Sum-of-Squares Certificates for Ma-Trudinger-Wang Regularity in Optimal Transport

Sachin Shivakumar[1], Georgiy A. Bondar[2], Gabriel Khan[3], Abhishek Halder[3]

[1] Los Alamos National Laboratory
[2] UC Santa Cruz
[3] Iowa State University

A Brief History of OT

What is optimal transport?

The Monge problem

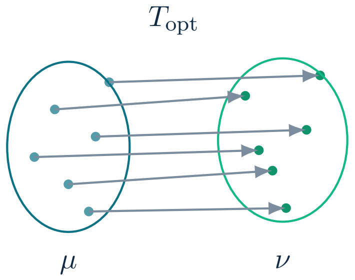
Given a source distribution μ and a target distribution ν, find a map

$$T : \mathcal{X} \to \mathcal{Y}, \qquad T_{\#}\mu = \nu,$$

that transports μ to ν at minimum cost:

$$\min_{T} \int_{\mathcal{X}} c(x, T(x))\, d\mu(x).$$

Interpretation: move mass from one distribution to another while paying a transportation cost $c(x, y)$.



Dormant till 20^{th} century

Tosltoi, Kantorovich, Hitchcock, Brenier, etc. picked it up.

Where is optimal transport useful?

Data, images, and learning

- **Imaging:** color transfer, morphing, registration.
- **Machine learning:** distributional distances, generative models, domain adaptation.
- **Statistics:** comparing probability distributions robustly.

Economics and allocation

- allocating resources to demand.
- market equilibrium and assignment problems.

PDEs and physics

- fluid motion and density evolution.
- gradient flows.
- diffusion and aggregation models.

Geometry

- curvature-dimension inequalities.
- comparison geometry.
- geometric and functional inequalities.

Many numerical methods to solve OT: relaxation, regularization, etc.

What about analysis of the solutions?

Exitense/Uniqueness

Comes from twist, non-degeneracy, c-convexity conditions, μ-a.c.

Regularity is stability

A regular optimal map sends nearby points to nearby points:

$$x \approx x' \qquad \Longrightarrow \qquad T_{\text{opt}}(x) \approx T_{\text{opt}}(x').$$

This means the map does not tear, fold, collapse, or create singular behavior.

Why this matters in practice

- stable predictions.
- interpretable maps.
- reliable numerical schemes and training efficiency.

Why this matters mathematically

Regularity turns an optimizer into a structured geometric object, allowing PDE, curvature, and variational tools to apply.

When is T_{opt} continuous, differentiable, or smooth? MTW might answer!!

MTW positivity guarantees regularity, but is not tractable

MTW tensor

Given $c(x, y)$, define

$$C(x,y) := D^2_{xy}c(x,y), \quad C^{ij}(x,y) := \left(C(x,y)^{-1}\right)_{ij}.$$

The Ma–Trudinger–Wang tensor is

$$S_{(x,y)}(\xi,\eta) := \Big(-c_{ij,kl} + c_{ij,r}\, C^{rs}\, c_{s,kl}\Big)\xi^i\xi^j\eta^k\eta^l,$$

where

$$c_{ij,kl} = \frac{\partial^4 c}{\partial x_i \partial x_j \partial y_k \partial y_l}, \qquad c_{ij,r} = \frac{\partial^3 c}{\partial x_i \partial x_j \partial y_r}.$$

MTW condition

For orthogonal directions $\eta^\top \xi = 0$, one asks for

$$S_{(x,y)}(\xi,\eta) \geq \kappa \|\xi\|^2 \|\eta\|^2.$$

The Challenge

For a new non-Euclidean cost, verifying nonnegativity of S is often a functional inequality problem.

Some Examples of Interest

Manifold	$c(x,y)$	Application
$\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$	$-\log \\|x-y\\|$	Reflector Antenna
$\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$	$b_1 - \sqrt{b_2 + b_3 \\|x-y\\|^2}$	Reflector Antenna
$\mathbb{R}^{n} \times \mathbb{S}^{n-1}$	$-\langle x, y\rangle / \\|x\\|$	Text-to-image Generation
$\mathbb{H}^{n} \times \mathbb{H}^{n}$	$-\cosh \circ d_{\mathbb{H}^n}(x,y) = -\left(1 + 2\frac{\\|x-y\\|^2}{(1-\\|x\\|^2)(1-\\|y\\|^2)}\right)$	Word Embeddings
$\mathbb{R}^{n} \times \mathbb{R}^{n}$	$\inf_{\gamma \in \mathcal{C}(x,y)} \int_0^1 \mathcal{L}(\gamma_t, \dot{\gamma}_t)\, \mathrm{d}t$	Diffusion Models
Unknown	$\frac{1}{2} d^2(x,y)$ learnt from data	??
$\mathbb{R}^{n} \times \mathbb{R}^{n}$	$\log(1 + \sum_{i=1}^{n} \exp(x_i - y_i))$	Information Theory

