TWO STAGE ALGORITHM FOR SEMI-DISCRETE OPTIMAL TRANSPORT ON DISCONNECTED DOMAINS

MOHIT BANSIL

ABSTRACT. In this paper we present a two-stage algorithm to solve the semi-discrete optimal transport problem in the case where the support of the source measure is disconnected. We establish global linear convergence and local superlinear convergence. We also find convergence of the associated Laguerre cells both in measure and in the sense of Hausdorff distance.

CONTENTS

1. Introduction 1
2. Setup 2
2.1. Notations and Conventions 2
2.2. Previous Results 4
2.3. Outline of the Paper 5
3. Spectral Estimates 5
4. Convergence of Newton Algorithm 7
5. Convergence of Gradient Descent Algorithm 10
6. Two-Stage Algorithm 13
7. Convergence of Cells 14
8. Estimation of $d_\lambda$ 15
9. Acknowledgments 19
References 19

1. INTRODUCTION

We begin by recalling the semi-discrete optimal transport problem. Let $X \subset \mathbb{R}^n$, $n \geq 2$ be compact and $Y := \{y_i\}_{i=1}^N \subset \mathbb{R}^n$ a fixed collection of finite points, along with a cost function $c : X \times Y \to \mathbb{R}$. We also fix Borel probability measures $\mu, \nu$ with $\text{spt } \mu \subset X$ and $\text{spt } \nu \subset Y$. Furthermore $\mu$ is absolutely continuous with respect to Lebesgue measure and we denote its density by $\rho$. We set $\lambda = (\nu(y_1), \ldots, \nu(y_N))$.

We want to find a measurable mapping $T : X \to Y$ such that $T_\# \mu(E) := \mu(T^{-1}(E)) = \nu(E)$ for any measurable $E \subset Y$, and $T$ satisfies

$$\int_X c(x, T(x)) \, d\mu = \min_{T_\# \mu = \nu} \int_X c(x, \tilde{T}(x)) \, d\mu.$$  \hspace{1cm} (1.1)

In this paper our goal is to propose and show convergence of a two stage algorithm. The first stage is a regularized gradient descent which achieves global linear convergence with respect to the regularized problem. The second stage is a damped Newton algorithm similar to that of [KMT19] which has superlinear local convergence. The main difference between
our algorithm and that of [KMT19] is ours doesn’t require spt μ to be connected. Without a connected support the algorithm in [KMT19] may not converge at all.

Furthermore, we will show that the Laguerre cells, \(T^{-1}(y_i)\), for the approximate optimizers constructed by our algorithm converge both in the \(L^1(\mu)\) sense and in the sense of Hausdorff distance. This result is analogous to that of [BK19]. While the \(L^1(\mu)\) convergence follows directly from the results there, the Hausdorff convergence result in [BK19] heavily relies on the connectedness of spt \(\mu\).

2. Setup

2.1. Notations and Conventions. Here we collect some notations and conventions for the remainder of the paper. We fix positive integers \(N, M, n\) with \(N, n \geq 2\) and a collection \(Y := \{y_i\}_{i=1}^N \subset \mathbb{R}^n\). We will write \(1\) to denote the vector in \(\mathbb{R}^N\) whose components are all 1. For any vector \(V \in \mathbb{R}^k\), we will write its components as superscripts so \(V^i\) is the \(i\)-th component of \(V\). We reserve the notation \(||V||\) for the \(l^2\)-Euclidean norm, i.e. \(||V|| = \sqrt{\sum_{i=1}^k |V^i|^2}\). We use \(||V||_1, ||V||_\infty\) for the \(l^1\) and \(l^\infty\) Euclidean norms, i.e. \(||V||_1 = \sum_{i=1}^k |V^i|\) and \(||V||_\infty = \max_{i \in \{1, \ldots, k\}} |V^i|\). We use \(\mathcal{L}\) to denote \(n\)-dimensional Lebesgue measure and \(\mathcal{H}^k\) to denote \(k\)-dimensional Hausdorff measure. Also we use the notation

\[
\Lambda = \{\lambda \in \mathbb{R}^N : \lambda^i \in [0, 1], \sum_{i=1}^N \lambda^i = 1\}
\]

for the set of all admissible weight vectors.

We will assume that the cost function \(c\) satisfies the following standard conditions:

(Reg) \(c(\cdot, y_i) \in C^2(X), \forall i \in \{1, \ldots, N\}\),

(Twist) \(\nabla_x c(x, y_i) \neq \nabla_x c(x, y_k), \forall x \in X, i \neq k\).

We also assume the following condition, originally studied by Loeper in [Loc09].

**Definition 2.1.** The cost, \(c\), is said to satisfy Loeper’s condition if for each \(i \in \{1, \ldots, N\}\) there exists a convex set \(Y_i \subset \mathbb{R}^n\) and a \(C^2\) diffeomorphism \(\exp_i(\cdot) : Y_i \rightarrow X\) such that

(QC) \(\forall t \in \mathbb{R}, 1 \leq k, i \leq N, \{p \in Y_i \mid -c(\exp_i^t(p), y_k) + c(\exp_i^t(p), y_i) \leq t\}\) is convex.

See Remark 2.3 below for further discussion of these conditions.

We also say that a set \(X \subset \mathbb{R}^n\) is \(c\)-convex with respect to \(Y\) if \((\exp_i^{-1})^{-1}(X)\) is a convex set for every \(i \in \{1, \ldots, N\}\).

**Definition 2.2.** For any \(\psi \in \mathbb{R}^N\) and \(i \in \{1, \ldots, N\}\), we define the \(i\)th Laguerre cell associated to \(\psi\) as the set

\[
\text{Lag}_i(\psi) := \{x \in X \mid c(x, y_i) + \psi^i = \min_i c(x, y_i) + \psi^i\}.
\]

We also define the function \(G : \mathbb{R}^n \rightarrow \Lambda\) by

\[
G(\psi) := (G^1(\psi), \ldots, G^N(\psi)) = (\mu(\text{Lag}_1(\psi)), \ldots, \mu(\text{Lag}_N(\psi)));
\]

and denote for any \(\epsilon \geq 0\),

\[
\mathcal{K}^\epsilon := \{\psi \in \mathbb{R}^N \mid G^i(\psi) > \epsilon, \forall i \in \{1, \ldots, N\}\}.
\]
Remark 2.3. The above conditions ([Reg], [Twist], [QC]) are the same ones assumed in [KMT19] and [BK19]. As is also mentioned in those papers, ([Reg]) and ([Twist]) are standard conditions in optimal transport. Furthermore, ([QC]) holds if $Y$ is a finite set sampled from a continuous space, and $c$ is a $C^4$ cost function satisfying what is known as the Ma-Trudinger-Wang condition (first introduced in a strong form in [MTW05], and in [TW09] in a weaker form).

If $\mu$ is absolutely continuous with respect to Lebesgue measure (or even if $\mu$ doesn’t give mass to small sets), then ([Twist]) implies that the Laguerre cells are pairwise $\mu$-almost disjoint. In this case the generalized Brenier’s theorem [Vil09, Theorem 10.28], tells us that for any vector $\psi \in \mathbb{R}^N$, the map $T_\psi : X \rightarrow Y$ defined by $T_\psi(x) = y_i$ whenever $x \in \text{Lag}_i(\psi)$ is a minimizer in the optimal transport problem (1.1), where the source measure is $\mu$ and the target measure is defined by $\nu = \nu_{G(\psi)}$.

We define $C_\nabla := \sup_{x,y \in Y} \| \nabla_c c(x,y) \|$. Throughout the rest of the paper we will assume that $X$ is compact and $c$-convex with respect to $Y$ and that $\rho$, the density of $\mu$, is $\alpha$-Hölder continuous.

In order to obtain our convergence results we will need spt $\mu$ to be at least locally connected in a quantitative way. First we recall a definition.

Definition 2.4. A non-zero measure $\mu$ on a compact set $Z \subset \mathbb{R}^n$ satisfies a $(q,1)$-Poincaré-Wirtinger inequality for some $1 \leq q \leq \infty$ if there exists a constant $C_{pw} > 0$ such that for any $f \in C^1(Z)$,

$$\|f - \int_Z f \, d\mu\|_{L^q(\mu)} \leq C_{pw}\|\nabla f\|_{L^1(\mu)}.$$ 

For brevity, we will write this as “$\mu$ satisfies a $(q,1)$-PW inequality on $Z$”. If $\mu$ is defined on a larger set then $Z$ then the phrase “$\mu$ satisfies a $(q,1)$-PW inequality on $Z$” is understood to mean that $\mu(Z) > 0$ and the restriction of $\mu$ to $Z$ satisfies a $(q,1)$-PW inequality.

Remark 2.5. We remark that some version of all of our results hold with only $(1,1)$-PW inequalities but $q > 1$ will give better constants.

We now define a local version of the PW inequality that will serve as our quantitative measure of the local connectivity of spt $\mu$.

