{R}^d} |y|^2 \nu(dy) - 2\langle \text{Id}_{\mathbb{R}^d}, \bar{\gamma} \rangle_{L^2(\mu; \mathbb{R}^d)} \\ &\geq \|\bar{\gamma} - \text{Id}_{\mathbb{R}^d}\|_{L^2(\mu; \mathbb{R}^d)}^2, \end{aligned}$$

which finishes the proof. $\square$

We will now prove a very helpful proposition, which shows that relaxed Lagrangian solutions in physical space produce dual space solutions that are Lipschitz in time (with respect to the Wasserstein distance); this is due to the renormalization property.

Theorem 3.7. *Let (P, σ) be a relaxed Lagrangian solution as in Definition 3.5 and let $\alpha_t := \nabla P_{t\#} h_t$, where $h_t := P_t - |\text{Id}_\Omega|^2/2$. Then:*

- (i) α is a distributional solution of the initial-value problem (1.2)–(1.6) with $\alpha_0 = \nabla P_{0\#} h_0$ in dual space, i.e.

$$\int_0^T \int_{\mathbb{R}^2} \{ \partial_t \zeta(t, X) + J[X - \bar{\gamma}_t(X)] \cdot \nabla \zeta(t, X) \} \alpha_t(dX) dt + \int_{\mathbb{R}^2} \zeta(0, X) \alpha_0(dX) = 0$$

for all $\zeta \in C_c^\infty([0, T] \times \mathbb{R}^2)$, where $\bar{\gamma}_t$ is the barycentric projection onto α_t of the optimal plan between α_t and h_t (see (1.7)).

- (ii) $\alpha \in AC^\infty(0, T; \mathcal{P}_2(\mathbb{R}^2))$ (see [5] for the definition of this set).

Proof. Pick arbitrary $\varphi \in C_c^1[0, T]$ and $\xi \in C_c^1(\mathbb{R}^2)$. For the test function $\zeta(t, X) = \varphi(t)\xi(X)$ in dual space, we have, by using (3.8),

$$\begin{aligned} I &:= \int_0^T \int_{\mathbb{R}^2} \dot{\varphi}(t) \xi(X) \alpha_t(dX) dt = \int_0^T \int_{\Omega} \dot{\varphi}(t) \xi(\nabla P_t(y)) h_t(y) dy dt \\ &= \int_0^T \int_{\Omega \times \Omega} \dot{\varphi}(t) \xi(\nabla P_t(y)) \sigma(dt, dx, dy). \end{aligned}$$

Furthermore, by (3.9),

$$\begin{aligned} -I &= \int_0^T \int_{\Omega \times \Omega} \varphi(t) \nabla \xi(\nabla P_t(y)) \cdot J \left(\nabla P_t(y) - y \right) \sigma(dt, dx, dy) \\ &\quad + \int_{\Omega} \varphi(0) \xi(\nabla P_0(x)) h_0(x) dx \\ &= \int_0^T \int_{\Omega} \varphi(t) \nabla \xi(\nabla P_t(y)) \cdot J \left(\nabla P_t(y) - y \right) h_t(y) dy dt + \int_{\mathbb{R}^2} \varphi(0) \xi(X) \alpha_0(dX) \\ &= I_1 + I_2 + \int_{\mathbb{R}^2} \varphi(0) \xi(X) \alpha_0(dX). \end{aligned}$$

Then,

$$\begin{aligned} I_1 &= \int_0^T \int_{\Omega} \varphi(t) \nabla \xi(\nabla P_t(y)) \cdot J \nabla P_t(y) h_t(y) dy dt \\ &= \int_0^T \int_{\mathbb{R}^2} \varphi(t) \nabla \xi(X) \cdot J X \alpha_t(dX) dt. \end{aligned}$$

Let $\gamma_t := (\text{Id}_{\Omega} \times \nabla P_t)_{\#} h_t$ to get

$$\begin{aligned} I_2 &= - \int_0^T \int_{\Omega} \varphi(t) \nabla \xi(\nabla P_t(y)) \cdot J y h_t(y) dy dt \\ &= - \int_0^T \int_{\Omega \times \mathbb{R}^2} \varphi(t) \nabla \xi(X) \cdot J y \gamma_t(dy, dX) dt \\ &= - \int_0^T \int_{\mathbb{R}^2} \varphi(t) \nabla \xi(X) \cdot J \bar{\gamma}_t(X) \alpha_t(dX) dt. \end{aligned}$$

Combine the above to obtain the equation from (i) for $\zeta(t, x) = \varphi(t)\xi(X)$. Thus, (i) is proved.

Note that, according to (3.11), the minimization principle and the triangle inequality, we have, after setting $U_t := J[\text{Id}_{\mathbb{R}^2} - \bar{\gamma}_t]$,

$$\|U_t\|_{L^2(\alpha_t, \mathbb{R}^2)} \leq W_2(\alpha_t, h_t) \leq W_2(\alpha_t, \chi) + \sqrt{\mathcal{L}^2(\Omega)}. \quad (3.12)$$

So, according to Theorem 8.3.1 [5], we have

$$W_2(\alpha_s, \alpha_t) \leq \int_s^t \|U_{\tau}\|_{L^2(\alpha_{\tau}, \mathbb{R}^2)} d\tau \leq T \sqrt{\mathcal{L}^2(\Omega)} + \int_s^t W_2(\alpha_{\tau}, \chi) d\tau \quad (3.13)$$

for all $0 \leq s \leq t \leq T$. By the triangle inequality, we get

$$W_2(\alpha_t, \chi) \leq T\sqrt{\mathcal{L}^2(\Omega)} + W_2(\alpha_s, \chi) + \int_s^t W_2(\alpha_\tau, \chi) d\tau \text{ if } 0 \leq s \leq t \leq T.$$

Gronwall's Lemma implies $[0, T] \ni t \mapsto W_2(\alpha_t, \chi)$ is bounded, which, in view of (3.12) and the first inequality in (3.13), yields (ii). $\square$

Note that Definition 3.5 takes into account the key features of Definition 3.1. Indeed, the property $F(t, \cdot)_\# h_0 = h_t$ for all $t \in [0, T]$ is accounted for by (3.8), and (3.3) follows from (3.5) which is accounted for by (3.9). The time-continuity conditions from Definition 3.1 can be translated in terms of the relaxed Lagrangian solutions of Definition 3.5 (see remark below). However, it is not clear whether they will be satisfied in general.

