SUPPLEMENTARY MATERIAL

Proof of Proposition 2.1

If we denote \( \gamma := (T_{\mu}, T_{\nu})\#\rho \), then \( \gamma \in \Pi(\mu, \nu) \). The change of variable formula gives

\[
W_2^2(\mu, \nu) \leq \int_{\mathbb{R}^2} ||y - y'||^2 d\gamma(y, y') \\
= \int_{\mathbb{R}^2} ||T_{\mu}(x) - T_{\nu}(x)||^2_{L^2(\rho)} dx = ||T_{\mu} - T_{\nu}||^2_{L^2(\rho)}.
\]

The continuity of the map \( \mu \mapsto T_{\mu} \) follows from e.g. Exercise 2.17 in [Villani 2003]. To prove (iii), we use the following lemma:

**Lemma 5.1.** Let \( \rho \) be uniform on the unit disc \( X \subseteq \mathbb{R}^2 \). Then, there is a curve \( \theta \in [0, 2\pi] \to \mu_{\theta} \in \mathcal{P}(X) \) and \( C > 0 \) such that \( ||T_{\mu_{\theta}} - T_{\mu_{\theta}}||_{L^2(\rho)} \geq CW_2(\mu_{\theta}, \mu_0)^{1/2} \).

**Proof.** Given \( \theta \in \mathbb{R} \), we denote \( x_{\theta} = (\cos \theta, \sin \theta) \) and \( \mu_{\theta} = \frac{1}{2}(\delta_{x_{\theta}} + \delta_{-x_{\theta}}) \). Then, the optimal transport map between \( \rho \) and \( \mu_{\theta} \) is given by

\[
T_{\mu_{\theta}}(x) = \begin{cases} 
 x_{\theta} & \text{if } (x|x_{\theta}) \geq 0 \\
 -x_{\theta} & \text{if not.}
\end{cases}
\]

One can easily check that for \( \theta \) one has \( W_2(\mu_{\theta}, \mu_0) \leq \sqrt{|\theta|} \). For \( \theta > 0 \) we set

\[
D_\theta = \{ x \in \mathbb{R}^2 \mid (x|x_{\theta}) \geq 0 \text{ and } (x|x_{\theta}) \leq 0 \}.
\]

Then, on \( D_\theta \), \( T_{\mu_{\theta}} = x_{-\theta} \) and \( T_{\mu_0} = x_0 \), giving

\[
||T_{\mu_{\theta}} - T_{\mu_0}||_{L^2(\rho)} \geq \int_{D_\theta} ||x_{-\theta} - x_0||^2 dx = |D_\theta| ||x_{-\theta} - x_0||^2.
\]

Moreover, if \( |\theta| \leq \frac{\pi}{2} \) one has \( ||x_{-\theta} - x_0||^2 \geq 2 \), thus giving \( ||T_{\mu_{\theta}} - T_{\mu_0}||_{L^2(\rho)} \geq 2 |D_\theta| \geq \frac{|\theta|}{\pi} \). \( \square \)

Proof of Theorem 2.2

The proof of Theorem 2.2 is based on the following lemma.

**Lemma 5.2.** Under the assumptions of Theorem 2.2,

\[
||T_{\mu} - T_{\nu}||_{L^2(\rho)}^2 \leq 2K \int_X (\psi_{\nu} - \psi_{\mu}) d(\mu - \nu).
\]

**Proof.** From convex analysis, the map \( T_{\mu} = \nabla \phi_{\mu} \) is K-Lipschitz if and only if \( \psi_{\mu} = \phi_{\mu}^* \) is \( \frac{1}{K} \)-strongly convex. We denote \( A = \int_X \psi_{\nu} d(\mu - \nu) \) and \( B = \int_Y \psi_{\nu} d(\nu - \mu) \).

We use that \( (\nabla \phi_{\mu})\#\rho = \mu \) (resp. \( (\nabla \phi_{\nu})\#\rho = \nu \)) and \( \nabla \phi_{\mu} = \nabla \psi_{\mu} \) (resp. \( \nabla \phi_{\nu} = \nabla \psi_{\nu} \)) by convexity of \( \phi_{\mu} \) (resp. \( \phi_{\nu} \)) to do the following change of variable:

\[
A = \int_X (\psi_{\nu}(\nabla \psi_{\mu}^* - \psi_{\nu}(\nabla \psi_{\nu}^*)d\rho.
\]