Definition 2.6. A probability measure $\mu$ on a compact set $X \subset \mathbb{R}^n$ satisfies a local $(q,1)$-Poincaré-Wirtinger inequality if there are compact $X_i \subset X$ which are pairwise disjoint and $\mu$-almost cover $X$, i.e. $\mu(X \setminus (X_1 \cup \cdots \cup X_M)) = 0$ such that $\mu$ satisfies a $(q,1)$-PW inequality on each $X_i$ with constant $\kappa_i$.

We define the local PW constant $\kappa$ to be the $q$-th power mean of the $\kappa_i$, i.e. $\kappa = (\frac{1}{M} \sum_{i=1}^{M} \kappa_i q)^{1/q}$. Note that $\kappa \leq \max_i \kappa_i$.

When we have $M = 1$ we recover the global connectivity assumption of [KMT19, Definition 1.3]. Throughout the paper we use the notation $\chi^i = \mu(X_i)$ and the vector $\chi = (\chi^1, \ldots, \chi^M) \in \mathbb{R}^M$.

In order to obtain convergence of our algorithm we require that optimal map $T$, “splits up” the $X_i$. In order to assure this we assume that the subsets sums of $\chi$ and $\lambda$ are separated. For any vector $V \in \mathbb{R}^k$ we introduce the notation $S_V$ for the set of subset sums of $V$, i.e.

$$S_V = \left\{ \sum_{a \in A} a : A \subset \{ V^1, \ldots, V^k \} \right\}$$
and the notation $S^\text{prop}_V$ for the set of proper subset sums of $V$, i.e.

$$S^\text{prop}_V = \left\{ \sum_{a \in A} a : A \subset \{V^1, \ldots, V^k\}, A \neq \emptyset \right\}.$$ 

With this notation, we say that the subset sums of $\chi$ and $\lambda$ are separated if $S^\text{prop}_\lambda \cap S^\text{prop}_\chi = \emptyset$.

**Remark 2.7.** Note that for any fixed $X$ with fixed decomposition $X_1, \ldots, X_M$, the collection of all $\lambda \in \Lambda$ such that our condition is not satisfied (i.e. $S^\text{prop}_\lambda \cap S^\text{prop}_\chi \neq \emptyset$) is a “small set” in the sense that it has Hausdorff dimension $n - 2$ whereas $\Lambda$ has Hausdorff dimension $n - 1$. In particular if $\lambda$ is randomly chosen from $\Lambda$ (with respect to any probability measure that is absolutely continuous with respect to $\mathcal{H}^{n-1}$), then with probability 1, we have $S^\text{prop}_\lambda \cap S^\text{prop}_\chi = \emptyset$. We defer a formal proof of this till Proposition 8.1.

For the remainder of the paper we use the notation $d_\lambda := \text{dist}(S^\text{prop}_\lambda, S^\text{prop}_\chi)$ for any $\lambda \in \Lambda$. Note that since $N > 1$, $S^\text{prop}_\lambda$ is never empty and so $d_\lambda$ is always defined. We defer a discussion of methods to compute and/or bound $d_\lambda$ to section 8.

It is well-known that the optimal transport problem has a dual problem with strong duality (we refer the reader to [Vil03, Chapter 1] for background),

$$\min_{T, \mu = \nu} \int_X c(x, T(x)) d\mu = \sup_{\psi \in \mathbb{R}^N} \int_X (\min_j c(x, y_j) + \psi^j) d\mu - \langle \psi, \lambda \rangle. \quad (2.1)$$

Furthermore given any dual optimizer $\psi$ the map $T_\psi$ from Remark 2.3 is an optimal transport map for the primal problem. Because of this we are able to optimize the finite dimensional dual problem instead of the infinite dimensional primal problem. We define Kantorovich’s functional as in [KMT19, Theorem 1.1]

$$\Phi(\psi) := \int_X (\min_i c(x, y_i) + \psi^i) d\mu - \langle \psi, \lambda \rangle = \sum_i \int_{\text{Lag}_i(\psi)} (c(x, y_i) + \psi^i) d\mu - \langle \psi, \lambda \rangle.$$

Our algorithm will find approximate maximizers of $\Phi$. We also recall that $\nabla \Phi(\psi) = G(\psi) - \lambda$.

### 2.2. Previous Results

There are many papers that apply a Newton-type algorithm to solve semi-discrete optimal transport problems and Monge-Ampère equations. In [OP88] the authors prove local convergence of a Newton Method for solving a semi-discrete Monge-Ampère equation. Global convergence is proved in [Mir15].

For the semi-discrete optimal transport problem the authors of [KMT19] give a Newton method similar to our Algorithm 1 in the case when the source measure is given by a continuous probability density and has connected support. An analogous result for the case when $\mu$ is supported on a union of simplexes is given in [MMT18], however this paper also requires a connectedness condition on the support of $\mu$.

In [BK19], Kitagawa and the author present a Newton method to solve a generalization of the semi-discrete optimal transport problem in which there is a storage fee. Our results do not require a connectedness condition on the support of $\mu$, however we require a strict convexity assumption on the storage fee function that doesn’t hold for the classical semi-discrete optimal transport problem. Furthermore we show that the Laguerre cells associated to the approximate transport maps converge quantitatively both in the sense of $\mu$-symmetric distance and in the sense of Hausdorff distance.
2.3. Outline of the Paper. In section 3 we obtain quantitative bounds on the strong concavity of \( \Phi \). In section 4 we exploit these bounds to obtain local convergence of a damped Newton algorithm. In section 5 we obtain global convergence of a gradient descent algorithm applied to a regularized version of \( \Phi \). In section 6 we discuss convergence of our two stage algorithm. In section 7 we obtain quantitative convergence of the associated Laguerre cells. In section 8 we describe some methods to compute and/or estimate the parameter \( d_\lambda \) from the initial data.

3. Spectral Estimates

In this section we work toward quantitative bounds on the strong concavity of \( \Phi \).

To start off we recall some notation in order to remain consistent with [KMT19].

Definition 3.1. We will write \( \text{int}(X_i) \) to denote the interior of the set \( X_i \). Given an absolutely continuous measure \( \mu = \rho dx \), where \( \rho \) is continuous, and a set \( A \subset X_i \) with Lipschitz boundary, we will write

\[
|\partial A|_{\rho, X_i} := \int_{\partial A \cap \text{int}(X_i)} \rho dH^{n-1}, \quad |A|_{\rho} := \mu(A).
\]

The next lemma is virtually identical to [BK19, Lemma 7.4]. We omit the proof.

Lemma 3.2. Suppose that \( \mu = \rho dx \) satisfies a \((q, 1)\)-PW inequality on \( X_i \). Then

\[
\inf_{A \subset X_i} \frac{|\partial A|_{\rho, X_i}}{\min(|A|_{\rho}, |X_i \setminus A|_{\rho})^{1/q}} \geq \frac{1}{2^{1/q} K_i},
\]

where the infimum is over \( A \subset \text{int}(X_i) \) whose boundary is Lipschitz with finite \( H^{n-1} \)-measure, and \( \min(|A|_{\rho}, |X_i \setminus A|_{\rho}) > 0 \).

Recall from [KMT19] that \( DG \) is negative semidefinite (as it is the Hessian of a concave function). Furthermore, recall that \( 1 \) is in the kernel of \( DG \), so the largest eigenvalue of \( DG \) is 0. We now work toward the following estimate on the second largest eigenvalue.

Theorem 3.3. Suppose that \( \mu \) satisfies a local \((q, 1)\)-PW inequality on \( X \) with constant \( \kappa \). Let \( \psi \in \mathbb{R}^N \) be fixed. Assume that the subset sums of \( G(\psi) \) and \( \chi \) are separated i.e., \( S_{G(\psi)}^{\text{prop}} \cap S_{\chi} = \emptyset \). Then the second largest eigenvalue of \( DG(\psi) \) is bounded away from zero. In particular it is bounded by

\[
-\frac{2^{3-1/q}}{C_{\psi} M^{1/q} N^{2} \kappa} d_{G(\psi)}^{1/q} < 0
\]

where \( d_{G(\psi)} = \text{dist}(S_{G(\psi)}^{\text{prop}}, S_{\chi}) > 0 \).

At this point, given some \( \psi \in \mathbb{R}^N \) we recall the construction of \( W \) from [KMT19, Section 5.3]. \( W \) is the simple weighted graph whose vertex set is the collection \( Y \), and for any \( y_i \) and \( y_j, \ i \neq j \) there is an edge between \( y_i, y_j \) of weight \( w_{ij} \) given by

\[
w_{ij} := D_i G^j(\psi) = D_j G^i(\psi) = \int_{\text{Lag}_{i,j}(\psi)} \frac{\rho(x)}{\|\nabla_x c(x, y_i) - \nabla_x c(x, y_j)\|} dH^{n-1}(x)
\]

where we have used the notation

\[
\text{Lag}_{i,j}(\psi) := \text{Lag}_i(\psi) \cap \text{Lag}_j(\psi)
\]

for \( i, j \in \{1, \ldots, N\} \).