Lemma 3.8. *Let (P, σ) be a relaxed Lagrangian solution for SGSW with initial P_0 satisfying (2.2). Let $F : [0, T] \times \Omega \rightarrow \Omega$ be a Borel map such that $F_{t\#} h_0 = h_t$ for all $t \in [0, T]$, where $h_t := P_t - |\text{Id}_\Omega|^2/2$ in Ω . Then $F \in C([0, T]; L^2(h_0; \mathbb{R}^2))$ is equivalent to $F \in C([0, T]; L^2_{weak}(h_0; \mathbb{R}^2))$.*

Proof. By Theorem 3.7, $\alpha_t := \nabla P_{t\#} h_t$ is a dual-space solution $\alpha \in AC^\infty(0, T; \mathcal{P}_2(\mathbb{R}^2))$ corresponding to the initial α_0 . Thus, for any $t_0 \in [0, T]$ we have that $W_2(\alpha_t, \alpha_{t_0}) \rightarrow 0$ as $t \rightarrow t_0$. By Proposition 2.1 (ii) we deduce $\|h_t - h_{t_0}\|_{L^2(\Omega)} \rightarrow 0$, so

$$\|F_t\|_{L^2(h_0; \mathbb{R}^2)}^2 = \int_\Omega |y|^2 h_t(y) dy \xrightarrow{n \rightarrow \infty} \int_\Omega |y|^2 h_{t_0}(y) dy = \|F_{t_0}\|_{L^2(h_0; \mathbb{R}^2)}^2,$$

which finishes the proof. $\square$

Remark 3.9. *From Lemma 3.8, the $C([0, T]; L^2(h_0; \mathbb{R}^2))$ continuity of F reads (in terms of σ defined by (3.6))*

$$\lim_{t \rightarrow t_0} \int_{\Omega^2} \zeta(x) \cdot y \sigma_t(dx, dy) = \int_{\Omega^2} \zeta(x) \cdot y \sigma_{t_0}(dx, dy) \text{ for all } \zeta \in C_c(\Omega; \mathbb{R}^2). \quad (3.14)$$

Since σ_t is only unambiguously defined (from σ) for a.e. $t \in [0, T]$, the above equality should be understood in the sense that there exists a Borel disintegration $d\sigma = d\sigma_t dt$ for which (3.14) holds. Consequently, the condition $F(0, \cdot) \equiv \text{Id}_\Omega$ translates to

$$\int_{\Omega^2} \zeta(x) \cdot y \sigma_0(dx, dy) = \int_\Omega \zeta(x) \cdot x h_0(x) dx \text{ for all } \zeta \in C_c(\Omega; \mathbb{R}^2). \quad (3.15)$$

If we denote by $\bar{\sigma}_t$ the barycentric projection of σ_t as defined in (1.7), then (3.14) and (3.15) combine into

$$[0, T] \ni t \mapsto \bar{\sigma}_t \in L^2_{weak}(h_0; \mathbb{R}^2) \text{ is continuous and } \bar{\sigma}_0 \equiv \text{Id}_\Omega \text{ } h_0\text{-a.e.} \quad (3.16)$$

However, as the example in [16] clearly shows, the corresponding property for the SG system cannot be expected from our relaxed solutions in general. It is not difficult to explicitly construct dual-space solutions to SGSW that are point masses for all times; however, we have not been able to explicitly construct absolutely continuous approximations of such solutions.

3.1. Consistency. Here we shall argue that relaxed Lagrangian solutions with some extra regularity determine classical solutions of SGSW (1.1). We start by showing that relaxed Lagrangian solutions with σ_t supported on graphs and satisfying (3.14) and (3.15) give rise to weak Lagrangian solutions as in [9], except the almost everywhere invertibility of the measure preserving maps may fail.

Theorem 3.10. *Let (P, σ) be a relaxed Lagrangian solution for SGSW. Suppose (3.16) is satisfied, $P \in L^\infty(0, T; W^{1,\infty}(\Omega))$, and $\sigma_t := (\text{Id}_\Omega \times F_t)_\# h_0$ for all $t \in [0, T)$ and some Borel map $F : [0, T) \times \Omega \rightarrow \Omega$. Then (P, F) is a weak Lagrangian solution as in Definition 3.1 except, possibly, (iii) and $P \in C([0, T); H^1(\Omega))$.*

Proof. Note that the push-forward properties of F_t are direct consequences of (3.8). Next, we have that (3.14) implies the weak convergence of F_t to F_{t_0} in $L^2(h_0)$ as $t \rightarrow t_0$. By Lemma 3.8 we get that F_t converges to F_{t_0} strongly in $L^2(h_0; \mathbb{R}^2)$. Then (3.15) yields $F_0 \equiv \text{Id}_\Omega$ h_0 -a.e. in Ω .