Two regularity certificates: NNCC and MTW

NNCC

$$S_{(x,y)}(\xi,\eta) \geq 0$$

$$\forall (x,y) \in (\mathcal{X}\times\mathcal{Y}), \ \ \forall\, \xi,\eta.$$

- Stronger algebraic condition.
- No orthogonality constraint.
- Can be checked through PSD structure of a matrix representation.

MTW

$$S_{(x,y)}(\xi,\eta) \geq \kappa \left\|\xi\right\|^2 \left\|\eta\right\|^2$$

$$\forall (x,y) \in (\mathcal{X}\times\mathcal{Y}), \ \ \forall \xi,\eta, \ \eta^\top \xi = 0.$$

- Weak condition has $\kappa = 0$.
- Orthogonality makes the algebra subtler.
- Still SOS-checkable using an equality multiplier.

Sum-of-squares relaxation of non-negativity conditions

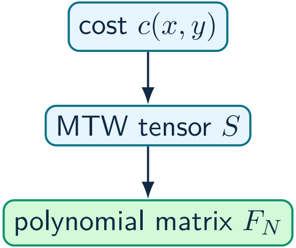
MTW positivity is polynomial positivity for rational cost functions

Represent the MTW tensor as a biquadratic form

For many costs, the tensor can be written as

$$S_{(x,y)}(\xi, \eta) = (\xi \otimes \eta)^\top F(x,y)(\xi \otimes \eta), \qquad F(x,y) = \frac{F_N(x,y)}{F_D(x,y)}.$$

If F_N, F_D are polynomial and $F_D > 0$, nonnegativity of S becomes polynomial nonnegativity.



SOS relaxation of polynomial positivity certificate

Polynomial nonnegativity

Checking $p(z) \geq 0$ globally or on a set is hard in general.

SOS sufficient certificate

If

$$p(z) = \sum_i q_i(z)^2,$$

then $p(z) \geq 0$ everywhere.

What about $\mathcal{X} \times \mathcal{Y}$?

SDP representation

Let $Z_d(z)$ collect monomials up to degree d.

$$p(z) = Z_d(z)^\top Q Z_d(z), \qquad Q \succeq 0.$$

Matching coefficients gives linear constraints; $Q \succeq 0$ gives an SDP.

Local non-negativity certificates using Positivstellansatz

Domain model

$$\mathcal{X} \times \mathcal{Y} = \{(x, y) : m_i(x, y) \leq 0,\ i = 1, \ldots, \ell\}.$$

Putinar-style certificate

To prove $p(z) \geq 0$ on $m_i(z) \leq 0$, seek SOS multipliers s_i such that

$$p(z) = s_0(z) - \sum_{i=1}^{\ell} s_i(z) m_i(z), \qquad s_i \in \Sigma[z].$$

This requires p-compactness of the set, but that can be enforced by adding a large, but finite ball constraint.

What if positivity certificates are unprovable?

Forward problem

Given $c, \mathcal{X}, \mathcal{Y}$, verify whether the cost satisfies

$$\mathrm{NNCC}, \quad \mathrm{MTW}(0), \quad \text{or} \quad \mathrm{MTW}(\kappa)$$

on all of $\mathcal{X} \times \mathcal{Y}$.

Inverse problem

Given $c, \mathcal{X}, \mathcal{Y}$, find a large semialgebraic region

$$\mathcal{U} \times \mathcal{V} \subseteq \mathcal{X} \times \mathcal{Y}$$

where the desired condition holds locally.

global certificate → if global fails → largest certified region

Forward problem: A global certificate

NNCC certificate

If there exist SOS multipliers s_i such that

$$\big(F_N(x,y) + F_N^\top(x,y)\big) - s_0(x,y)F_D(x,y) + \sum_{i\in\ell} s_i(x,y)m_i(x,y) \in \Sigma^{n^2}[x,y],$$

then c satisfies NNCC on $\mathcal{X}\times\mathcal{Y}$.

