# Supplemental material for *Dynamical Optimal Transport on Discrete Surfaces*

## 1 PROOF OF PROPOSITION 3.2

Recall that the Wasserstein distance is defined as

$$W_d^2(\bar\mu_0,\bar\mu_1)=\begin{cases}\sup_\varphi & \sum_{v\in V}|v|\varphi_v^1\bar\mu_v^1-\sum_{v\in V}|v|\varphi_v^0\bar\mu_v^0\\ \text{s.t.} & \partial_t\varphi_v^t+\dfrac{1}{2}\dfrac{\sum_{f\in T_v}|f|\|(G\varphi)_f^t\|^2}{3|v|}\leqslant 0\\ & \text{for all }(t,v)\in[0,1]\times V.\end{cases}$$

The constraint $\partial_t\varphi_v^t+\dfrac{1}{2}\dfrac{\sum_{f\in T_v}|f|\|(G\varphi)_f^t\|^2}{3|v|}\leqslant 0$ can be rewritten using the identity

$$\inf_\mu\int_0^1\sum_{v\in V}|v|\mu_v^t\left(\partial_t\varphi_v^t+\frac{1}{2}\frac{\sum_{f\in T_v}|f|\|(G\varphi)_f^t\|^2}{3|v|}\right)\mathrm{d}t$$
$$=\begin{cases}0 & \text{if }\partial_t\varphi_v^t+\dfrac{1}{2}\dfrac{\sum_{f\in T_v}|f|\|(G\varphi)_f^t\|^2}{3|v|}\leqslant 0\\ & \text{for all }(t,v)\in[0,1]\times V\\ -\infty & \text{else}\end{cases}$$

where the infimum is taken over all non negative $\mu:[0,1]\times V\to\mathbb{R}$. Hence the discrete Wasserstein distance $W_d$ is the saddle point of the following Lagrangian:

$$W_d^2(\bar\mu_0,\bar\mu_1)=\sup_\varphi\inf_\mu\sum_{v\in V}|v|\varphi_v^1\bar\mu_v^1-\sum_{v\in V}|v|\varphi_v^0\bar\mu_v^0$$
$$+\int_0^1\sum_{v\in V}|v|\mu_v^t\left(\partial_t\varphi_v^t+\frac{1}{2}\frac{\sum_{f\in T_v}|f|\|(G\varphi)_f^t\|^2}{3|v|}\right)\mathrm{d}t.$$

Rearranging the last sum so that it is indexed by triangles and using the identity $-\frac{1}{2}\|(G\varphi)_t\|^2=\inf_{\mathbf{v}}[-\mathbf{v}\cdot(G\varphi)_t+\frac{1}{2}\|\mathbf{v}\|^2]$, we see that

$$W_d^2(\bar\mu^0,\bar\mu^1)=\sup_\varphi\inf_{\mu,\mathbf{v}}\sum_{v\in V}\varphi_v^1\bar\mu_v^1-\sum_{v\in V}\varphi_v^0\bar\mu_v^0$$
$$-\int_0^1\left(\sum_{v\in V}|v|\mu_v^t\partial_t\varphi_v^t\right)\mathrm{d}t$$
$$+\int_0^1\left(\sum_{f\in T}|f|\underbrace{\left(\frac{1}{3}\sum_{v\in V_f}\mu_v^s\right)}_{\hat\mu_f^t}\left(-\mathbf{v}_f^t\cdot(G\varphi)_f^t+\frac{1}{2}\|\mathbf{v}_f^t\|^2\right)\right)\mathrm{d}t.$$

Then, one can exchange the sup and the inf; a rigorous proof would involve the Rockefeller–Fenchel duality Theorem [Theorem 1.9, *Topics in Optimal Transportation*, Villani]. Taking the supremum in $\varphi$ yields a weak formulation of the discrete continuity equation with the boundary conditions $\bar\mu^0,\bar\mu^1$. In particular, it implies that the mass is conserved, and hence $\mu$ is valued in $\mathcal{P}(S)$. The remaining term is nothing that the integral over time of the kinetic energy.

## 2 THE WASSERSTEIN DISTANCE AS A GEODESIC DISTANCE: THE CONTINUOUS CASE

Recall that if $(\mathcal{M},\langle\ ,\ \rangle)$ is a Riemannian manifold, then the geodesic distance between $x,y\in\mathcal{M}$ is defined as

$$d(x,y)^2=\inf_\gamma\int_0^1\langle\dot\gamma,\dot\gamma\rangle\mathrm{d}t,$$

where the infimum is taken over all curves $\gamma:[0,1]\to\mathcal{M}$ such that $\gamma^0=x$ and $\gamma^1=y$.

