9/1/2026 AMPO UP presentation

I. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem

A. Intro.

1. Monge's problem
   a. M = set of mines
   b. F = set of factories
   c. cost of transporting output of ~~fac~~ mine m to factor f: $c(m, f)$
   d. assume we don't want to split output of any mine
   e. assume each factory needs output of exactly one mine.
   f. What is cheapest way to assign mines' outputs to factories?

$$\min_{T: M \to F} \sum_{m \in M} c(m, T(m))$$

2. More generally:
   a. What if factory f needs i_f inputs and mine m makes o_m outputs?
   b. then we want

$$\min_{T: M \to F} \sum_{m \in M} c(m, T(m))$$

s.t. $\sum_{m \text{ s.t. } T(m)=f} o_m = i_f$

3. We could normalize so $\sum \tilde{o}_m = 1$, $\sum \tilde{i}_f = 1$
   a. then this would be a problem about sending probability distr'ns to probability distr'ns

4. Arrive at full Monge problem:
   a. have distr'ns ρ_0, ρ_T on $\mathbb{R}^d$:

$$\int_{\mathbb{R}^d} \rho_0(x)\,dx = 1 = \int_{\mathbb{R}^d} \rho_T(x)\,dx$$

   b. Consider $M: \mathbb{R}^d \to \mathbb{R}^d$ that transfers ρ_0 to ρ_T, which means for all $A \subset \mathbb{R}^d$ Borel:

$$\int_{x \in A} \rho_T(x)\,dx = \int_{M(x) \in A} \rho_0(x)\,dx$$

   i. equivalently, $M_\# \rho_0 = \rho_T$, so $\rho_0(M^{-1}(A)) = \rho_T(A)$
   then for any test function, $\int \varphi(x) \rho_T(x)\,dx = \int (\varphi \circ M)(x) \rho_0(x)\,dx$

Kantorovich

4. our goal: find M to minimize a distance
$$d_2(\rho_0, \rho_T)^2 = \inf_M \int |M(x)-x|^2 \rho_0(x)\, dx$$

5. general approach
a. look at "Kantorovich relaxed problem"
b. look at dual to Kantorovich
c. prove sol'n exists for dual
d. show that (for this L^2 case), this sol'n is given by a sol'n to Monge problem

6. idea in this paper
a. instead of jumping straight from ρ_0 to ρ_T, let's consider a gradual change from ρ_0 to ρ_T given by $\rho(t,x)$ for $0 \le t \le T$ where $\rho(0,x) = \rho_0(x)$, $\rho(T,x) = \rho_T(x)$.

7. In particular, let's consider densities satisfying the continuity eq'n:
$$\partial_t \rho + \nabla \cdot (\rho v) = 0$$

probability p.o.v.