Proposition 3.4. Under the conditions of Theorem 3.3, \( W \) is connected by edges of weight at least

\[
\frac{2^{1-1/q}}{C_{\psi} M^{1/q} N^{2} \kappa} d_{G(\psi)}^{1/q}.
\]
Proof. Suppose by contradiction that the proposition is false. This implies that removing all edges with weight strictly less than $\frac{2^{1-1/q}}{C \gamma N^{2K}} d_{G(\psi)}^{1/q}$ yields a disconnected graph. In other words, we can write $W = W_1 \cup W_2$ where $W_1, W_2 \neq \emptyset$ and are disjoint, such that every edge connecting a vertex in $W_1$ to a vertex in $W_2$ has weight strictly less than $\frac{2^{1-1/q}}{C \gamma N^{2K}} d_{G(\psi)}^{1/q}$. Letting $A = \cup_{y_i \in W_1} \text{Lag}_i(\psi)$ we see that

$$|\partial A|_{\rho, X} \leq \sum_{\{(i,j) | y_i \in W_1, y_j \in W_2\}} 2C \gamma w_{ij} < \frac{2^{2-1/q}}{M^{1/q} N^{2K}} d_{G(\psi)}^{1/q} |W_1| |W_2| \leq \frac{2^{2-1/q}}{M^{1/q} N^{2K}} d_{G(\psi)}^{1/q} \frac{N^2}{4} = \frac{2^{-1/q}}{M^{1/q} N^{2K}} d_{G(\psi)}^{1/q}.$$  

Now for each $i$ we claim that $\min(|A \cap X_i|_\rho, |X_i \setminus A|_\rho) < \frac{k_i^q}{M K^q} d_{G(\psi)}$. If $|A \cap X_i|_\rho = 0$ or $|X_i \setminus A|_\rho = 0$ then the claim is trivial. Otherwise the claim follows from Lemma 3.2, after noting that $|\partial A|_{\rho, X_i} \leq |\partial A|_{\rho, X}$. Now note that by construction

$$\mu(A) = \sum_{y_i \in W_1} \mu(\text{Lag}_i(\psi)) = \sum_{y_i \in W_1} G(\psi)^i \in S_{G(\psi)}.$$  

On the other hand we can partition the $X_i$ into two types. Let $I$ be the index set so that $i \in I$ if and only if $|X_i \setminus A|_\rho < \frac{k_i^q}{M K^q} d_{G(\psi)}$. We see that

$$\mu(A) = \sum_{i=1}^M \mu(A \cap X_i) = \sum_{i \in I} \mu(A \cap X_i) + \sum_{i \notin I} \mu(A \cap X_i) = \sum_{i \in I} \mu(X_i) - \mu(X_i \setminus A) + \sum_{i \notin I} \mu(A \cap X_i).$$  

Hence

$$\left| \mu(A) - \sum_{i \in I} \mu(X_i) \right| \leq \sum_{i \notin I} \mu(X_i \setminus A) + \sum_{i \notin I} \mu(A \cap X_i)$$

$$< \sum_{i \in I} \frac{k_i^q}{M K^q} d_{G(\psi)} + \sum_{i \notin I} \frac{k_i^q}{M K^q} d_{G(\psi)} = \frac{1}{M K^q} \sum_{i=1}^M k_i^q d_{G(\psi)} = d_{G(\psi)}.$$  

By definition $\sum_{i \in I} \mu(X_i) \in S_\chi$. Furthermore since both $W_1$ and $W_2$ are nonempty we have $\mu(A) \in S_{G(\psi)}^{\text{pro}}$, recalling (3.1). Hence we see that

$$d_{G(\psi)} = \text{dist}(S_{G(\psi)}^{\text{pro}}, S_\chi) < d_{G(\psi)}$$

which is a clear contradiction. \[\square\]

At this point Theorem 3.3 follows via the exact same method of proof as [BK19 Theorem 8.1]. We omit the proof.
4. Convergence of Newton Algorithm

First we recall the standard damped Newton Algorithm (similar to the one proposed in [KMT19]).

**Input:** A tolerance $\zeta > 0$ and an initial $\psi_0 \in \mathbb{R}^N$.

**while** $\|\nabla \Phi(\psi_k)\| \geq \zeta$ **do**

**Step 1:** Compute $\tilde{d}_k = -[D^2 \Phi(\psi_k)]^{-1}(\nabla \Phi(\psi_k))$

**Step 2:** Determine the minimum $\ell \in \mathbb{N}$ such that $\psi_{k+1,\ell} := \psi_k + 2^{-\ell} \tilde{d}_k$ satisfies

$$
\|\nabla \Phi(\psi_{k+1,\ell})\| \leq (1 - 2^{-\ell+1})\|\nabla \Phi(\psi_k)\|.
$$

**Step 3:** Set $\psi_{k+1} = \psi_k + 2^{-\ell} \tilde{d}_k$ and $k \leftarrow k + 1$.

end

**Algorithm 1:** Damped Newton’s algorithm

We remark that unlike the algorithm proposed in [KMT19] we do not impose a condition on the cells not collapsing. This is because we will prove local convergence for the above algorithm and within the zone of convergence the error reduction requirement will automatically imply that the cells don’t collapse.

Our main result for this section is that Algorithm 1 converges locally with superlinear rate.

**Theorem 4.1.** Suppose that $\mu$ satisfies a local $(q,1)$-PW inequality on $X$. If $\|\nabla \Phi(\psi_0)\| \leq \frac{d_\lambda}{2\sqrt{N}}$ then Algorithm 1 converges with linear rate and locally with rate $1 + \alpha$.

We start with a simple lemma.

**Lemma 4.2.** For $\lambda_1, \lambda_2 \in \Lambda$, we have $|d_{\lambda_1} - d_{\lambda_2}| \leq \|\lambda_1 - \lambda_2\|_1$.

**Proof.** Let $x \in S_{\lambda_1}^{prop}$. We claim that $\text{dist}(x, S_{\lambda_2}^{prop}) \leq \|\lambda_1 - \lambda_2\|_1$. Indeed if $x = \sum_{i \in I} \lambda_i^1$ then

$$
|x - \sum_{i \in I} \lambda_i^2| = \sum_{i \in I} (\lambda_i^1 - \lambda_i^2) \leq \sum_{i \in I} |\lambda_i^1 - \lambda_i^2| \leq \|\lambda_1 - \lambda_2\|_1.
$$

If we apply this to $x = \arg\min_{\tilde{x} \in S_{\lambda_1}^{prop}} \text{dist}(\tilde{x}, S_{\lambda})$ then we get that there is $y \in S_{\lambda_2}^{prop}$ so that $|x - y| \leq \|\lambda_1 - \lambda_2\|_1$. Hence

$$
\text{dist}(S_{\lambda_2}^{prop}, S_{\lambda}) \leq \text{dist}(y, S_{\lambda}) \leq |x - y| + \text{dist}(x, S_{\lambda}) = \text{dist}(S_{\lambda_1}^{prop}, S_{\lambda}) + |x - y| \leq \text{dist}(S_{\lambda_1}^{prop}, S_{\lambda}) + \|\lambda_1 - \lambda_2\|_1
$$

and so we get $d_{\lambda_2} - d_{\lambda_1} = \text{dist}(S_{\lambda_2}^{prop}, S_{\lambda}) - \text{dist}(S_{\lambda_1}^{prop}, S_{\lambda}) \leq \|\lambda_1 - \lambda_2\|_1$. A symmetric argument proves the opposite inequality.

With the above lemma, Theorem 3.3 gives us that $\Phi$ is locally well-conditioned near its maximum.

**Proposition 4.3.** Suppose that $\mu$ satisfies a local $(q,1)$-PW inequality on $X$ and $\eta < d_\lambda$. On the set $J^\eta = \{\psi \in \mathbb{R}^N : \|\nabla \Phi(\psi)\|_1 < \eta\}$ we have that $\Phi$ is uniformly $C^{2, \alpha}$ and strongly concave, except in the direction $1$. In particular on $J^\eta$

$$
\|\Phi\|_{C^{2, \alpha}} < C(d_\lambda - \eta)^{-2}
$$
and

\[ D^2\Phi \leq -\frac{2^{3-1/q}}{C_\gamma M^{1/q}N^4} (d_\lambda - \eta)^{1/q} P \]

where \( P \) is the orthogonal projection onto the hyperplane perpendicular to \( 1 \) and \( C \) is a constant depending only on \( \|\rho\|_\infty, \text{diam}(X) \), and \( c \) (the cost).