Finally, let $\varphi \in C_c^1([0, T) \times \Omega; \mathbb{R}^2)$. Denote $Z(t, x) := \nabla P(t, F(t, x))$ and let $\xi(X) = X_i$ (this choice is admissible now that $P \in L^\infty(0, T; W^{1,\infty}(\Omega))$, as previously discussed), $\zeta(t, x) := \varphi_i(t, x)$ for $i = 1, 2, 3$ in (3.9). Then add the resulting equations to obtain (3.3). $\square$

As a result, we have (see [9, Lemma 3.7]) the following:

Corollary 3.11. *Under the same hypotheses from Theorem 3.10 and the extra assumptions that $\partial_t F \in L^\infty((0, T) \times \Omega; \mathbb{R}^2)$ and F^* as in Definition 3.1 exists, let us define $\mathbf{u}(t, \cdot) := \partial_t F(t, F^*(t, \cdot))$. Then $(P, \mathbf{u})$ is a weak (Eulerian) solution for SGSW in physical space, i.e. for (1.1), in the following sense:*

$$\begin{aligned} & \int_0^T \int_\Omega \{ \nabla P(t, x) \cdot [\partial_t \phi(t, x) + \nabla \phi(t, x) \mathbf{u}(t, x)] \\ & \quad + J[\nabla P(t, x) - x] \cdot \phi(t, x) \} h(t, x) dx dt + \int_\Omega \nabla P_0(x) \cdot \phi(0, x) h_0(x) dx = 0; \\ & \int_0^T \int_\Omega \mathbf{u}(t, x) \cdot [\partial_t \psi(t, x) + \nabla \psi(t, x)] h(t, x) dx dt + \int_\Omega \psi(0, x) h_0(x) dx = 0 \end{aligned}$$

for any $\phi \in C_c^1([0, T) \times \Omega; \mathbb{R}^2)$, $\psi \in C_c^1([0, T) \times \overline{\Omega})$.

Furthermore, if $(F, F^*, P) \in C^2([0, T] \times \overline{\Omega})$, then the function $\mathbf{u}(t, \cdot) := \partial_t F(t, F^*(t, \cdot))$ satisfies $\mathbf{u} \in C^1([0, T] \times \overline{\Omega}; \mathbb{R}^2)$, and $(P, \mathbf{u})$ is a classical solution of (1.1) in $[0, T] \times \Omega$.

Our approach to proving existence of such solutions consists of approximating a given P_0 which satisfies (2.2) in the following manner: we first approximate $\alpha_0 := \nabla P_0 \# h_0$ by measures $\alpha_0^n \ll \mathcal{L}^2$, for which solutions in the sense of the above definition can easily be constructed as shown above from the stronger solutions of Cullen & Feldman, Theorem 3.1 [9] in the sense of Definition 3.4 [9]. Then we wish to pass to the limit as $n \rightarrow \infty$ by using Theorem 4.1 in order to obtain a solution with initial data α_0 . All this will be achieved in Section 5.

4. Weak stability of solutions in dual space. Let us recall [5] the definition of the set of absolutely continuous paths/curves $AC^p(0, T; \mathcal{P}_2(\mathbb{R}^2))$ (for $1 \leq p \leq \infty$) as the set of all paths $\mu : [0, T] \ni t \rightarrow \mu_t \in \mathcal{P}_2(\mathbb{R}^2)$ for which there exists $\beta \in L^p(0, T)$ such that

$$W_2(\mu_s, \mu_t) \leq \int_s^t \beta(\tau) d\tau \text{ for all } 0 \leq s \leq t \leq T,$$

where W_2 is the quadratic Wasserstein distance [19].

Now we prove the stability of dual-space solutions that are Lipschitz in time (i.e., lying in $AC^\infty(0, T; \mathcal{P}_2(\mathbb{R}^2))$). Since the convex functions P_t in (1.3) and the optimal transport plans $\gamma_t = (\nabla P_t \times \text{Id}_\Omega)_\# h_t$ between α_t and h_t are unambiguously defined by α_t (by uniqueness of the minimizer $h_t := \mathcal{H}[\alpha_t]$; see Proposition 2.1), we will discuss the dual problem (1.2) – (1.6) in terms of α_t only.

Theorem 4.1. *Let $\{\alpha_0\} \cup \{\alpha_0^n\}_{n \geq 1} \subset \mathcal{P}_2(\mathbb{R}^2)$ be such that*

$$\alpha_0^n \rightharpoonup \alpha_0 \text{ narrowly, and } \sup_{n \geq 1} W_2(\alpha_0^n, \alpha_0) < \infty. \quad (4.1)$$

For all integers $n \geq 1$, let $\alpha^n \in AC^\infty(0, T; \mathcal{P}_2(\mathbb{R}^2))$ be solutions of (1.2)–(1.6), corresponding to the initial data α_0^n , respectively. Then a subsequence of α^n converges to a solution $\alpha \in AC^\infty(0, T; \mathcal{P}_2(\mathbb{R}^2))$ with initial data α_0 . More precisely, we have

(i) $\alpha_t^n \rightarrow \alpha_t$ in $\mathcal{P}_p(\mathbb{R}^2)$ for all $t \in [0, T]$ and all $p \in [1, 2)$. Also,

$$\sup \{W_2(\alpha_t^n, \alpha_0) : n \geq 1, t \in [0, T]\} < \infty.$$

(ii) $\mathcal{H}[\alpha_t^n] \rightharpoonup \mathcal{H}[\alpha_t]$ weakly in $L^2(\Omega)$ for all $t \in [0, T]$.

(ii) $\alpha \in AC^\infty(0, T; \mathcal{P}_2(\mathbb{R}^2))$ is a solution of (1.2)–(1.6) with initial data α_0 .