We now use the inequality \( \psi_{\nu}(y) - \psi_{\nu}(z) \geq (y - z|v) \), which holds for all \( v \) in the subdifferential \( \partial \psi_{\nu}(z) \). The convex functions \( \psi_{\nu}, \psi_{\mu} \) are differentiable \( \rho \)-almost everywhere. Taking \( z = \nabla \psi_{\nu}^*(x) \) and \( y = \nabla \psi_{\mu}^*(x) \), and using \( x \in \partial \psi_{\nu}(z) \), we obtain

\[
A \geq \int_X (\nabla \psi_{\mu}^* - \nabla \psi_{\nu}^*) d\rho.
\]
Using the strong convexity of $\psi_\mu$, we get a similar lower bound on $B$, with an extra quadratic term

\[ B = \int_\mathcal{X}(\psi_\mu(\nabla \psi_\mu^*) - \psi_\mu(\nabla \psi_\mu^*))d\rho \]
\[ \geq \int_\mathcal{X}(\langle \text{id}, \nabla \psi_\mu^* - \nabla \psi_\mu^* \rangle + \frac{1}{2K} \| \nabla \psi_\mu^* - \nabla \psi_\mu^* \|^2_2)d\rho. \]

Summing up these inequalities we get:
\[ \int_\mathcal{Y}(\psi_\nu - \psi_\mu)d(\mu - \nu) \geq \frac{1}{2K} \int_\mathcal{X} \| \nabla \psi_\mu^* - \nabla \psi_\mu^* \|^2_2d\rho \]
\[ = \frac{1}{2K} \| T_\nu - T_\mu \|^2_{L^2(\rho)}. \]

**Proof of Theorem 2.2.** Formula (3) clearly shows that $\text{Lip}(\psi_\mu) \leq M_X$, where $\text{Lip}(f)$ denotes the Lipschitz constant of $f$. Combining this with Lemma 5.2,
\[ \| T_\mu - T_\nu \|^2_{L^2(\rho)} \leq 2K \int_\mathcal{Y}(\psi_\nu - \psi_\mu)d(\mu - \nu) \]
\[ \leq 2K \max_{\text{Lip}(f) \leq 2M_X} \int_\mathcal{Y} f d(\mu - \nu) \]
\[ = 4K M_X \max_{\text{Lip}(f) \leq 1} \int_\mathcal{Y} f d(\mu - \nu) \]
\[ = 4K M_X W_1(\mu, \nu), \]
where we used Kantorovich-Rubinstein’s theorem to get the last equality.

**Proof of Corollary 2.4**

We first state a simple lemma that links the uniform norm of a Lipschitz function to its $L^2(\rho)$ norm:

**Lemma 5.3.** If $f$ is $L$-Lipschitz on $\mathcal{X}$, then
\[ ||f||_\infty \leq C ||f||_{L^2(\mathcal{X})}^{\frac{2}{d+2}}, \]
for some $C$ depending on $L$, $d$ and $\mathcal{X}$ only.

**Proof.** If $||f||_\infty = \eta$, then there exists $x_0 \in \mathcal{X}$ such that for all $x \in B(x_0, \frac{\eta}{2L}) \cap \mathcal{X}$ we have $|f(x)| \geq \frac{\eta}{2}$. This implies that
\[ ||f||_{L^2(\mathcal{X})}^2 \geq \int_{B(x_0, \frac{\eta}{2L}) \cap \mathcal{X}} f(x)^2dy \geq \beta_d \left( \frac{\eta}{2L} \right)^d \eta^2 = \frac{\beta_d}{(2L)^d} ||f||_\infty^{d+2}, \]
where $\beta_d$ is the volume of the $d$-dimensional unit ball.

**Proof of Corollary 2.4.** Theorem 2.3 implies
\[ ||\nabla \psi_\mu - \nabla \psi_\nu||_{L^2(\mathcal{Y})} \leq C \left( \int_{\mathcal{Y}} (\psi_\nu - \psi_\mu)d(\mu - \nu) \right)^{\frac{1}{d+2}}, \]
and as in Theorem 2.2, the quantity in the parenthesis can be upper bounded by $2M_X W_1(\mu, \nu)$. Adding a constant to $\psi_\mu$ if necessary, we can assume that $\int_{\mathcal{Y}} \psi_\mu(y)dy = \int_{\mathcal{Y}} \psi_\nu(y)dy$. The Poincaré-Wirtinger inequality on $\mathcal{Y}$ then implies
\[ ||\psi_\mu - \psi_\nu||_{L^2(\mathcal{Y})} \leq C' W_1(\mu, \nu)^{\frac{1}{d+2}}, \]
for some $C'$ depending only on $\rho$, $\mathcal{X}$ and $\mathcal{Y}$.