**Proof.** Fix \( \psi \in \mathcal{J}^n \). Since \( 0 \in S_\chi \) and each \( \lambda^i \in S_\lambda^{\text{prop}} \), we have that each \( \lambda^i \geq d_\lambda \). Furthermore, since

\[ |G^i(\psi) - \lambda^i| \leq \|\nabla \Phi(\psi)\|_\infty \leq \|\nabla \Phi(\psi)\|_1 < \eta \]

we obtain that \( G^i(\psi) - \lambda^i > -\eta \) and so \( G^i(\psi) > d_\lambda - \eta \). In other words

(4.1)

\[ \mathcal{J}^n \subset K^{d_\lambda - \eta}. \]

Hence the first result follows by careful tracing through the proof of [KMT19, Theorem 4.1]. Now for the second claim Lemma 4.2 gives that

\[ |d_{G(\psi)} - d_\lambda| \leq \|G(\psi) - \lambda\|_1 = \|\nabla \Phi(\psi)\|_1 \leq \eta. \]

In particular \( d_{G(\psi)} \geq d_\lambda - \eta \). The second result now follows from Theorem 3.3.

We briefly note that this proposition gives us uniqueness of the maximizer of \( \Phi \) up to a multiple of \( 1 \).

**Corollary 4.4.** If \( \psi_1, \psi_2 \) are two maximizers of \( \Phi \) then there is \( r \in \mathbb{R} \) so that \( \psi_1 - \psi_2 = r1 \).

**Proof.** Since \( \Phi \) is concave and \( \psi_1, \psi_2 \) are two maximizers, \( \Phi \) must be constant along the line segment joining \( \psi_1 \) and \( \psi_2 \). In particular if \( t \in [0, 1] \) then \( \Phi(t\psi_1 + (1-t)\psi_2) = \Phi(\psi_1) \) and so \( t\psi_1 + (1-t)\psi_2 \) is a maximizer of \( \Phi \). Hence \( \nabla \Phi(t\psi_1 + (1-t)\psi_2) = 0 \) for all \( t \in [0, 1] \). In particular we can conclude \([D^2\Phi(\psi_1)](\psi_2 - \psi_1) = 0 \).

Now since \( \psi_1 \) is a maximizer we have \( \nabla \Phi(\psi_1) = 0 \) so the conditions of Proposition 4.3 are satisfied. Hence we have \( D^2\Phi(\psi_1) \leq -\frac{2^{3-1/q}}{C_\gamma M^{1/q}N^4} (d_\lambda - \eta)^{1/q} P \) where \( P \) is the orthogonal projection onto the hyperplane perpendicular to \( 1 \) and we get \( \|P(\psi_2 - \psi_1)\| = 0 \) which implies the claim.

**Remark 4.5.** Given any maximizer, \( \psi \), of \( \Phi \) we see that \( \psi + r1 \) is also a maximizer. In particular given any maximizer, \( \psi \), we can construct the maximizer \( \psi_{\max} := \psi - (\sum_{i=1}^{N} \psi^i)1 \) which is a maximizer of \( \Phi \) that satisfies \( \sum_{i=1}^{N} \psi_{\max}^i = 0 \).

Conversely 4.3 tells us that there is a unique \( \psi_{\max} \) that maximizes \( \Phi \) and satisfies \( \sum_{i=1}^{N} \psi_{\max}^i = 0 \). For the remainder of the paper we shall refer to this unique maximizer as \( \psi_{\max} \).

It is well-known that Proposition 4.3 gives local \( 1 + \alpha \) convergence of the standard damped Newton Algorithm. However for convenience of the reader we will give a self contained proof.

**Proposition 4.6.** If \( \|\nabla \Phi(\psi_0)\| < \frac{d_\lambda}{2\sqrt{N}} \) then the iterates of Algorithm 1 satisfy

\[ \|\nabla \Phi(\psi_{k+1})\| \leq (1 - \frac{\tau_k}{2}) \|\nabla \Phi(\psi_k)\| \]
where
\[ \tau_k = \min\left( \frac{\beta^{1+\frac{1}{2}}\delta}{L^2 \| \nabla \Phi(\psi) \| 2^{\frac{1}{2}}} , 1 \right) \]
and
\[ \delta := \frac{1}{2} \min \lambda^2 \]
\[ L \leq C\delta^{-2} \]
\[ \beta \geq \frac{2^{3-2/q}}{C\sqrt{M^{1/q}N^4K}} d_\lambda^{1/q} \]

where \( C \) is a constant depending only on \( \| \rho \|_\infty \), \( \text{diam}(X) \), and \( c \) (the cost).
Furthermore once we have \( \tau_k = 1 \) we get
\[ \| \nabla \Phi(\psi_{k+1}) \| \leq \frac{L\| \nabla \Phi(\psi) \|^{1+\alpha}}{\beta^{1+\alpha}}. \]

**Proof.** The proof is very similar to that of [KMT19, Proposition 6.1]. However we will give all the details.

Note that \( \| \nabla \Phi(\psi_0) \| < \frac{d\lambda}{2\sqrt{N}} \) implies \( \| \nabla \Phi(\psi_0) \|_1 < \frac{d\lambda}{2} \).

For this proof define \( \epsilon := \frac{d\lambda}{2}, \delta := \frac{1}{2} \min \lambda^2 \). Note \( \epsilon \leq \delta \) and so, by (1.1), \( J^\epsilon \subset K^\epsilon \subset K^\delta \) (Note that it is NOT true in general that \( J^\eta \subset K^\eta \); our choice of \( \epsilon \) is the largest choice for which this is guaranteed to work). Furthermore \( L \) is the \( C^{1,\alpha} \) norm of \( G \) on \( K^{\delta/2} \) and \( -\beta \) is the bound of the second largest eigenvalue of \( DG \) on \( J^\epsilon \). By Proposition 4.3 we have
\[ \beta \geq \frac{2^{3-2/q}}{C\sqrt{M^{1/q}N^4K}} d_\lambda^{1/q} \cdot \]

Also by carefully tracing through the proofs in [KMT19] we have \( L \leq C\delta^{-2} \).
We analyze a single iteration of the Algorithm. Define \( \psi := \psi_k \in J^\epsilon \subset K^\epsilon \subset K^\delta \).
Let \( v := [D^2 \Phi(\psi)]^+(\nabla \Phi(\psi)) \). We see that
\[ \| v \| \leq \frac{\| \nabla \Phi(\psi) \|}{\beta}, \]
as \( \Phi \) is \( \beta \)-concave in the direction orthogonal to \( 1 \).

Also define \( \psi_\tau = \psi - \tau v \). Let \( \tau_1 \) be the first exit time from \( K^{\delta/2} \). We have that
\[ \frac{\delta}{2} \leq \| G(\psi_{\tau_1}) - G(\psi) \| \leq L\tau_1 \| v \| \leq \frac{L}{\beta} \| \nabla \Phi(\psi) \| \tau_1 \]
and so
\[ \tau_1 \geq \frac{\beta \delta}{2L \| \nabla \Phi(\psi) \|}. \]

Applying Taylor’s formula to \( \nabla \Phi \) we get
\[ (4.2) \quad \nabla \Phi(\psi_\tau) = \nabla \Phi(\psi - \tau v) = \nabla \Phi(\psi) - \tau (D \nabla \Phi(\psi)) v + R(\tau) = \nabla \Phi(\psi) - \tau \nabla \Phi(\psi) + R(\tau) \]
where

\[ \| R(\tau) \| = \| \int_0^\tau (D\nabla \Phi(\psi) - D\nabla \Phi(\psi))v d\sigma \| \]
\[ = \| \int_0^\tau D\Phi(\psi)v - D\Phi(\psi)v d\sigma \| \]
\[ \leq \int_0^\tau L\| \psi_\sigma - \psi \| \alpha \| v \| d\sigma \]
\[ = \int_0^\tau L\| \sigma v \| \alpha \| v \| d\sigma \]
\[ = L\| v \|^{\alpha + 1} \int_0^\tau \sigma^\alpha d\sigma \]
\[ = L\| v \|^{\alpha + 1} \frac{\tau^{\alpha + 1}}{\alpha + 1} \]
\[ \leq \frac{L\| \nabla \Phi(\psi) \|^2}{\beta^{1+\alpha}} \tau^{1+\alpha} \]

(4.3)

for \( \tau \leq 1 \).

Finally we establish the error reduction estimates. (4.2) gives

\[ \nabla \Phi(\psi_\tau) = (1 - \tau)\nabla \Phi(\psi) + R(\tau) \]

so we have

\[ \| \nabla \Phi(\psi_\tau) \| \leq (1 - \frac{\tau}{2})\| \nabla \Phi(\psi) \| \]

provided that

\[ \| R(\tau) \| \leq \frac{\tau}{2} \| \nabla \Phi(\psi) \|. \]

Again using (4.3) this will be true provided

\[ \tau \leq \min(\tau_1, \frac{\beta^{1+\frac{1}{\alpha}}}{L^{\frac{1}{\alpha}}\| \nabla \Phi(\psi) \|^2}) \]

\( =: \tau_2 \).