Proof. First, note that we can use (1.2) satisfied by α^n , Theorem 8.3.1 [5] and (3.11), to write

$$W_2(\alpha_s^n, \alpha_t^n) \leq \int_s^t \|J[\text{Id}_{\mathbb{R}^2} - \bar{\gamma}_\tau^n]\|_{L^2(\alpha_\tau^n; \mathbb{R}^2)} d\tau \leq \int_s^t W_2(\alpha_\tau^n, h_\tau^n) d\tau \quad (4.2)$$

for all $0 \leq s \leq t \leq T$. By the variational principle for h_τ^n , we have

$$W_2(\alpha_\tau^n, h_\tau^n) \leq W_2(\alpha_\tau^n, \chi) + \sqrt{\mathcal{L}^2(\Omega)} \text{ for all } \tau \in [0, T] \text{ and all integers } n \geq 1.$$

Thus,

$$W_2(\alpha_t^n, \chi) \leq t\sqrt{\mathcal{L}^2(\Omega)} + W_2(\alpha_s^n, \chi) + \int_s^t W_2(\alpha_\tau^n, \chi) d\tau$$

for all $0 \leq s \leq t \leq T$. By Gronwall's Lemma we get

$$W_2(\alpha_t^n, \chi) \leq e^T [W_2(\alpha_0^n, \chi) + T\sqrt{\mathcal{L}^2(\Omega)}] \text{ for all } n \geq 1 \text{ and } t \in [0, T].$$

But (4.1) ensures that $W_2(\alpha_0^n, \mu)$ is bounded for any $\mu \in \mathcal{P}_2(\mathbb{R}^2)$, uniformly in n . Thus,

$$W_2(\alpha_t^n, \chi) \text{ is bounded uniformly in } n \geq 1 \text{ and } t \in [0, T]. \quad (4.3)$$

Since

$$\int_{\mathbb{R}^2} |X|^2 \alpha_t^n(dX) = W_2^2(\alpha_t^n, \delta_0) \leq 2W_2^2(\alpha_t^n, \chi) + 2 \int_{\Omega} |x|^2 dx,$$

we infer that

$$\alpha_t^n \text{ has bounded second moment, uniformly in } t \in [0, T], n \geq 1. \quad (4.4)$$

(Note that (4.3) and (4.4) are equivalent; however, we record them both since further reference may prefer one form over the other.) Fix $p \in [1, 2)$. By a standard diagonalization argument, a subsequence (not relabeled) can be extracted such that $\alpha_t^n \rightarrow \alpha_t$ in $\mathcal{P}_p(\mathbb{R}^2)$ for every $t \in [0, T] \cap \mathbb{Q}$ and for some Borel probabilities α_t . Again by diagonalization (and, possibly, a further subsequence extraction), one can ensure that this convergence holds for all $p \in [1, 2)$, thus

$$\alpha_t^n \rightarrow \alpha_t \text{ in } \mathcal{P}_p(\mathbb{R}^2) \text{ for every } t \in [0, T] \cap \mathbb{Q}, p \in [1, 2). \quad (4.5)$$

Another look at (4.2) and (4.3) reveals

$$W_2(\alpha_s^n, \alpha_t^n) \leq C|s - t|, \text{ with } C \in \mathbb{R} \text{ independent of } n \geq 1, s, t \in [0, T]. \quad (4.6)$$

This leads to

$$W_p(\alpha_s, \alpha_t) \leq W_p(\alpha_s^n, \alpha_s) + W_p(\alpha_t^n, \alpha_t) + C|s - t|$$

for all $s, t \in [0, T] \cap \mathbb{Q}$, $n \geq 1$, $p \in [1, 2)$. Let $n \rightarrow \infty$ and use the completeness of $\mathcal{P}_p(\mathbb{R}^2)$ to infer that there exists α_t for all times $t \in [0, T]$ such that

$$W_p(\alpha_s, \alpha_t) \leq C|s - t| \text{ for all } p \in [1, 2) \text{ and all } s, t \in [0, T]. \quad (4.7)$$

Then, the inequality

$$W_p(\alpha_t^n, \alpha_t) \leq W_p(\alpha_s^n, \alpha_s) + W_p(\alpha_s^n, \alpha_t^n) + W_p(\alpha_s, \alpha_t)$$

yields, by (4.5), (4.6) and (4.7), that α_t^n converges to α_t in $\mathcal{P}_p(\mathbb{R}^2)$ for all $t \in [0, T]$ and all $1 \leq p < 2$. Obviously, α_t has bounded second moment uniformly in $t \in [0, T]$; indeed, this follows from (4.4) and the lower-semicontinuity property

$$\int_{\mathbb{R}^2} |X|^2 \alpha_t(dX) \leq \liminf_{n \rightarrow \infty} \int_{\mathbb{R}^2} |X|^2 \alpha_t^n(dX), \quad (4.8)$$

which holds due to the narrow convergence of α_t^n to α_t (see [5], subsection 5.1.1). Thus, we can let $p \rightarrow 2^-$ in (4.7) (or, alternatively, use the lower-semicontinuity of W_2 with respect to narrow convergence) to get

$$W_2(\alpha_s, \alpha_t) \leq C|s - t| \text{ for all } s, t \in [0, T]. \quad (4.9)$$

So, $\alpha \in AC^\infty(0, T; \mathcal{P}_2(\mathbb{R}^2))$. By Proposition 2.5, we have that $\{h_t^n\}_n$ converges weakly in $L^2(\Omega)$ to $h_t := \mathcal{H}[\alpha_t]$, which, by Corollary 2.2 and Proposition 2.3, is equivalent to (1.3) plus (1.4). From (1.7) we get

$$\int_{\mathbb{R}^2} (\bar{\gamma}_t^n(X) - X) \cdot \xi(X) \alpha_t^n(dX) = \iint_{\mathbb{R}^2 \times \Omega} (y - X) \cdot \xi(X) \gamma_t^n(dX, dy)$$

for all $\xi \in C_c(\mathbb{R}^2; \mathbb{R}^2)$, where $\gamma_t^n = (\nabla P_t^n \times \text{Id}_\Omega)_\# h_t^n$ is the optimal plan between α_t^n and h_t^n . Due to (4.3) and the narrow convergence of α_t^n to α_t and of h_t^n to h_t , we can use Proposition 7.1.3 [5] (stability of optimal plans with respect to narrow convergence) and the uniqueness of the optimal plan $\gamma_t = (\nabla P_t \times \text{Id}_\Omega)_\# h_t$ between α_t and h_t to deduce that the left hand side of the last displayed equation converges to $\langle \bar{\gamma}_t - \text{Id}_{\mathbb{R}^2}, \xi \rangle_{L^2(\alpha_t; \mathbb{R}^2)}$. The equation (1.2) (with initial α_0^n) satisfied by α^n in the sense of distributions reads