We reuse the fact that $\psi_\mu - \psi_\nu$ is Lipschitz with a constant at most $2M_X$ to use Lemma 5.3
\[ ||\psi_\mu - \psi_\nu||_\infty \leq C'' W_1(\mu, \nu)^{\frac{2}{d+2}}. \]
Since $\phi_\mu = \psi_\mu^*$ and $\phi_\nu = \psi_\nu^*$, the definition of the Legendre transform yields

$$||\phi_\mu - \phi_\nu||_\infty \leq C''W_1(\mu, \nu)^{\frac{2}{d+2}}.$$  

We conclude using Proposition 3.6 and the fact that $\phi_\mu$ is diam($\mathcal{Y}$)-Lipschitz (as the Legendre transform of the function $\psi_\mu$ on $\mathcal{Y}$): there exists a constant $C$ depending only on $\rho$, $\mathcal{X}$ and $\mathcal{Y}$ such that

$$||T_\mu - T_\nu||_{L^2(\rho)} \leq CW_1(\mu, \nu)^{\frac{2}{d+2}}. \quad \square$$

**Proof of Lemma 3.2**

For any $N > 0$, we consider a finite partition $\mathcal{Y} = \sqcup_{1 \leq i \leq N} Y_i^N$, we let $\varepsilon_N = \max_i \text{diam}(Y_i^N)$ and we assume that $\lim_{N \to +\infty} \varepsilon_N = 0$. Then, we define

$$\mu_k^N = \sum_{1 \leq i \leq N} \left(1 - \frac{1}{N}\right) \mu^k(Y_i^N) + \frac{1}{N^2} \delta_{y_i^N},$$

where $y_i^N \in Y_i^N$. Then, it is easy to check that the support of the measures $\mu_0^N$ and $\mu_1^N$ is the set $\{y_1^N, \ldots, y_N^N\}$. Moreover,

$$||\mu_N - \mu_0^N||_{TV} \leq ||\mu_1 - \mu_0||_{TV}.$$  

In addition, $W_1(\mu_k^N, \mu_0^N) \leq \varepsilon_N \frac{N \to +\infty}{N}$. Combined with the triangle inequality, we deduce

$$|W_1(\mu_0^N, \mu_N^N) - W_1(\mu_0, \mu_1)| = |W_1(\mu_0^N, \mu_N^N) - W_1(\mu_0^N, \mu_1^N) + W_1(\mu_N^N, \mu_1^N) - W_1(\mu_0^N, \mu_1)|$$

$$\leq |W_1(\mu_0^N, \mu_N^N) - W_1(\mu_0^N, \mu_1^N)| + |W_1(\mu_N^N, \mu_1^N) - W_1(\mu_0^N, \mu_1)|$$

$$\leq W_1(\mu_N^N, \mu_1^N) + W_1(\mu_0^N, \mu_0)$$

$$\leq 2\varepsilon_N \frac{N \to +\infty}{N} W_1(\mu_0^N, \mu_1^N).$$

Using the stability of optimal transport maps (Proposition 2.1), we finally deduce that

$$\lim_{N \to +\infty} ||T_{\mu_k^N} - T_{\mu_0^N}||_{L^2(\rho)} = ||T_{\mu_1} - T_{\mu_0}||_{L^2(\rho)}.$$  

**Proof of Lemma 3.3**

Let $x^0 \in V_i(\psi_0)$ and $x^1 \in V_i(\psi_1)$. Then, for all $j \in \{1, \ldots, N\}$,

$$\begin{cases} 
\psi_0(y_j) \geq \psi_0(y_i) + (y_j - y_i)x^0, \\
\psi_1(y_j) \geq \psi_1(y_i) + (y_j - y_i)x^1.
\end{cases}$$

Taking the convex combination of these inequalities we get for all $j \in \{1, \ldots, N\}$,

$$\psi^t(y_j) \geq \psi^t(y_i) + (y_j - y_i)(1-t)x^0 + tx^1.$$  

This shows that $(1-t)x^0 + tx^1 \in V_i(\psi^t)$ (note that we use the convexity of $\mathcal{X}$ here). Thus, $$(1-t)V_i(\psi_0) + tV_i(\psi_1) \subseteq V_i(\psi^t).$$