Hence we see that if we set \( \tau_k := \tau_2 \), then the claim is true. Furthermore as the error goes to zero, eventually we must have \( \tau_k = 1 \). When this happens (4.2) gives

\[ \nabla \Phi(\psi_1) = R(1) \]

and so we get the super-linear convergence.

\[ \square \]

5. CONVERGENCE OF GRADIENT DESCENT ALGORITHM

In order to remedy the fact that the above algorithm only has local convergence we propose a regularized version in order to get within a close enough error.

We define the regularized Kantorovich’s functional

\[ \Phi(\psi) = \Phi(\psi) - \frac{\gamma}{2} \| \psi \|^2 \]

where \( \gamma > 0 \) is some parameter.
We see that
\[ \nabla \tilde{\Phi}(\psi) = \nabla \Phi(\psi) - \gamma \psi = G(\psi) - \lambda - \gamma \psi \]
and
\[ D^2 \tilde{\Phi}(\psi) = D^2 \Phi(\psi) - \gamma I = DG(\psi) - \gamma I. \]

In particular we note that \( \tilde{\Phi}(\psi) \) is strongly \( \gamma \)-concave. If we let \( C_L \) be the Lipschitz constant of \( G \) \((C_L = C N \) for some universal constant \( C \)) then \( \tilde{\Phi} \) has Lipschitz gradient with constant \( C_L + \gamma \). Hence \( \tilde{\Phi} \) is well-conditioned.

**Lemma 5.1.** The projection of \( K^0 \) in the direction \( \mathbf{1} \) is bounded by
\[
M_0 := \sqrt{N} \left( \max_{x \in X} (\max_{y \in Y} c(x, y) - \min_{y \in Y} c(x, y)) \right).
\]
In other words if \( \psi \in K^0 \) and \( \sum_i \psi^i = 0 \) then \( \|\psi\| \leq M_0 \).

**Proof.** Since \( \sum_i \psi^i = 0 \) there are some indices \( j_1, j_2 \in \{1, \ldots, N\} \) so that \( \psi^{j_1} \leq 0 \) and \( \psi^{j_2} \geq 0 \). Suppose for sake of contradiction that \( \|\psi\| > M_0 \). Then we must have \( \|\psi\|_\infty > M_0 \).

We first look at the case where \( \psi^j > \max_{x \in X} (\max_{y \in Y} c(x, y) - \min_{y \in Y} c(x, y)) \). By the fundamental lemma \( \|\psi\|_\infty > M_0 \). We wish to show that \( \|\psi\|_\infty \leq M_0 \).

We remark that since \( X \) is compact and \( c \) is continuous, \( M_0 < +\infty \) and so the bound isn’t vacuous.

Since \( \Phi \) is strongly concave it has a unique maximizer. Although it may not be true that this maximizer is in \( K^0 \) we show that it is still bounded.

**Lemma 5.2.** Let \( \psi^* \) be the unique maximizer of \( \Phi \). Then \( \|\psi^*\| \leq M_0 \).

**Proof.** Let \( \psi_{\max} \) be the maximizer of \( \Phi \) constructed in Remark 4.5 so that \( \sum_i \psi^j \leq 0 \).

Then \( G(\psi_{\max}) = \lambda \) and so \( \psi_{\max} \in K^0 \). In particular \( \|\psi_{\max}\| \leq M_0 \). But since \( \psi^* \) is the maximizer of \( \Phi \) we have \( \Phi(\psi_{\max}) \leq \Phi(\psi^*) \). Hence:

\[
\Phi(\psi_{\max}) - \gamma \|\psi_{\max}\|^2 \leq \Phi(\psi^*) - \gamma \|\psi^*\|^2 \\
-\gamma \|\psi_{\max}\|^2 \leq -\gamma \|\psi^*\|^2 \\
\|\psi_{\max}\|^2 \geq \|\psi^*\|^2
\]

and so the claim follows.

\( \square \)
We now recall the standard gradient descent algorithm.

**Parameters:** A regularization constant \( \gamma > 0 \) and a step size \( h > 0 \).

**Input:** A tolerance \( \zeta > 0 \) and an initial \( \psi_0 \in \mathbb{R}^N \) such that \( \sum \psi_i = 0 \).

while \( \| \nabla \tilde{\Phi}(\psi_k) \| \geq \zeta \) do

1. Compute \( \tilde{d}_k = -\nabla \tilde{\Phi}(\psi_k) \)
2. Set \( \psi_{k+1} = \psi_k + h\tilde{d}_k \) and \( k \leftarrow k + 1 \).

end

**Algorithm 2:** Regularized Gradient Descent

It is well known that Algorithm 2 converges with linear rate as \( \tilde{\Phi} \) is well conditioned. We recall the convergence rate.

**Proposition 5.3.** Let \( \psi^* \) be the unique maximizer of \( \tilde{\Phi} \). Suppose that \( \gamma < \min \lambda_i M_0 \). If we set \( h = \frac{2}{C_L + 2\gamma} \) then the iterates of Algorithm 2 satisfy

\[
\| \psi_k - \psi^* \| \leq \left( \frac{Q_f - 1}{Q_f + 1} \right)^k (M_0 + \| \psi_0 \|)
\]

and

\[
\| \nabla \Phi(\psi_k) \| \leq (C_L + \gamma) \left( \frac{Q_f - 1}{Q_f + 1} \right)^k (M_0 + \| \psi_0 \|)
\]

where \( Q_f = \frac{C_L + \gamma}{\gamma} \) is the condition number and \( C_L \) is the Lipschitz constant of \( G \).

**Proof.** From the above lemma \( \| \psi^* \| \leq M_0 \) hence \( \| \psi_0 - \psi^* \| \leq \| \psi_0 \| + \| \psi^* \| \leq M_0 + \| \psi_0 \| \). The first claim now follows directly from [Nes18, Theorem 2.1.15]. For the second claim since \( \nabla \tilde{\Phi} \) is Lipschitz we have

\[
\| \nabla \tilde{\Phi}(\psi_k) - \nabla \tilde{\Phi}(\psi^*) \| \leq (C_L + \gamma) \| \psi_k - \psi^* \| \leq (C_L + \gamma) \left( \frac{Q_f - 1}{Q_f + 1} \right)^k (M_0 + \| \psi_0 \|)
\]

and the result follows from recalling that \( \nabla \tilde{\Phi}(\psi^*) = 0 \) as \( \psi^* \) is a maximizer.

\[\square\]

Using the convergence rate we give a quick estimate of how many iterations it takes our algorithm to converge.

**Corollary 5.4.** With tolerance \( \zeta \), parameter \( \gamma \), starting input \( \psi_0 \), and step size \( h = \frac{2}{C_L + 2\gamma} \), Algorithm 2 terminates in at most \( \log_b \left( \frac{M_0 + \| \psi_0 \|}{(M_0 + \| \psi_0 \|)(C_L + \gamma)} \right) \) iterations where \( b = \frac{C_L}{C_L + 2\gamma} \).

**Proof.** This follows from the second inequality in the above prop after noting that \( b = \frac{Q_f - 1}{Q_f + 1} \).

\[\square\]

We conclude this section with some quantitative stability bounds that estimate how much error our regularization introduced.

**Proposition 5.5.** Let \( \tilde{\psi} \) be the result of running Algorithm 2 with tolerance \( \zeta \), parameter \( \gamma \), and starting input \( \psi_0 \). Then

\[
\| G(\tilde{\psi}) - \lambda \| < \zeta + \gamma (2M_0 + \| \psi_0 \|)
\]

where \( M_0 \) is as defined above.
Proof. We have
\[ G(\tilde{\psi}) - \lambda = \nabla \tilde{\Phi}(\tilde{\psi}) + \gamma \tilde{\psi} = \nabla \tilde{\Phi}(\tilde{\psi}) + \gamma (\tilde{\psi} - \psi^* + \psi^*), \]
where $\psi^*$ is the unique maximizer of $\tilde{\Phi}$. By Proposition 5.3 we have $\|\tilde{\psi} - \psi^*\| \leq M_0 + \|\psi_0\|$ and by Lemma 5.2 we have $\|\psi^*\| \leq M_0$. Hence we get
\[ \|G(\tilde{\psi}) - \lambda\| \leq \|\nabla \tilde{\Phi}(\tilde{\psi})\| + \gamma (\|\tilde{\psi} - \psi^*\| + \|\psi^*\|) < \zeta + \gamma (2M_0 + \|\psi_0\|) \]
as desired. \qed

6. Two-Stage Algorithm

We now propose a two-stage algorithm.

**Input:** A tolerance $\tilde{\zeta} > 0$.

**Stage 1:** Run Algorithm 2 with tolerance $\zeta = \frac{d_\lambda}{4\sqrt{N}}$, parameter $\gamma = \frac{d_\lambda}{8M_0\sqrt{N}}$, starting input $\psi_0 = 0$, and step size $h = \frac{2}{C_L + 2\gamma}$. Let $\tilde{\psi}$ be the output.