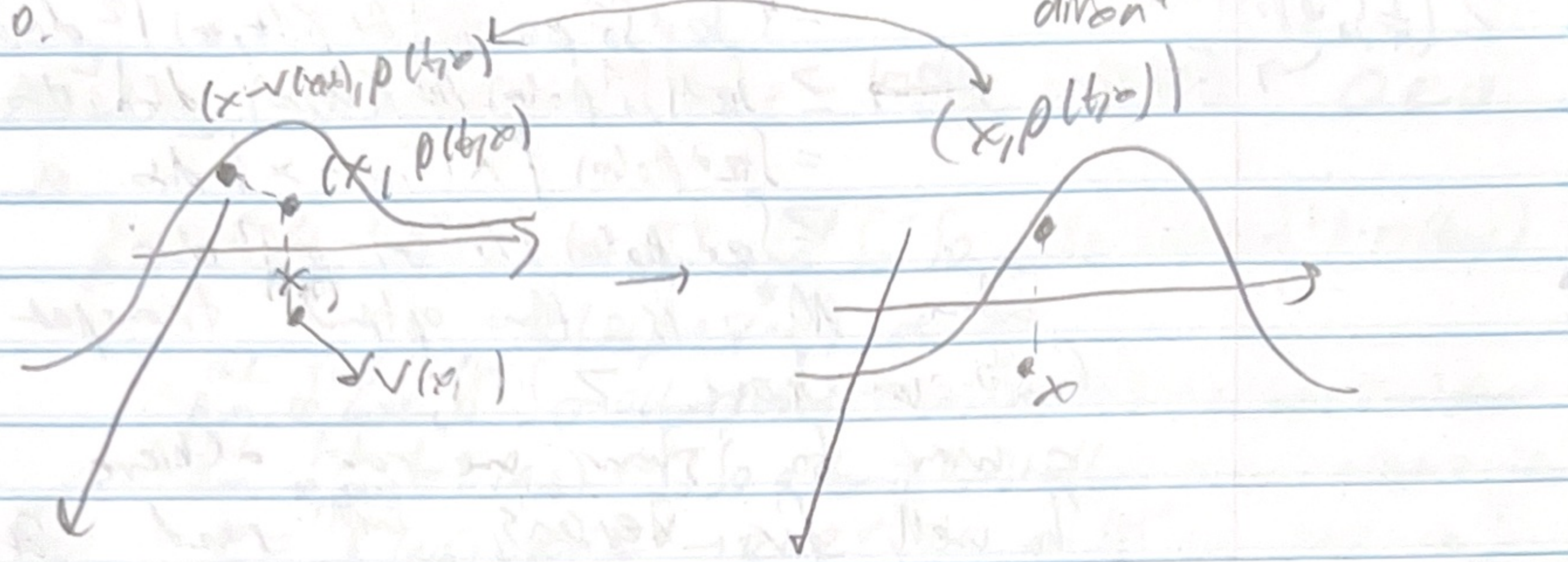
8. Background: the continuity equation
a. say we have some ~~particle~~ r.v. evolving per
$$\dot{X}_t = v(X_t, t)$$
b. say this particle is initially drawn from density $X_0 \sim \rho_0$
c. claim: law $\rho(t,x)$ of X_t will satisfy
$$\partial_t \rho + \nabla \cdot (\rho v) = 0$$
d. check
i. for test fcn φ,
$$\partial_t \mathbb{E}(\varphi(X_t)) = \mathbb{E}(\nabla\varphi(X_t) \cdot \partial_t X_t)$$
$$\int \varphi(x)\, \partial_t \rho(x,t)\, dx \qquad \int \nabla\varphi(x) \cdot v(x,t)\, \rho(x,t)\, dx$$
$$-\int \varphi(x)\, \nabla \cdot (v(x,t)\, \rho(x,t))\, dx$$
e. so we're looking at densities that correspond to the particles whose density is being described evolving per $\dot{X} = v(X)$.

+ A a. alt. background: physics

i. particles evolve in direction V, can calculate their density through a small cube, get result

10.

for fixed q density moves in "opposite direction"



a. we shift the coord's per v — result in shift of fcn by $-v$.

B. Main ~~thm~~ result

1. Proposition: The square of the L^2 Kantorovich distance is equal to the infimum of

$$T \int_{\mathbb{R}^d} \int_0^T \rho(t,x)\,|v(t,x)|^2\, dx$$

among all (ρ, v) satisfying above continuity eq'n and initial and final conditions

better: set $\rho_t = (X_t)_\# \rho_0$, among them this is by def'n and the continuity eq'n is easy to prove.