$$\int_0^T \int_{\mathbb{R}^2} \{ \partial_t \zeta(t, X) + J[X - \bar{\gamma}_t^n(X)] \cdot \nabla \zeta(t, X) \} \alpha_t^n(dX) dt + \int_{\mathbb{R}^2} \zeta(0, X) \alpha_0^n(dX) = 0$$

for all $\zeta \in C_c^\infty([0, T] \times \mathbb{R}^2)$. According to the above considerations, we can pass to the limit (use Dominated Convergence for the time integral in view of (3.11) and (4.3)) to obtain that α is a distributional solution for (1.2), (1.5), (1.6) with initial measure α_0 . $\square$

Remark 4.2. *The above result can be improved by first proving that we are in the correct setting for (the conservation of the Hamiltonian) Theorem 5.2 [4] to hold. Indeed, if*

$$J[\alpha_t^n](h_t^n) = J[\alpha_0^n](h_0^n) \text{ for all } t \in [0, T], \quad n \geq 1$$

and

$$J[\alpha_t](h_t) = J[\alpha_0](h_0) \text{ for all } t \in [0, T],$$

then we can take the initial approximations α_0^n such that $W_2(\alpha_0^n, \alpha_0) \rightarrow 0$ to get that h_t^n converges strongly to h_t in $L^2(\Omega)$ for all $t \in [0, T]$. To see this, note that Proposition 2.1 and the above displayed identities yield

$$\lim_{n \rightarrow \infty} J[\alpha_t^n](h_t^n) = J[\alpha_t](h_t) \text{ for all } t \in [0, T]. \quad (4.10)$$

But, as we have seen in the proof of Theorem 4.1, we have α_t^n converges narrowly to α_t and $h_t^n \rightharpoonup h_t$ in $L^2(\Omega)$ (which implies $W_2(h_t^n, h_t) \rightarrow 0$), which gives

$$\liminf_{n \rightarrow \infty} W_2(\alpha_t^n, h_t^n) \geq W_2(\alpha_t, h_t) \text{ and } \liminf_{n \rightarrow \infty} \|h_t^n\|_{L^2(\Omega)} \geq \|h_t\|_{L^2(\Omega)}.$$

Now (4.10) comes in to show that the last two $\liminf$'s are, in fact, limits, and the two inequalities above are equalities. We conclude that $\|h_t^n - h_t\|_{L^2(\Omega)} \rightarrow 0$.

Even though our goal is to simply use this dual-space stability result to prove a similar result for relaxed Lagrangian solutions in physical space, we

finish this section by noting that Theorem 4.1 implies, in view of the existence result for “nice” initial data in [4] (or, alternatively, see the first part of the proof of Theorem 5.1):

Corollary 4.3. *The dual-problem (1.2)–(1.6) admits a distributional solution $\alpha \in AC^\infty(0, T; \mathcal{P}_2(\mathbb{R}^2))$ for any initial $\alpha_0 \in \mathcal{P}_2(\mathbb{R}^2)$.*

5. Main results. Here we state and prove the main results of the paper. In the first subsection we make the connection with dual space solutions. Then we move on to the stability and existence results on the relaxed Lagrangian solutions in physical space.

5.1. Weak stability and existence of relaxed Lagrangian solutions. Before our existence result, we need to prove what we call a “weak stability” result, i.e. for appropriate approximations of the initial data and corresponding relaxed Lagrangian solutions, we obtain convergence (up to a subsequence) to a relaxed Lagrangian solution.

Theorem 5.1. *Let P_0, P_0^n satisfy (2.2) for all positive integers n . Assume (P^n, σ^n) are relaxed solutions for SGSW in physical space corresponding to the initial data P_0^n . Then, possibly up to a subsequence, (P^n, σ^n) converges to a relaxed solution (P, σ) corresponding to the initial datum P_0 . The convergence is in the following sense:*

- (i) $P_t^n \rightharpoonup P_t$ weakly in $L^2(\Omega)$ and locally uniformly in Ω for all $t \in [0, T]$;
- (ii) $\nabla P_t^n \rightarrow \nabla P_t$ a.e. in Ω for all $t \in [0, T]$, and $\{\nabla P_t\}_n$ is locally bounded uniformly with respect to $t \in [0, T]$ and $n \geq 1$;
- (iii) σ^n converges narrowly to σ .

Furthermore, the corresponding dual space solutions satisfy $W_p(\alpha_t^n, \alpha_t) \rightarrow 0$ for all $t \in [0, T]$ and all $1 \leq p < 2$.

Proof. The measures σ^n have same mass and bounded moments of any order, uniformly in n . By a diagonalization argument, we can extract a subsequence (not relabeled) such that σ^n converges in the Wasserstein distance W_r of any order $1 \leq r < \infty$ to some measure σ on $(0, T) \times \Omega \times \Omega$ with finite moments of any order. Let $\alpha_0^n := \nabla P_{0\#}^n h_0^n$; according to Theorem 3.7, for each $t \in [0, T]$ the measures $\alpha_t^n := \nabla P_{t\#}^n h_t^n$ are dual space solutions (corresponding to the initial data α_0^n , respectively) in the sense of Theorem 4.1 (i.e., distributional solutions in dual-space which also lie in $AC^\infty(0, T; \mathcal{P}_2(\mathbb{R}^2))$). Therefore, according to Theorem 4.1, possibly up to a further subsequence, we obtain a limiting $\alpha \in AC^\infty(0, T; \mathcal{P}_2(\mathbb{R}^2))$ solving SGSW in dual space. From Theorem 4.1 (ii) we get that $h_t^n := \mathcal{H}[\alpha_t^n]$ converges weakly in $L^2(\Omega)$ to $h_t := \mathcal{H}[\alpha_t]$. This, combined with the fact that σ^n disintegrates like $d\sigma_t^n dt$ in the sense of (3.7) and the σ_t^n ’s satisfy (3.8) with marginals h_0^n and h_t^n for each integer $n \geq 1$, shows that for the limiting σ , disintegrating $d\sigma = d\sigma_t dt$, we get that σ_t satisfies (3.8) with marginals h_0 and h_t . By Proposition 2.5, we infer that $\nabla P_t^n \rightarrow \nabla P_t$ a.e. in Ω for all $t \in [0, T]$. Let $\xi \in C_c^1(\mathbb{R}^2)$ and $\phi \in C_b([0, T] \times \Omega \times \Omega)$. Since ξ is bounded,