**Stage 2:** Run Algorithm 1 with parameters $\zeta = \tilde{\zeta}$ and $\psi_0 = \tilde{\psi}$.

**Algorithm 3:** Two Stage Algorithm

**Proposition 6.1.** Stage 1 of Algorithm 3 terminates in at most $\frac{32NM_0^2C_L^2}{d_\lambda}$ steps (note that this is a constant independent of $\tilde{\zeta}$).

Let $E = \frac{\beta^{1+\frac{1}{2}\gamma}}{L + 2\gamma}$. If $\tilde{\zeta} \geq E$ then Stage 2 terminates in at most
\[ \max \left( \frac{2\left(\frac{d_\lambda}{4\sqrt{N}} - \tilde{\zeta}\right)}{E}, 0 \right) \]
steps.

Otherwise Stage 2 terminates in at most
\[ \max \left( \frac{2\left(\frac{d_\lambda}{4\sqrt{N}} - E\right)}{E}, 0 \right) + \log_{1+\alpha} \frac{\log \frac{L^{\frac{1}{\alpha}}}{\beta^{1+\frac{1}{2}\gamma}}}{\log \frac{L^{\frac{1}{\alpha}}}{\beta^{1+\frac{1}{2}\gamma}} E} \]
steps.

**Proof.** For the first claim we have using Corollary 5.4 that the maximum number of iterations is
\[ \log \frac{\tilde{\zeta}}{M_0(C_L + \gamma)} = \log \frac{\frac{d_\lambda}{4\sqrt{N}}}{M_0(C_L + \frac{d_\lambda}{8M_0\sqrt{N}})} = \log \frac{M_0(C_L + \frac{d_\lambda}{8M_0\sqrt{N}})}{\frac{d_\lambda}{4\sqrt{N}}} = \log \frac{C_L}{C_L + 2\frac{d_\lambda}{8M_0\sqrt{N}}} = \log \frac{\frac{d_\lambda}{C_L + 2\frac{d_\lambda}{8M_0\sqrt{N}}}}{\frac{d_\lambda}{4\sqrt{N}}} = \log \frac{\frac{1}{2} + \frac{4\sqrt{N}M_0C_L}{d_\lambda}}{\log (1 + \frac{\frac{d_\lambda}{4\sqrt{N}M_0C_L}}{d_\lambda})}. \]

Now using the classical inequalities, $\log(1 + x) \geq \frac{x}{1 + x}$ and $\log(\frac{1}{2} + x) \leq \log(1 + x) \leq x$ we get
\[ \frac{\log(\frac{1}{2} + \frac{4\sqrt{N}M_0C_L}{d_\lambda})}{\log(1 + \frac{\frac{d_\lambda}{4\sqrt{N}M_0C_L}}{d_\lambda})} \leq \frac{4\sqrt{N}M_0C_L}{d_\lambda} \leq \frac{4\sqrt{N}M_0C_L}{d_\lambda} = \frac{32NM_0^2C_L^2}{d_\lambda^2}, \]
where we have used $M_0, C_L \geq 1$ and $d_\lambda \leq 1$ for the last inequality.
Next by Proposition 5.5 we see that
\[ \| \nabla \Phi(\tilde{\psi}) \| = \| G(\tilde{\psi}) - \lambda \| < \frac{d_\lambda}{4\sqrt{N}} + \frac{d_\lambda}{8M_0\sqrt{N}} (2M_0 + \| \psi_0 \|) = \frac{d_\lambda}{2\sqrt{N}} \]
where \( \tilde{\psi} \) is the initial input for Stage 2 in Algorithm 3. Hence the result of Proposition 4.6 applies. In particular as long as \( \frac{E}{\| \nabla \Phi(\psi_k) \|} \leq 1 \) we have
\[ \| \nabla \Phi(\psi_{k+1}) \| \leq (1 - \frac{\tau_k}{2}) \| \nabla \Phi(\psi_k) \| = (1 - \frac{E}{2\| \nabla \Phi(\psi_k) \|}) \| \nabla \Phi(\psi_k) \| = \| \nabla \Phi(\psi_k) \| - \frac{E}{2} \]
which shows the second claim. Finally once we have \( \| \nabla \Phi(\psi_k) \| < E \), we see that \( \tau_k = 1 \) and so by Proposition 4.6 we get
\[ \| \nabla \Phi(\psi_{k+1}) \| \leq \frac{L \| \nabla \Phi(\psi_k) \|^{1+\alpha}}{\beta^{1+\alpha}} \]
and so
\[ \| \nabla \Phi(\psi_{k+j}) \| \leq \left( \frac{L}{\beta^{1+\alpha}} \right)^{(1+\alpha)^j - 1} \| \nabla \Phi(\psi_k) \|^{(1+\alpha)^j} \leq \left( \frac{L}{\beta^{1+\alpha}} \right)^{(1+\alpha)^j - 1} E^{(1+\alpha)^j}. \]
Hence if
\[ j \geq \log_{1+\alpha} \frac{\log \left( \frac{L}{\beta^{1+\alpha}} \right)}{\log \left( \frac{L}{\beta^{1+\alpha}} E \right)} \]
then we will have \( \| \nabla \Phi(\psi_{k+j}) \| < \tilde{\zeta} \) as desired. This completes the proof of the third and final claim.

\[ \square \]

7. Convergence of Cells

In this section we obtain quantitative convergence of the Laguerre cells associated with our problem. We recall \( \psi_{\text{max}} \), which is the maximizer of \( \Phi \) constructed in Remark 4.5 so that \( \sum \psi_{\text{max}} = 0 \).

The result on the \( \mu \)-symmetric convergence of the cells follows immediately from [BK19].

**Proposition 7.1.**
\[ \sum_{i=1}^{N} \Delta_{\mu}(\text{Lag}_i(\psi), \text{Lag}_i(\psi_{\text{max}})) \leq 4N \| \nabla \Phi(\psi) \|_1 \]

**Proof.** This is an immediate consequence of [BK19, Corollary 5.7] after applying Remark 5.8 there.

\[ \square \]

The Hausdorff convergence result in [BK19] strongly depends on the global PW inequality. Hence we have to do more work in our case to recover the result.

**Theorem 7.2.** Let \( \psi \in \mathbb{R}^N \) be so that \( \langle \psi, 1 \rangle = 0 \). Suppose that
\[ \| \nabla \Phi(\psi) \| < \min \left( \frac{2^{3-2/q}d_\lambda^{1/q} \delta}{\| \rho \|_\infty C_\Delta C_\nabla M^{1/q} N^5 \kappa}, \frac{d_\lambda}{2\sqrt{N}} \right) \]
where \( \delta = \frac{1}{2} \min \lambda^i \). Then
\[
\| \psi_{\text{max}} - \psi \| \leq \frac{C_\nabla M^{1/q} N^4 \kappa}{2^{3-2/q} d_{\lambda}^{1/q}} \| \nabla \Phi(\psi) \|
\]
and
\[
d_H(\text{Lag}_i(\psi_{\text{max}}), \text{Lag}_i(\psi))^n \leq \frac{C_1 C_\nabla M^{1/q} N^5 \kappa}{2^{3-2/q} d_{\lambda}^{1/q} (\arccos(1 - C_2 \mathcal{L}(\text{Lag}_i(\psi_{\text{max}}))^2))^{n-1}} \| \nabla \Phi(\psi) \|
\]
where \( C_1, C_2 \) are the constants described in \([BK19, \text{Theorem 6.6}]\).