2. proof

a. first, let's consider

$$X(0,x) = x, \quad \partial_t X(t,x) = v(t, X(t,x))$$

b. claim: for all test functions φ,

$$\int \varphi(t,x)\,\rho(t,x)\, dx\, dt = \int \varphi(t, X(t,x))\,\rho_0(x)\, dx\, dt$$

where ρ, v satisfy ~~cont'n~~ cond's.

i. write $= \int \varphi(t,x)\,\rho(t,x|y,0)\,\rho_0(y)\, dy\, dx\, dt$

ii. I claim that $\rho(t,x|y,0) = \delta(x - X(t,y))$

iii. this follows by integration against test functions as above

iv. so $= \int \varphi(t, X(t,y))\,\rho_0(y)\, dy\, dt$

c. in particular,

$$\int \varphi(t,x)\,|v(t,x)|^2\,\rho(t,x)\, dx\, dt = \int \varphi(t, X(t,x))\,|v(t, X(t,x))|^2\,\rho_0(x)$$

d. now, by our "alt." formulation, this means that

$X(T,x)$ is a transport map.

to id. next,

$T\int_{\mathbb{R}^d}\int_0^T \rho(t,x)\,|v(t,x)|^2\,dx\,dt$

$\frac{1}{T}\int_0^T \varphi^2\,\rho_0(x)\,dx$

$= T\int_{\mathbb{R}^d}\int_0^T \rho_0(x)\,|v(t,X(t,x))|^2\,dx\,dt$

$\geq \left(\frac{1}{T}\int_0^T \varphi\,\rho_0(x)\,dt\right)^2$

$= T\int_{\mathbb{R}^d}\int_0^T \rho_0(x)\,|\partial_t X(t,x)|^2\,dx\,dt$

Jensen $\longrightarrow \geq \int_{\mathbb{R}^d}\left|\int_0^T \rho_0(x)\,\partial_t X(t,x)\,dt\right|^2 dx$

$= \int_{\mathbb{R}^d}\rho_0(x)\,|X(T,x)-x|^2\,dx$

$\geq \int_{\mathbb{R}^d}\rho_0(x)\,|M^*(x)-x|^2\,dx$

where M^* is the optimal transport

f. so we have $\geq$

g. want to show we can achieve $=$

h. well, for Jensen's, will need $\partial_t X = \text{const in } t$, and for the last bit, will want

$X(T,x) = M^*(x)$

i. ~~thus, we take~~ and, of course, need $X(0,x)$

j. solve this, s.t.

$X(t,x) = x + \frac{t}{T}(M^*(x) - x).$

b. i.e., take $v(t,x) = M^*(x) - x$, and then you have optimal solution.

2. So if we can find optimal v, ρ, we can get optimal transport using

$M^* = v + x$

C. Numerical method:

1. Introduce new notation $m = \rho v$ (momentum)

2. so let's look at objective function

$L(\varphi,\rho,m) = \int_0^T\int_D \frac{|m|^2}{2\rho} + \varphi(\partial_t\rho + \nabla\cdot m)\,dx\,dt$

a. φ = Lagrange multiplier

b. change out T for $\frac{1}{2}$ since doesn't change where minimum occurs

3. Claim: $\frac{|m(t,x)|^2}{2\rho(t,x)} = \sup_{(a,b)\in K}\left(a\rho(t,x) + b\cdot m(t,x)\right)$

for $K = \{(a,b)\in\mathbb{R}\times\mathbb{R}^d \text{ s.t. } a + \frac{|b|^2}{2} \leq 0\}$

4. proof

a. this is linear in a,b, so optimal on boundary $\left(a = -\frac{|b|^2}{2}\right)$

4. so use lagrange multipliers:
$$\begin{pmatrix}\rho(t,x)\\ m(t,x)\end{pmatrix} + \lambda\begin{pmatrix}1\\ b\end{pmatrix} = 0$$
c. so $\lambda = -\rho(t,x)$

d. then $b = \frac{m(t,x)}{\rho(t,x)}$

e. so $a = -\frac{|m(t,x)|^2}{2\rho(t,x)^2}$

f. and $a\,\rho(t,x) + b\cdot m(t,x) = -\frac{|m(t,x)|^2}{2\rho(t,x)} + \frac{|m(t,x)|^2}{\rho(t,x)}$ Q.E.D.