Ω is bounded and $h, h^n \in L^1((0, T) \times \Omega)$, we have $f := \xi \circ \nabla P \in L^2(h)$, $f^n := \xi \circ \nabla P^n \in L^2(h^n)$, and also $f \in L^2(h^n)$, $f^n \in L^2(h)$. As in the proof of Lemma 3.4, we deduce $f, f^n \in L^2(\sigma) \cap L^2(\sigma^n)$ for all integers $n \geq 1$. So, the integrals below make sense. We use the marginal property (3.8) to see that for any $\varepsilon > 0$ and sufficiently large n we have

$$\begin{aligned} & \left| \int_0^T \int_{\Omega^2} \phi(t, x, y) [f^n(t, y) - f(t, y)] \sigma^n(dt, dx, dy) \right| \\ & \leq \|\phi\|_\infty \int_0^T \int_\Omega |f^n(t, y) - f(t, y)| h_t^n(y) dy dt \leq \frac{\varepsilon}{2}. \end{aligned} \quad (5.1)$$

Indeed, since h_t^n is bounded in $L^2(\Omega)$ uniformly in $t \in [0, T]$ and $n \geq 1$, by Dominated Convergence we get that the above quantity vanishes in the limit as $n \rightarrow \infty$. Next, let $\varepsilon > 0$ and approximate f in $L^2(h)$ by $f_\varepsilon \in C_c((0, T) \times \Omega)$ such that

$$\begin{aligned} & \left| \int_0^T \int_{\Omega^2} \phi(t, x, y) [f_\varepsilon(t, y) - f(t, y)] \sigma(dt, dx, dy) \right| \\ & \leq \|\phi\|_\infty \int_0^T \int_\Omega |f_\varepsilon(t, y) - f(t, y)| h_t(y) dy dt \leq \varepsilon/7 \end{aligned}$$

(since the (t, y) -marginal of σ is h). Since, as seen above, $f \in L^2(\sigma^n)$, it makes sense to write

$$\begin{aligned} & \left| \int_0^T \int_{\Omega^2} \phi(t, x, y) [f_\varepsilon(t, y) - f(t, y)] \sigma^n(dt, dx, dy) \right| \\ & \leq \|\phi\|_\infty \int_0^T \int_\Omega |f_\varepsilon(t, y) - f(t, y)| h_t^n(y) dy dt \leq \varepsilon/6 \end{aligned}$$

for sufficiently large n (as $f_\varepsilon - f \in L^\infty((0, T) \times \Omega)$ and $h_t^n \rightharpoonup h_t$ weakly in $L^2(\Omega)$ with uniform (in n and t) $L^2(\Omega)$ bounds. By the narrow convergence of σ^n to σ , we also have

$$\left| \int_0^T \int_{\Omega^2} \phi(t, x, y) f_\varepsilon(t, y) d(\sigma^n - \sigma) \right| \leq \varepsilon/6 \text{ for sufficiently large } n.$$

We combine the last three displayed inequalities to get

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_0^T \int_{\Omega^2} \phi(t, x, y) \xi(\nabla P(t, y)) \sigma^n(dt, dx, dy) \\ & = \int_0^T \int_{\Omega^2} \phi(t, x, y) \xi(\nabla P(t, y)) \sigma(dt, dx, dy), \end{aligned}$$

which, in by (5.1), yields that

$$[\xi \circ \nabla P^n(t, y)] \sigma^n \text{ converges narrowly to } [\xi \circ \nabla P(t, y)] \sigma. \quad (5.2)$$

Since $Jy \cdot \nabla \xi(\nabla P(t, y)) \in L^\infty((0, T) \times \Omega)$, we can repeat the same argument as above to obtain

$$[Jy \cdot \nabla \xi \circ \nabla P^n(t, y)] \sigma^n \text{ converges narrowly to } [Jy \cdot \nabla \xi \circ \nabla P(t, y)] \sigma.$$

As far as the term $(t, y) \mapsto J \nabla P(t, y) \cdot \nabla \xi(\nabla P(t, y))$ is concerned, we now use the fact that not only is $\nabla \xi$ bounded, but it also has compact support, so the same argument as for the first term above applies (with $\zeta(X) := X \cdot \nabla \xi(X)$ instead of ξ , as $\zeta \in C_c(\mathbb{R}^2)$ as well); so

$$[J \nabla P^n(t, y) \cdot \nabla \xi \circ \nabla P^n(t, y)] \sigma^n \rightarrow [J \nabla P(t, y) \cdot \nabla \xi \circ \nabla P(t, y)] \sigma \text{ narrowly.}$$

Finally, we use the a.e. convergence of ∇P_0^n to ∇P_0 , the weak $L^2(\Omega)$ convergence of h_0^n to h_0 , and the fact that $\xi \in C_c(\mathbb{R}^2)$ to infer

$$\langle \xi \circ \nabla P_0^n, \zeta(0, \cdot) h_0^n \rangle_{L^2(\Omega)} \rightarrow \langle \xi \circ \nabla P_0, \zeta(0, \cdot) h_0 \rangle_{L^2(\Omega)}.$$

We now have all the ingredients to pass to the limit in (3.9) as satisfied by (P^n, σ^n) and obtain it for (P, σ) . $\square$

Clearly, Theorem 5.1 will give existence of relaxed solutions under the general assumptions (2.2) provided that we produce approximations of P_0 for which relaxed solutions exist.