Proof. Setting \( \epsilon := \frac{d_{\lambda}}{2} \) we see that \( \| \nabla \Phi(\psi) \| < \frac{d_{\lambda}}{2} \) implies that \( \psi \in \mathcal{J}^\epsilon \), where \( \mathcal{J}^\epsilon \) is as in Proposition 1.3. We recall by Proposition 1.3 that \( \Phi \) is strongly concave on \( \mathcal{J}^\epsilon \) except in the direction 1 with parameter \( C_\nabla M^{1/q} N^4 \kappa \). Since we have that \( \psi \) and \( \psi_{\text{max}} \) are normal to 1 we get
\[
\langle \nabla \Phi(\psi), \psi - \psi_{\text{max}} \rangle = \int_0^1 \langle D^2 \Phi((1-t)\psi_{\text{max}} + t\psi)(\psi - \psi_{\text{max}}), \psi - \psi_{\text{max}} \rangle dt
\]
\[
\leq \frac{-2^{3-2/q} d_{\lambda}^{1/q}}{C_\nabla M^{1/q} N^4 \kappa} \| \psi - \psi_{\text{max}} \|^2.
\]
From the Cauchy-Schwarz inequality we obtain \( \| \nabla \Phi(\psi) \| \geq \frac{2^{3-2/q} d_{\lambda}^{1/q}}{C_\nabla M^{1/q} N^4 \kappa} \| \psi - \psi_{\text{max}} \| \). Hence
\[
\| \psi - \psi_{\text{max}} \| \leq \frac{C_\nabla M^{1/q} N^4 \kappa}{2^{3-2/q} d_{\lambda}^{1/q}} \| \nabla \Phi(\psi) \|. \quad \text{Now using the assumed bound on } \| \nabla \Phi(\psi) \| \text{ we get}
\]
\[
\| \psi - \psi_{\text{max}} \| \leq \| \psi - \psi_{\text{max}} \| < \frac{C_\nabla M^{1/q} N^4 \kappa}{2^{3-2/q} d_{\lambda}^{1/q} (\| \rho \|_\infty C_\Delta C_\nabla M^{1/q} N^5 \kappa)} \| \| \rho \|_\infty C_\Delta N\|
\]
Since \( \lambda^i \leq \| \rho \|_\infty \mathcal{L}(\text{Lag}_i(\psi_{\text{max}})) \) we get \( \| \psi - \psi_{\text{max}} \| \leq \frac{\text{min}_i \mathcal{L}(\text{Lag}_i(\psi_{\text{max}}))}{2 C_\Delta N} \). Hence the conditions of \([BK19, \text{Theorem 6.6}]\) are satisfied and we get
\[
d_H(\text{Lag}_i(\psi_{\text{max}}), \text{Lag}_i(\psi))^n \leq \frac{C_1 N \| \psi_{\text{max}} - \psi \|_\infty}{(\arccos(1 - C_2 \mathcal{L}(\text{Lag}_i(\psi_{\text{max}}))^2))^{n-1}} \leq \frac{C_1 C_\nabla M^{1/q} N^5 \kappa}{2^{3-2/q} d_{\lambda}^{1/q} (\arccos(1 - C_2 \mathcal{L}(\text{Lag}_i(\psi_{\text{max}}))^2))^{n-1}} \| \nabla \Phi(\psi) \|
\]
as desired. \( \square \)

Remark 7.3. This shows that Algorithm 3 has local superlinear convergence even when the error is measured by the Hausdorff distances between the cells. Furthermore, the size of the zone of local convergence is explicitly determined from the initial data.

8. Estimation of \( d_{\lambda} \)

In this section we give some methods to estimate or compute \( d_{\lambda} \). Recall our notation, \( \chi^j = \mu(X_j) \).
First we prove the claim made in Remark 2.7.
Proposition 8.1. Let $X_1, \ldots, X_M$ be any fixed decomposition of $X$. Let $B := \{ \lambda \in \Lambda : d^i = 0 \} = \{ \lambda \in \Lambda : S^\lambda \cap S_{\chi} \neq \emptyset \}$. Then $\dim B = N - 2$ where dimension is understood in the sense of Hausdorff dimension.

Proof. Note that since $\chi$ is a vector with a finite number of entries, $S_{\chi}$ is a finite set. Say $S_{\chi} = \{ s_1, \ldots, s_l \}$. We denote the set of all length $N$ vectors in $\mathbb{R}^N$ whose entries are all 0 or 1 by $\mathbb{F}_2^N$. We let $\mathbb{F}_2^{N,\text{prop}} = \mathbb{F}_2^N \setminus \{(0, \ldots, 0), (1, \ldots, 1)\}$.

Now for any vector $v \in \mathbb{F}_2^N$ and $i \in \{1, \ldots, l\}$ define $A_{v,i} = \{ \lambda \in \mathbb{R}^N : \langle v, \lambda \rangle = s_i \}$. We see that

$$B = \bigcup_{(v,i) \in \mathbb{F}_2^{N,\text{prop}} \times \{1, \ldots, l\}} A_{v,i} \cap \Lambda$$

as for any $v \in \mathbb{F}_2^N$,

$$A_{v,i} \cap \Lambda = \{ \lambda \in \Lambda : \sum_{\{j: v^j = 1\}} \lambda^j = s_i \}.$$

Note that as long as $v \neq (0, \ldots, 0)$, $A_{v,i}$ is a hyperplane normal to $v$. Since $\Lambda$ is contained in a hyperplane normal to $(1, \ldots, 1)$ we see that for every $v \in \mathbb{F}_2^{N,\text{prop}}$, $A_{v,i}$ is a hyperplane transversal to $\Lambda$ (as $v \neq (1, \ldots, 1)$) and hence $A_{v,i} \cap \Lambda$ has dimension at most $N - 2$. Since $B$ is a finite union of sets of this form we see that $B$ has dimension at most $N - 2$. Note that since $B$ always contains $\{ \lambda \in \Lambda : \lambda^1 = 0 \}$, $\dim B \geq N - 2$.

Next we examine the case where $\lambda^i$ and $\chi^j$ are all rational numbers and the denominators of $\lambda^i$ are relatively prime to those of $\chi^j$. Note that the probability of two large integers being relatively prime is $\frac{\pi^2}{6} \approx 60\%$.

Proposition 8.2. Suppose that there are relatively prime positive integers $\bar{N}, \bar{M}$ so that $\bar{N}\lambda^i, \bar{M}\chi^j \in \mathbb{N}$. Then $d^\lambda \geq \frac{1}{\bar{M}\bar{N}}$.

Proof. Let $I, J$ be the minimizing index sets so that

$$d^\lambda = \left| \sum_{i \in I} \lambda^i - \sum_{j \in J} \chi^j \right|$$

where $I \neq \emptyset, \{1, \ldots, N\}$.

Then

$$\bar{M}\bar{N}d^\lambda = \left| \bar{M} \sum_{i \in I} \bar{N}\lambda^i - \bar{N} \sum_{j \in J} \bar{M}\chi^j \right|.$$

We claim that $\bar{M} \sum_{i \in I} \bar{N}\lambda^i \neq \bar{N} \sum_{j \in J} \bar{M}\chi^j$. If it did then we would have that $\bar{N}$ divides $\bar{M} \sum_{i \in I} \bar{N}\lambda^i$. Since $\bar{M}, \bar{N}$ are relatively prime this gives us that $\bar{N}$ divides $\sum_{i \in I} \bar{N}\lambda^i$. In particular either $\sum_{i \in I} \bar{N}\lambda^i = 0$ or $\sum_{i \in I} \bar{N}\lambda^i \geq \bar{N}$ which would imply either $\sum_{i \in I} \lambda^i = 0$ or $\sum_{i \in I} \lambda^i \geq 1$ respectively. The first case gives $I = \emptyset$ and the second gives $I = \{1, \ldots, N\}$ both of which are contradictions.
Now since we have $\tilde{M} \sum_{i \in I} \tilde{N} \lambda^i \neq \tilde{N} \sum_{j \in J} \tilde{M} \chi^j$ we get $\tilde{M} \sum_{i \in I} \tilde{N} \lambda^i - \tilde{N} \sum_{j \in J} \tilde{M} \chi^j \neq 0$. Since this is an integer we must then have $|\tilde{M} \sum_{i \in I} \tilde{N} \lambda^i - \tilde{N} \sum_{j \in J} \tilde{M} \chi^j| \geq 1$. Hence

$$\tilde{M} \tilde{N} d_\lambda = \left| \tilde{M} \sum_{i \in I} \tilde{N} \lambda^i - \tilde{N} \sum_{j \in J} \tilde{M} \chi^j \right| \geq 1$$

and so $d_\lambda \geq \frac{1}{\tilde{M} \tilde{N}}$ as desired.

One particularly useful case of this is when we have equal weights.

**Corollary 8.3.** If we have $\gcd(M, N) = 1$, $\lambda^i = \frac{1}{N}$, and $\chi^j = \frac{1}{M}$ then $d_\lambda \geq \frac{1}{M N}$.

In general we can see that computation of $d_\lambda$ is equivalent to a variant of the subset sum problem.

**Proposition 8.4.** Let $S = \{\lambda^2, \ldots, \lambda^N, -\lambda^1, \ldots, -\chi^M\}$ (note that $S$ doesn’t include $\lambda^1$). Then

$$d_\lambda = \min_{A \subseteq S} \left| \sum_{a \in A} a \right|$$

where the minimum is taken over subsets $A$ of $S$ that contain some $\lambda^i$.

**Proof.** Let $\Lambda = \{\lambda^1, \ldots, \lambda^N\}$ and $X = \{-\chi^1, \ldots, -\chi^M\}$. Note that if $A \subseteq S$ and $A \cap \Lambda \neq \emptyset$ then

$$\left| \sum_{a \in A} a \right| = \left| \sum_{p \in A \cap \Lambda} p + \sum_{q \in A \cap X} q \right| \geq d_\lambda$$

as $\sum_{q \in A \cap X} q + \sum_{p \in A \cap \Lambda} p$ is the difference between a subset sum of $\lambda$ and $\chi$.