5. so
$$\inf_{\rho,m}\sup_{\varphi} L(\varphi,\rho,m) = \inf_{\rho,m}\sup_{\varphi}\int_0^T\int_D \sup_{(a,b)\in K}\left(a\rho(t,x) + b\cdot m(t,x)\right) + \varphi(\partial_t\rho + \nabla\cdot m)\,dx\,dt$$
$$= \inf_{\rho,m}\sup_{\varphi,(a,b)\in K}\int_0^T\int_D a(t,x)\rho(t,x) + b(t,x)\cdot m(t,x) + \varphi(\partial_t\rho + \nabla\cdot m)\,dx\,dt$$
$$= \inf_{\rho,m}\sup_{\varphi,(a,b)}\left(-F(a,b) + \int_0^T\int_D a(t,x)\rho(t,x) + b(t,x)\cdot m(t,x) + \varphi(\partial_t\rho + \nabla\cdot m)\,dx\,dt\right)$$
for $K = \{(a,b) : \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}\times\mathbb{R}^d$ s.t. $a + \frac{|b|^2}{2} \leq 0$ pointwise$\}$

and $F(a,b) = \begin{cases}\infty & \text{if } (a,b)\notin K\\ 0 & \text{if } (a,b)\in K\end{cases}$

a.
$$= \inf_{\rho,m}\sup_{\varphi,(a,b)}\left(-F(a,b) - G(\varphi) + \int_0^T\int_D a\rho + b\cdot m - \partial_t\varphi\rho - \nabla_x\varphi\cdot m\,dx\,dt\right)$$
$$= \inf_{\mu}\sup_{\varphi,q}\left(-F(q) - G(\varphi) - \langle\mu, \nabla_{t,x}\varphi - q\rangle\right)$$
for

i. $G(\varphi) = \int_D \varphi(0,x)\rho_0(x) - \varphi(T,x)\rho_T(x)\,dx$

ii. $\mu = (\rho, m)$

iii. $q = (a,b)$

iv. $\langle\mu, q\rangle = \iint_D \mu\cdot q\,dx\,dt$
$$= -\sup_{\mu}\inf_{\varphi,q}\left(F(q) + G(\varphi) + \langle\mu, \nabla_{t,x}\varphi - q\rangle\right)$$
6. now, we can change to
$$L_r(\varphi,q,\mu) = F(q) + G(\varphi) + \langle\mu, \nabla_{t,x}\varphi - q\rangle + \frac{r}{2}\langle\nabla_{t,x}\varphi - q, \nabla_{t,x}\varphi - q\rangle$$
w/out changing where optima occur.

7. see clearer notes for alg.:

11. Call a, think of μ as Lagrange multiplier and $\nabla_{t,x} \varphi = q$ as constraint

12. Augmented Lagrangian:

$$L_r(\varphi, q, \mu) = F(q) + G(\varphi) + \langle \mu, \nabla_{t,x}\varphi - q \rangle + \frac{r}{2}\langle \nabla_{t,x}\varphi - q, \nabla_{t,x}\varphi - q \rangle$$

13. then look for

$$\sup_{\mu} \inf_{\varphi, q} L_r(\varphi, q, \mu)$$

14. The algorithm

a. Find φ^n s.t.

$$L_r(\varphi^n, q^{n-1}, \mu^n) \le L_r(\varphi, q^{n-1}, \mu^n) \quad \forall \varphi$$

b. Find q^n s.t.

$$L_r(\varphi^n, q^n, \mu^n) \le L_r(\varphi^n, q, \mu^n) \quad \forall q$$

c. update

$$\mu^{n+1} = \mu^n + r(\nabla_{t,x}\varphi^n - q^n)$$

15. How do get φ^n?

a. need

$$L_r(\varphi + \delta\varphi, q, \mu) = L(\varphi, q, \mu) + G(\delta\varphi) + \langle \mu, \nabla_{t,x}\delta\varphi \rangle + r\langle \nabla_{t,x}\varphi - q, \nabla_{t,x}\delta\varphi \rangle + O(\delta\varphi^2)$$

b. so want φ^n s.t.