Corollary 5.2. *Let $\Omega \subset \mathbb{R}^2$ be open, bounded and connected. Let P_0 satisfy (2.2) with respect to Ω . Then there exists a relaxed Lagrangian solution for SGSW corresponding to the initial data P_0 .*

Proof. Let $\alpha_0 := \nabla P_0 \# h_0$, where h_0 is defined as in (2.2). Since $\alpha_0 \in \mathcal{P}_2(\mathbb{R}^2)$, we can approximate it in the Wasserstein distance W_2 by $\alpha_0^n \in \mathcal{P}(\mathbb{R}^2) \cap C_c^\infty(\mathbb{R})$ as in the proof of Corollary 2.2. By Proposition 2.1 (iii) there exists $P_0^n : \mathbb{R}^2 \rightarrow \mathbb{R} \cup \{\infty\}$ convex and finite on an open ball containing $\bar{\Omega}$ (therefore, by convexity, continuous on the closure of a, possibly, smaller ball containing $\bar{\Omega}$) such that $\nabla P_0^n \# h_0^n = \alpha_0^n$, where $h_0^n := \mathcal{H}[\alpha_0^n] = P_0^n - |\text{Id}_\Omega|^2/2$ in Ω . By Proposition 2.3, we have that $h_0 = \mathcal{H}[\alpha_0]$, so we can apply Proposition 2.1 (iv) to the sequence $\{\alpha_0^n\}_n$ and its limit α_0 to get that P_0^n converges to P_0 strongly in $L^2(\Omega)$. At this stage, it is sufficient to know that relaxed solutions (P^n, σ^n) corresponding to the initial data P_0^n exist, in order to conclude by means of Theorem 5.1. Theorem 3.1 [9] applied to h_0^n (or, equivalently, to P_0^n) guarantees existence of a weak Lagrangian solution (P^n, F^n) in physical space in the sense of Definition 3.1 for the initial data P_0^n . But in the preamble to Definition 3.5 we showed exactly that the pair (P^n, F^n) gives rise to a relaxed Lagrangian solution by setting $\sigma^n := \sigma_t^n dt$ where $\sigma_t^n := (\text{Id}_\Omega \times F_t^n) \# h_0^n$. $\square$

5.2. Time continuity of relaxed Lagrangian solutions. As anticipated in Remark 3.9, we shall cover some weaker properties of the relaxed Lagrangian solutions that follow directly from their definition and the time continuity of P .

Remark 5.3. Define $G_t : \Omega \times \Omega \rightarrow \Omega \times \mathbb{R}^2$ by $G_t(x, y) = (x, \nabla P_t(y))$, then $t \mapsto G_{t\#}\sigma_t$ is continuous with respect to narrow convergence on $\Omega \times \mathbb{R}^2$: i.e., from the equation, for any $\zeta \in C_c^1(\Omega)$, $\xi \in C_c^1(\mathbb{R}^2)$

$$\Psi(t) := \int_{\Omega \times \Omega} \zeta(x) \xi(\nabla P_t(y)) \sigma_t(dx, dy) = \int_{\Omega \times \mathbb{R}^2} \zeta(x) \xi(Y) G_{t\#}\sigma_t(dx, dY)$$

is continuous on $[0, T]$.

Finally, we will make use of the time-regularity P enjoys in order to prove some kind of weak time-regularity for the relaxed Lagrangian solutions. Part (ii) of the proposition below makes a weaker claim, which is reminiscent of the initial condition $F(0, \cdot) \equiv \text{Id}_\Omega$ (yielding $\sigma_0 := (\text{Id}_\Omega \times \text{Id}_\Omega)_\# h_0$) from the regular case $\alpha_t \ll \mathcal{L}^2$ (recall Definition 3.1). More precisely, we have:

Proposition 5.4. (i) For any $t \in [0, T]$, we have that $s \mapsto G_{t\#}\sigma_s$ is continuous at $s = t$ with respect to narrow convergence on $\Omega \times \mathbb{R}^2$.

(ii) We have $G_{0\#}\sigma_0 = G_{0\#}\delta_{\{x=y\}}$, where $\delta_{\{x=y\}} := (\text{Id}_\Omega \times \text{Id}_\Omega)_\# h_0$, and $G_{0\#}\sigma_0$ is understood in the sense of convergence of $G_{t\#}\sigma_t$ as $t \rightarrow 0$ (Remark 5.3).

Proof. First, note that all quantities defined below make sense due to the marginal property (3.8) (a proof along the lines of the one for Lemma 3.4 can be readily written). Let $\zeta \in C_c(\Omega)$ and $\xi \in C_c^1(\mathbb{R}^2)$ to write:

$$\Phi(s; t) := \int_{\Omega} \int_{\mathbb{R}^2} \zeta(x) \xi(Y) G_{t\#}\sigma_s(dx, dY) = \int_{\Omega^2} \zeta(x) \xi(\nabla P_t(y)) \sigma_s(dx, dy).$$

So,

$$\begin{aligned} \Phi(s; t) - \Phi(t; t) &= \int_{\Omega^2} \zeta(x) \xi(\nabla P_t(y)) [\sigma_s - \sigma_t](dx, dy) \\ &= \int_{\Omega^2} \zeta(x) [\xi(\nabla P_t(y)) - \xi(\nabla P_s(y))] \sigma_s(dx, dy) \\ &\quad + \Psi(s) - \Psi(t), \end{aligned}$$

where Ψ is the function (depending on ζ and ξ) of t defined in Remark 5.3 above. According to same remark, it therefore suffices to prove that the first term in the right-hand-side vanishes in the limit as $s \rightarrow t$. To this effect, note that we can use (3.8) to write

$$\left| \int_{\Omega^2} \zeta(x) [\xi(\nabla P_t(y)) - \xi(\nabla P_s(y))] \sigma_s(dx, dy) \right| \leq \|\zeta\|_\infty \|\xi \circ \nabla P_t - \xi \circ \nabla P_s\|_{L^1(h_0)}.$$

We conclude the proof of (i) by Dominated Convergence using that $\nabla P_s \rightarrow \nabla P_t$ a.e. in Ω as $s \rightarrow t$, which follows by $P_s \rightarrow P_t$ strongly in $L^2(\Omega)$ (see proof of Lemma 5.3) and the convexity of these maps [13].