Conversely suppose that we are given $P \subseteq \Lambda$ and $Q \subseteq \{\chi^1, \ldots, \chi^M\}$ where $P \neq \emptyset$, $\Lambda$ such that $d_\lambda = \left| \sum_{p \in P} p - \sum_{q \in Q} q \right|$. If $\lambda^1 \notin P$ set $A = P \cup \{-Q\}$, where $-Q = \{-q : q \in Q\}$. We see that

$$\left| \sum_{a \in A} a \right| = \left| \sum_{p \in P} p - \sum_{q \in Q} q \right| = d_\lambda.$$

If $\lambda^1 \in P$ then set $A = (\Lambda \setminus P) \cup (X \setminus -Q)$.

We see that

$$\left| \sum_{a \in A} a \right| = \left| \sum_{p \in P} p - \sum_{q \in Q} q \right| = \left| (1 - \sum_{p \in P} p) - (1 - \sum_{q \in Q} q) \right| = \left| \sum_{p \in P} p - \sum_{q \in Q} q \right| = d_\lambda.$$

In either case we will have $\min_{A \subseteq S} \left| \sum_{a \in A} a \right| \leq d_\lambda$ where again the minimum is taken over subsets $A$ of $S$ that contain some $\lambda^i$.

\[ \Box \]

**Remark 8.5.** Unfortunately the subset sum problem is NP complete and so in general it may not be feasible to compute the exact value of $d_\lambda$. However in order to use the results of this paper all one needs is an approximation to $d_\lambda$. Although there are fast
approximate algorithms for the subset sum problem and an extensive literature, it seems
upon inspection that most papers require the elements of the set to be positive integers or
put other assumptions that are unreasonable for our situation. However, a slight variation
to the classical approximate subset sum algorithm works for our case. We present it for the
convenience of the reader.

First we describe a “Trim” function. This function takes an ordered list of real numbers and
throws out almost duplicates. Our Trim function differs from the classical Trim function
only in that it measures errors in absolute scale rather than log scale.

**Parameters:** A tolerance $\epsilon > 0$.

**Input:** A sorted list of real numbers, $\tilde{A} = (r_1, \ldots, r_m)$, where $r_1 \geq r_2 \geq \cdots \geq r_m$.

```
begin
\begin{align*}
A &= (r_1) \\
last &= r_1 \\
for i \in \{2, \ldots, m\} \ do \\
&\quad \text{if } r_i \geq last + \epsilon \ then \\
&\quad \quad \text{last } = r_i \\
&\quad \quad \text{Append } r_i \text{ to } A \\
end
end \\
return A
end
```

**Algorithm 4:** Trim Routine

It is clear that Algorithm 4 runs in $O(m)$ time where $m = |A|$, the size of the input list.
With the modified Trim function our approximate subset sum algorithm is virtually identical
to the classical one. We also recall that given two sorted lists $L_1, L_2$, we can compute the
sorted list given by their union in $O(|L_1| + |L_2|)$ time (see, for example, [CLRS09, Section
2.3.1]). Furthermore for a list $L$ and $r \in \mathbb{R}$ we use the notation $L + r$ to denote the list
obtained from adding $r$ to each element of $L$.

**Parameters:** A tolerance $\epsilon > 0$.

**Input:** An ordered list $S = (\lambda^2, \ldots, \lambda^N, -\chi^1, \ldots, -\chi^M)$, with $\lambda^i, \chi^j > 0$.

```
begin
\begin{align*}
L_1 &= (0) \\
\text{for } i \in \{2, \ldots, N\} \ do \\
&\quad \tilde{L}_i = L_{i-1} \cup (L_{i-1} + \lambda^i) \\
&\quad L_i = \text{Trim}(\tilde{L}_i, \frac{\epsilon}{N+M}) \\
end
L_N = L_N \setminus \{0\}. \\
\text{for } i \in \{N + 1, \ldots, N + M\} \ do \\
&\quad \tilde{L}_i = L_{i-1} \cup (L_{i-1} - \chi^{i-N}) \\
&\quad L_i = \text{Trim}(\tilde{L}_i, \frac{\epsilon}{N+M}) \\
end
\text{return } \min_{l \in L_{N+M}} |l|
end
```

**Algorithm 5:** Approximate Subset Sum
An analysis similar to that of the classical approximate subset sum algorithm (see for example, \cite[Theorem 35.8]{CLRS09}) shows that Algorithm 4 runs in polynomial time and indeed approximates $d_\lambda$. For the convenience of the reader we sketch the proof.

**Proposition 8.6.** Algorithm 3 terminates in at most $O\left(\frac{(N+M)^2}{\epsilon}\right)$ steps. Furthermore if $x$ is the value returned by the algorithm then $x - \epsilon \leq d_\lambda \leq x$.

**Proof.** Note that if $y \in \tilde{L}_i$ for any $i$ then $y \in [-1, 1]$. Hence we see by construction of our Trim Routine (Algorithm 3), that $|L_i| \leq \frac{2(N+M)}{\epsilon}$, where $|L_i|$ denotes the number of elements in $L_i$. This proves that Algorithm 3 terminates in at most $O\left(\frac{(N+M)^2}{\epsilon}\right)$ steps.

Next define $\hat{L}_i$ to be the collection of all subset sums of the first $i - 1$ elements of $S$ when $i \leq N - 1$ and define $\tilde{L}_i$ to be the collection of all subset sums of the first $i - 1$ elements of $S$ that contain some $\lambda^j$ when $i > N - 1$. It is clear that $L_i \subset \tilde{L}_i$. In particular $L_{N+M} \subset \hat{L}_{N+M}$. Since by Proposition 8.4, $d_\lambda = \min_{y \in L_{N+M}} |y|$, this gives us that $d_\lambda \leq x$.

Next note that by construction of our Trim Routine (Algorithm 4), for every $y \in \hat{L}_i$ there is a $z \in L_i$ so that $|y - z| \leq \frac{\epsilon}{M+N}$, i.e. $d_H(\hat{L}_i, L_i) \leq \frac{\epsilon}{M+N}$, where we use $d_H$ for Hausdorff distance. Since $L_i$ arises from $i$ applications of the Trim routine we see that $d_H(\hat{L}_i, L_i) \leq i \frac{\epsilon}{M+N}$. In particular $d_H(\hat{L}_{M+N}, L_{M+N}) \leq (M + N) \frac{\epsilon}{M+N} = \epsilon$ and so by Proposition 8.4, we get $d_\lambda \geq x - \epsilon$.

\[\square\]

9. **Acknowledgments**

I would like to thank Prof. Kitagawa for many helpful comments and suggestions on a previous version of this manuscript.

**References**

[BM19] Mohit Bansil and Jun Kitagawa. A Newton algorithm for semi-discrete optimal transport with storage fees and quantitative convergence of cells. *arXiv e-prints*, page arXiv:1905.01249, May 2019.

[CLRS09] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms, Third Edition*. The MIT Press, 3rd edition, 2009.

[KMT19] Jun Kitagawa, Quentin Mérigot, and Boris Thibert. Convergence of a Newton algorithm for semi-discrete optimal transport. *J. Eur. Math. Soc. (JEMS)*, Advance online publication, doi: 10.4171/JEMS/889, 2019.

[Loe09] Grégoire Loeper. On the regularity of solutions of optimal transportation problems. *Acta Math.*, 202(2):241–283, 2009.

[Mir15] Jean-Marie Mirebeau. Discretization of the 3D Monge-Ampere operator, between wide stencils and power diagrams. *ESAIM Math. Model. Numer. Anal.*, 49(5):1511–1523, 2015.

[MMT18] Quentin Mérigot, Jocelyn Meyron, and Boris Thibert. An algorithm for optimal transport between a simplex soup and a point cloud. *SIAM J. Imaging Sci.*, 11(2):1363–1389, 2018.

[MTW05] Xi-Nan Ma, Neil S. Trudinger, and Xu-Jia Wang. Regularity of potential functions of the optimal transportation problem. *Arch. Ration. Mech. Anal.*, 177(2):151–183, 2005.

[Nes18] Yurii Nesterov. *Lectures on convex optimization*, volume 137 of *Springer Optimization and Its Applications*. Springer, Cham, 2018. Second edition of [MR2142598].

[OP88] Vladimir Oliker and Laird Prussner. On the numerical solution of the equation $(\partial^2 z/\partial x^2)(\partial^2 z/\partial y^2) - ((\partial^2 z/\partial x \partial y))^2 = f$ and its discretizations, I. *Numer. Math.*, 54(3):271–293, 1988.
[TW09] Neil S. Trudinger and Xu-Jia Wang. On the second boundary value problem for Monge-Ampère type equations and optimal transportation. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 8(1):143–174, 2009.

[Vil03] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[Vil09] Cédric Villani. *Optimal transport: Old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009.

Department of Mathematics, Michigan State University

*E-mail address*: bansilmo@msu.edu