MTW(κ) certificate

If there exist SOS multipliers s_i and an equality multiplier t such that

$$\begin{aligned}(\xi\otimes\eta)^\top F_N(x,y)(\xi\otimes\eta) - \kappa F_D(x,y)\left\|\xi\right\|^2\left\|\eta\right\|^2 \\ + \sum_i s_i(x,y,\xi,\eta)m_i(x,y) + t(x,y,\xi,\eta)\eta^\top\xi \in \Sigma[x,y,\xi,\eta],\end{aligned}$$

then c satisfies $\mathrm{MTW}(\kappa)$.

Inverse problem: local regularity certificate

Parameterize the certified region

$$\mathcal{U} \times \mathcal{V} = \{(x, y) \in \Lambda : V(x, y) \leq 0\},$$

where Λ is compact and V is a polynomial function (decision variable).

Original Problem

$$\begin{aligned} \underset{\mathcal{U}\times\mathcal{V}\subseteq\mathcal{X}\times\mathcal{Y}}{\arg\max} \quad & \mathrm{vol}(\mathcal{U} \times \mathcal{V}) \\ \text{subject to} \quad & \mathfrak{S}_{(u,v)}(\xi, \eta) \geq 0, \quad \forall (u, v) \in \mathcal{U} \times \mathcal{V},\ \xi \in T_u\mathcal{U},\ \eta \in T_v^*\mathcal{V}. \end{aligned}$$

Sublevel-set Reformulation

$$\begin{aligned} \max_{V\in\mathbb{R}_d[x,y]} \quad & \mathrm{vol}\,\{(x, y) \mid V(x, y) \leq 0\} \\ \text{subject to} \quad & m_i(x, y) \leq V(x, y) \quad \forall\, i \in \ell. \\ & V(x, y) + \mathfrak{S}_{(x,y)}(\xi, \eta) \geq 0 \geq 0, \quad \forall (x, y) \in \mathcal{X} \times \mathcal{Y},\ \xi \in T_x\mathcal{X},\ \eta \in T_y^*\mathcal{Y}. \end{aligned}$$

Inverse problem: A proxy for volume

Parameterize the certified region

$$\mathcal{U} \times \mathcal{V} = \{(x, y) \in \Lambda : V(x, y) \leq 0\},$$

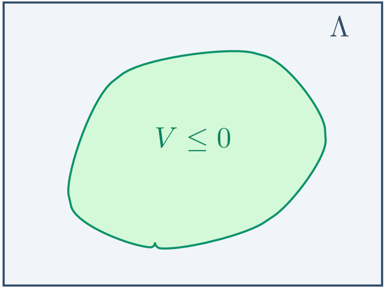
where Λ is a compact search box and V is a polynomial decision variable.

Volume surrogate

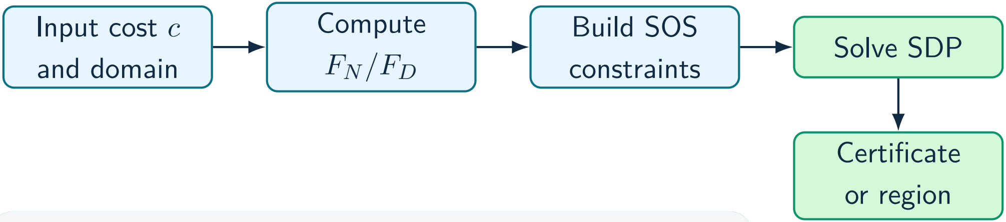
Instead of directly maximizing volume, minimize

$$\int_{\Lambda} V(x, y)\, dx dy$$

subject to SOS constraints that force the zero sublevel set to lie inside the regularity region.



Brief Summary: End-to-end workflow



Provable Guarantees: a feasible SOS program is a formal nonnegativity certificate under the modeling assumptions.

Can replace analytical verification!

Example 1: perturbed Euclidean cost

Cost

$$c(x, y) = \|x - y\|^2 - \varepsilon \|x - y\|^4, \qquad \varepsilon > 0.$$

- Use bisection on ε.
- For each ε, solve the relevant SOS feasibility problem.
- Recover analytically expected behavior.