Taking the Lebesgue measure on both sides and applying Brunn-Minkowski’s inequality we get

$$G_i(\psi^t)^{1/d} = \rho(V_i(\psi^t))^{1/d} \geq \rho((1-t)V_i(\psi_0) + tV_i(\psi_1))^{1/d}$$

$$\geq (1-t)\rho(V_i(\psi_0))^{1/d} + t\rho(V_i(\psi_1))^{1/d}$$

$$\geq (1-t)G_i(\psi_0)^{1/d} + tG_i(\psi_1)^{1/d}.$$  

This inequality directly implies

$$G_i(\psi^t) \geq \min(G_i(\psi_0), G_i(\psi_1)).$$
We consider the function \( u \) and conclude using the same formula as above:

\[
\frac{1}{2} \| G(\psi^t) - G(\psi^0) \|_1 = 1 - \sum_i \min(G_i(\psi^t), G_i(\psi^0))
\]

so it suffices to control the right hand side of this equality. Given \( (i,j) \) and \((x,y) \in \mathcal{X}^2 \), we denote

\[ \chi_{ij}(x,y) = \begin{cases} 1 & \text{if } V_i(\psi) \cap V_j(\psi) \cap [x,y] \neq \emptyset \text{ and } y_j - y_i |y - x| \geq 0 \\ 0 & \text{if not} \end{cases} \]

Then, \( u(y) - u(x) = \sum_{i \neq j} (v(y_j) - v(y_i)) \chi_{ij}(x,y) \). Denoting \( d_{ij} = \| y_j - y_i \| \), \( c_{ij,z} = \| \frac{y_j - y_i}{\| y_j - y_i \|} \| \) and applying Cauchy-Schwarz’s inequality we get

\[
(u(y) - u(x))^2 = \left( \sum_{i \neq j} (v(y_j) - v(y_i)) \chi_{ij}(x,y) \right)^2 \leq \sum_{i \neq j} \frac{(v(y_j) - v(y_i))^2}{d_{ij} c_{ij,y-x}} \chi_{ij}(x,y) \sum_{i \neq j} d_{ij} c_{ij,y-x} \chi_{ij}(x,y).
\]

In addition, when \( \chi_{ij}(x,y) = 1 \), we have \( y_j - y_i |y - x| \geq 0 \) so that

\[ d_{ij} c_{ij,y-x} = \| y_j - y_i \| \left( \frac{y - x}{\| y - x \|} \bigg| \frac{y_j - y_i}{\| y_j - y_i \|} \right) \geq 0, \]

and

\[
\sum_{i \neq j} d_{ij} c_{ij,y-x} \chi_{ij}(x,y) = \sum_{i \neq j} \left( \frac{y - x}{\| y - x \|} \bigg| y_j - y_i \right) \chi_{ij}(x,y) \leq \text{diam}(\mathcal{Y}).
\]

Therefore,

\[
\int_{\mathcal{X} \times \mathcal{X}} (u(y) - u(x))^2 dxdy \leq \text{diam}(\mathcal{Y}) \int_{\mathcal{X} \times \mathcal{X}} \sum_{i \neq j} \frac{(v(y_j) - v(y_i))^2}{d_{ij} c_{ij,y-x}} \chi_{ij}(x,y) dxdy
\]

\[
= \text{diam}(\mathcal{Y}) \int_{B(0,\text{diam}(\mathcal{X}))} \sum_{i \neq j} \frac{(v(y_j) - v(y_i))^2}{d_{ij} c_{ij,z}} \left( \int_{\mathcal{X}} \chi_{ij}(x,x+z) dx \right) dz.
\]
Moreover, denoting $m_{ij} = H^{d-1}(V_i(\psi) \cap V_j(\psi))$ we get
\[
\int_X \chi_{ij}(x, x + z)dx \leq m_{ij} \|z\|_{c_{ij,z}},
\]
thus giving
\[
\int_{X \times X} (u(y) - u(x))^2dxdy \leq C(d) \text{diam}(\mathcal{Y}) \text{diam}(X)^{d+1} \sum_{i \neq j} m_{ij} (v(y_j) - v(y_i))^2.
\]
Define $H_{ij} = \frac{m_{ij}}{d_{ij}}, H_{ii} = -\sum_{j \neq i} H_{ij}$. Then, $DG(\psi) = H$, and
\[
\langle DG(\psi)v | v \rangle = \sum_{i,j} H_{ij}v_i v_j
\]
\[
= \sum_i \left( H_{ii}v_i v_i + \sum_{j \neq i} H_{ij}v_i v_j \right)
\]
\[
= \sum_i \sum_{j \neq i} H_{ij}v_i (v_j - v_i)
\]
\[
= \sum_{j \neq i} H_{ij}v_i (v_j - v_i) := A.
\]
And
\[
\sum_{i \neq j} H_{ij} (v(y_j) - v(y_i))^2 = \sum_{i \neq j} H_{ij}v_j (v_j - v_i) - \sum_{i \neq j} H_{ij}v_i (v_j - v_i) = -2A.
\]
We finally obtain
\[
\iint (u(y) - u(x))^2dxdy \leq -C_{d,X,Y}(DG(\psi)v | v).
\]