To get (ii), take a test function $f(t)\zeta(x)$ instead of a general $\zeta(t, x)$ in (3.9) to see that, using (3.7), we have

$$\int_0^T [f'(t)\Psi(t) + f(t)\Lambda(t)] dt + f(0) \int_{\Omega} \xi(\nabla P_0(x)) \zeta(x) h_0(x) dx = 0 \quad (5.3)$$

for all $f \in C_c^1[0, T]$, $\zeta \in C_c(\Omega)$ and $\xi \in C_c^1(\mathbb{R}^2)$, where

$$\Lambda(t) := \int_{\Omega^2} \nabla \xi(\nabla P_t(y)) \cdot J[y - \nabla P_t(y)] \zeta(x) \sigma_t(dx, dy).$$

We use (3.8) again, the boundedness of Ω , and the fact that $\xi \in C_c^1(\mathbb{R}^2)$ to conclude that $\Lambda \in L^1(0, T)$. Thus, from (5.3) get $\Psi \in W^{1,1}(0, T)$ with $\Psi' = \Lambda$ a.e. in $(0, T)$ (thus $\Psi \in C[0, T]$), and

$$\Psi(0) = \int_{\Omega} \xi(\nabla P_0(x)) \zeta(x) h_0(x) dx,$$

which finishes the proof. $\square$

Acknowledgements. The authors would like to thank M. Cullen for his valuable suggestions and comments. The work of Mikhail Feldman was supported in part by the National Science Foundation under Grant DMS-1101260, and by the Simons Foundation under the Simons Fellows program. This work was partially supported by a grant from the Simons Foundation (#246063 to Adrian Tudorascu).

REFERENCES

- [1] L. Ambrosio. Transport equation and Cauchy problem for BV vector fields. *Invent. Math.*, **158** (2004), 227–260.
- [2] L. Ambrosio, M. Colombo, G. De Philippis, A. Figalli. Existence of Eulerian solutions to the semigeostrophic equations in physical space: the 2-dimensional periodic case. *Comm. Part. Diff. Eq.* **37**, No. 12 (2012), 2209–2227.
- [3] L. Ambrosio, M. Colombo, G. De Philippis, A. Figalli. A global existence result for the semigeostrophic equations in three dimensional convex domains, Preprint arXiv:1205.5435, 2012.
- [4] L. Ambrosio, W. Gangbo. Hamiltonian ODE in the Wasserstein spaces of probability measures. *Comm. Pure Appl. Math.* **61** (2008), 18–53.
- [5] L. Ambrosio, N. Gigli and G. Savaré. Gradient flows in metric spaces and the Wasserstein spaces of probability measures. Lectures in Mathematics, ETH Zurich, Birkhäuser, 2005.
- [6] J.-D. Benamou, Y. Brenier. Weak existence for the Semi-Geostrophic equations formulated as a coupled Monge-Ampere/transport problem. *SIAM J. Appl. Math.*, **58** (1998), no. 5, 1450–1461.
- [7] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* **44** (1991), no 4, 375–417.
- [8] M. Cullen. *Private communication*.
- [9] M. Cullen and M. Feldman. Lagrangian solutions of Semi-Geostrophic equations in physical space. *SIAM J. Math. Anal.* **37** (2006), no 5, 1371–1395.
- [10] M. Cullen, W. Gangbo. A variational approach for the 2-dimensional semi-geostrophic shallow water equations. *Arch. Rat. Mech. Anal.* **156** (2001), 241–273.
- [11] M. Cullen, H. Maroofi. The fully compressible semi-geostrophic system from meteorology. *Arch. Ration. Mech. Anal.* **167** (2003), no. 4, 309–336.
- [12] M. Cullen, R.J. Purser. An extended Lagrangian theory of semi-geostrophic frontogenesis. *J. Atmos. Sci.*, **41** (1984), 1477–1497

- [13] L. C. Evans, R. F. Gariepy. Measure theory and fine properties of functions. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.
- [14] J.C. O Faria, M.C. Lopes Filho, H.J. Nussenzweig Lopes. Weak stability of Lagrangian solutions to the Semi-Geostrophic equations. *Nonlinearity* **22** (2009), 2521–2539.
- [15] M. Feldman and A. Tudorascu. On Lagrangian solutions for the Semi-Geostrophic system with singular initial data. *SIAM J. Math. Anal.*, **45** (2013), no. 3, 1616–1640.
- [16] M. Feldman and A. Tudorascu. On the Semi-Geostrophic system in physical space with general initial data. *submitted*, (2014), 1–24.
- [17] W. Gangbo, T. Nguyen and A. Tudorascu. Euler-Poisson systems as action-minimizing paths in the Wasserstein space. *Arch. Rat. Mech. Anal.* **192**, No. 3 (2009), 419–452.
- [18] B. Hoskins. The Geostrophic Momentum Approximation and the Semi-Geostrophic Equations. *J. Atmos. Sci.* **32**, No. 2 (1975), 233–242.
- [19] C. Villani. Topics in optimal transportation. Graduate Studies in Mathematics **58**, American Mathematical Society, 2003.