Now, we consider $\mathcal{P}(\mathcal{M})$ as Riemannian manifold as follows. As explain in the article, if $\mu\in\mathcal{P}(\mathcal{M})$, the tangent space $T_\mu\mathcal{P}(\mathcal{M})$ is identified as the set of functions $\delta\mu:\mathcal{M}\to\mathbb{R}$ with 0-mean: $\delta\mu$ is the partial derivative w.r.t. time of a curve whose value at time 0 is $\mu$. The square of the norm of $\delta\mu\in T_\mu\mathcal{P}(\mathcal{M})$ is defined as

$$\|\delta\mu\|^2_{T_\mu\mathcal{P}(\mathcal{M})}:=\frac{1}{2}\int_{\mathcal{M}}\|\nabla\varphi\|^2\mathrm{d}\mu,\tag{1}$$

where $\varphi$ is the solution (unique up to a constant) of the elliptic equation

$$\nabla\cdot(\mu\nabla\varphi)=-\delta\mu.\tag{2}$$

Now, if we consider the geodesic distance induced on $\mathcal{P}(\mathcal{M})$ by the metric tensor defined by (1) and (2), we end up with the following problem of calculus of variations:

$$\begin{cases}\inf_\mu & \int_0^1\|\dot\mu\|^2_{T_\mu^t\mathcal{P}(\mathcal{M})}\mathrm{d}t\\ \text{s.t.} & \mu:[0,1]\to\mathcal{P}(\mathcal{M})\text{ and }\mu^0=\bar\mu^0,\mu^1=\bar\mu^1.\end{cases}\tag{3}$$

Taking in account (1) and (2) which give the definition of $\|\dot\mu\|^2_{T_\mu^t\mathcal{P}(\mathcal{M})}$, (3) can be rewritten

$$\begin{cases}\inf_\mu & \int_0^1\int_{\mathcal{M}}\frac{1}{2}\|\nabla\varphi^t\|^2\mathrm{d}\mu^t\mathrm{d}t\\ \text{s.t.} & \mu^0=\bar\mu^0,\mu^1=\bar\mu^1,\\ & -\partial_t\mu=\nabla\cdot(\mu\nabla\varphi).\end{cases}\tag{4}$$

Compared with the Benamou–Brenier formula (equation (2) in the article), and given the fact, as recalled in the article, that the minimum in the Benamou–Brenier formula is reached for a velocity field $\mathbf{v}=\nabla\varphi$ which is the gradient of a time-dependent potential $\varphi$, we reach the conclusion that (4) reads exactly as the Wasserstein distance as given in equation (2) of the article.

## 3 THE WASSERSTEIN DISTANCE AS A GEODESIC DISTANCE: THE DISCRETE CASE

Now we switch to the case of a triangle mesh $S$. If $\mu\in\mathcal{P}(S)$, the tangent space at $\mu$ is naturally $\{x\in\mathbb{R}^{|V|}\text{ s.t. }\sum_{v\in V}|v|x_v=0\}$. In analogy to the continuous case, the square of the norm of $\delta\mu\in T_\mu\mathcal{P}(S)$ is

$$\|\delta\mu\|^2_{T_\mu\mathcal{P}(S)}:=\frac{1}{2}\sum_{f\in T}\|(G\varphi)_f\|^2|f|\hat\mu_f,\tag{5}$$

where $\varphi$ is a solution of the equation

$$M_V\delta\mu=-(G^\top M_T M_{\hat\mu}G)\varphi.\tag{6}$$

Here, $M_{\hat{\mu}} \in \mathbb{R}^{3|T|\times 3|T|}$ is a diagonal matrix corresponding to multiplication on each triangle by $\hat{\mu}$.

The fact that the geodesic distance induced by (5) and (6) is exactly $W_d$, as expressed in Proposition 3.2 of the article, is proved in exactly the same way as in the previous section for the continuous case.

Then, we can give an explicit expression of the metric tensor, namely

$$P_\mu = \frac{1}{2} M_V^\top G^{-\top} (M_{\hat{\mu}} M_T)^{-1} G^{-1} M_V. \tag{7}$$

It just comes by rewriting of (5) and (6) with matrices. Indeed, (6) reads

$$\varphi = [-(G^\top M_T M_{\hat{\mu}} G)^{-1} M_V] \delta\mu. \tag{8}$$

But on the other hand,

$$\|\delta\mu\|^2_{T_\mu \mathcal{P}(S)} = \frac{1}{2} \sum_{f \in T} \|(G\varphi)_f\|^2 |f| \hat{\mu}_f = \frac{1}{2} \varphi^\top G^\top M_T M_{\hat{\mu}} G \varphi.$$

Replacing $\varphi$ in (5) by the value given in (8) yields

$$\begin{aligned}\|\delta\mu\|^2_{T_\mu \mathcal{P}(S)} = &\frac{1}{2} (\delta\mu)^\top [-(G^\top M_T M_{\hat{\mu}} G)^{-1} M_V]^\top \\ &G^\top M_T M_{\hat{\mu}} G [-(G^\top M_T M_{\hat{\mu}} G)^{-1} M_V](\delta\mu) = (\delta\mu)^\top P_\mu (\delta\mu),\end{aligned}$$

where

$$\begin{aligned}P_\mu &:= [(G^\top M_T M_{\hat{\mu}} G)^{-1} M_V]^\top G^\top M_T M_{\hat{\mu}} G [(G^\top M_T M_{\hat{\mu}} G)^{-1} M_V] \\ &= \frac{1}{2} M_V^{-\top} G^{-\top} (M_{\hat{\mu}} M_T)^{-1} G^{-1} G^\top M_{\hat{\mu}} M_T G G^{-1} (M_{\hat{\mu}} M_T)^{-1} G^{-\top} M_V \\ &= \frac{1}{2} M_V^\top G^{-\top} (M_{\hat{\mu}} M_T)^{-1} G^{-1} M_V\end{aligned}$$

indeed coincides with the (7).