Reported thresholds

n	1	2
$\varepsilon_{\max}$	0.67	0.67
Residual	$1.19 \cdot 10^{-7}$	$4.18 \cdot 10^{-7}$

Example 2: log-partition costs

Cost family

$$c(x, y) = \Psi_{\text{IsoMulNor}}(x - y),$$

$$\Psi(x) = \frac{n}{2}\left(-\log x_1 + \sum_{i=2}^{n} \frac{x_i^2}{x_1}\right).$$

SOS runs

n	3	4	5	6
Residual	10^{-7}	10^{-8}	10^{-8}	10^{-11}
Time	.72s	.81s	1.25s	1.67s

Even when the cost itself is not rational, the MTW expression can become rational. May require symbolic manipulation!

Example 3: inverse region for perturbed Euclidean cost

Setup

The same perturbed Euclidean cost may fail globally for larger ε, but still satisfy MTW locally.

What the inverse SOS returns

A degree-14 polynomial sublevel set $V \leq 0$ certifying a region where the MTW tensor is nonnegative.

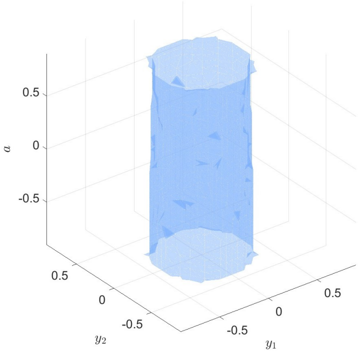


Figure 1: Inner approximation

Example 4: positive-curvature surface cost

$$c(x,y) = 3(x_1 - y_1)^2(x_2 + y_2) + 4(x_2^3 + y_2^3) - (4x_2y_2 - (x_1 - y_1)^2)^{\frac{3}{2}}$$

Non-Euclidean geometry

A scaled squared-distance cost induced by an incomplete Riemannian metric

$$ds^2 = x_2(dx_1^2 + dx_2^2)$$

has positive Gaussian curvature.

- Symmetries reduce the computation.
- Change of variable removes square-root.
- Systematic method in [1].

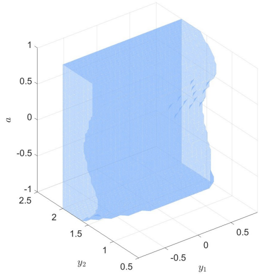


Figure 2: Inner Approximation

The method handles costs beyond simple rational examples.

Some Insights

What we know about NNCC and MTW complexities:

- NNCC: Sub-quadratic in constraints, polynomial in cost degree, exponential in dimensions
- MTW: Sub-quadratic in constraints, polynomial in cost degree, exponential in dimensions (naive case is worse than NNCC but has more symmetries)

Imposing additional symmetries via additional manifold constraints can be effective. Requires intuition and expertise.

Limitations and practical caveats

Computational scaling

SOS programs grow quickly with dimension, polynomial degree, and number of constraints.

Sufficient vs necessary

Failure of a chosen SOS relaxation does not always mean the condition is false. A failure is not as informative as a success!

Non-trivial Modeling choices

The certificate depends on the domain, rational representation, degree bounds, and local parameterization.

Inverse approximation: Volume proxy

Inner approximation is not necessarily the maximal true regularity region for any given polynomial. Choice of volume proxy is critical!

- Papachristodoulou, Antonis, and Stephen Prajna.
  ***Analysis of non-polynomial systems using the sum of squares decomposition*.**
  Positive polynomials in control. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005. 23-43.
- Sachin Shivakumar, Georgiy A. Bondar, Gabriel Khan, Abhishek Halder.
  ***Sum-of-Squares Programming for Ma-Trudinger-Wang Regularity of Optimal Transport Maps*.**
  arXiv:2412.13372, 2024.

Xi-Nan Ma, Neil Trudinger, Xu-Jia Wang.
***Regularity of potential functions of the optimal transportation problem*.**
Archive for Rational Mechanics and Analysis, 2005.

Pablo Parrilo.
***Semidefinite programming relaxations for semialgebraic problems*.**
Mathematical Programming, 2003.

Mihai Putinar.
***Positive polynomials on compact semi-algebraic sets*.**
Indiana University Mathematics Journal, 1993.

Johan Löfberg.
***YALMIP: A toolbox for modeling and optimization in MATLAB*.**
CACSD, 2004.