$$G(\delta\varphi) + \langle \mu, \nabla_{t,x}\delta\varphi \rangle + r\langle \nabla_{t,x}\varphi - q, \nabla_{t,x}\delta\varphi \rangle = 0 \quad \forall \delta\varphi$$

$$\cancel{G(\delta\varphi)} - \int_0^T\int \nabla_{t,x}\cdot\mu \, \delta\varphi \, dx\,dt - r\int_0^T\int \nabla_{t,x}\cdot(\nabla_{t,x}\varphi - q)\,\delta\varphi \, dx\,dt$$
$$+ \int_\Omega \rho_T^n \delta\varphi(x,T) - \rho_0^n \delta\varphi(x,0)\,dx + r\int_\Omega (\partial_t\varphi(x,T) - a^{n-1}(x,T))\delta\varphi(x,T) - (\partial_t\varphi(x,0) - a^n(x,0))\delta\varphi(x,0)\,dx$$

c. so we have

$$\begin{cases} -r\Delta_{t,x}\varphi^n = \nabla_{t,x}\cdot(\mu^n - r q^{n-1}) \\ r\partial_t\varphi^n(0,\cdot) = \rho_0 - \rho^n(0,\cdot) + r a^{n-1}(0,\cdot) \\ r\partial_t\varphi^n(T,\cdot) = \rho_T - \rho^n(T,\cdot) + r a^{n-1}(T,\cdot) \end{cases}$$

d. can add $\varepsilon \varphi^n$ onto LHS here.

16. How to get q^n?

a. $$\min_q \left(F(q) + \frac{r}{2}\langle \nabla_{t,x}\varphi^n - q, \nabla_{t,x}\varphi^n - q \rangle + \langle \mu^n, \nabla_{t,x}\varphi^n - q \rangle\right)$$

$$= \min_{q\in K}\left\{\langle \nabla_{t,x}\varphi^n + \frac{2\mu^n}{r} - q, \nabla_{t,x}\varphi^n - q \rangle\right\}$$

$$= \min_{q\in K}\left\{\langle \nabla_{t,x}\varphi^n + \frac{2\mu^n}{r} - q, \nabla_{t,x}\varphi^n + \frac{\mu^n}{r} - \frac{\mu^n}{r} - q \rangle\right\}$$

$$\Leftarrow \min_{q\in K}\left\{\langle \nabla_{t,x}\varphi^n + \frac{\mu^n}{r} - q, \nabla_{t,x}\varphi^n + \frac{\mu^n}{r} - q \rangle\right.$$

$-\langle \nabla_{\#x} e^n + \frac{\mu^n}{r} - q, \frac{\mu^n}{r} \rangle + \langle \frac{\mu^n}{r}, \nabla_{\#x} e^n + \frac{\mu^n}{r} - q \rangle$

$+ \langle \frac{\mu^n}{r}, \frac{\mu^n}{r} \rangle$

$= \inf_{q \in K} \{ \langle \nabla_{\#x} e^n + \frac{\mu^n}{r} - q, \nabla_{\#x} e^n + \frac{\mu^n}{r} - q \rangle$

$- \langle \frac{\mu^n}{r}, \frac{\mu^n}{r} \rangle \}$

can ignore since doesn't depend on q, so doesn't change ~~when~~ minimum occurs

I.e. 6-b-cd this is ~~equivalent~~ to just

$\inf \{ (\partial_t e^n + \frac{\rho^n}{r} - a)^2 + |\nabla_x e^n + \frac{m^n}{r} - b|^2 \mid a + \frac{|b|^2}{2} \le 0 \}$

i just use defn of q, M, K.

16 - how many steps to use?

well, $res^n = \partial_t e^n + \frac{|\nabla_x e^n|^2}{2}$ should go to 0

but let's normalize it:

$\text{crit}^n = \sqrt{\frac{\int_0^1 \int_\Omega \rho^n |res^n|}{\int_0^1 \int_\Omega \rho^n |\nabla_x e^n|^2}}$