Convex Relaxations for Manifold-Valued Markov Random Fields With Approximation Guarantees

Robin Kenis
robin.kenis@kuleuven.be
KU Leuven

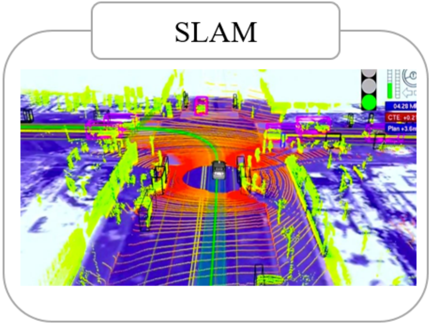
Emanuel Laude
emanuel.laude@kuleuven.be
KU Leuven

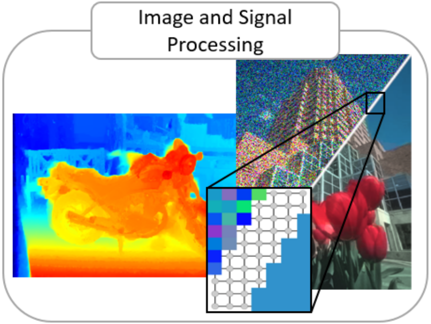
Panagiotis Patrinos
panos.patrinos@esat.kuleuven.be
KU Leuven

ECCV
October 4, 2024

MAP-MRF

$$\underset{x \in \Omega^{\mathcal{V}}}{\text{minimize}} \left\{ F(x) = \sum_{u \in \mathcal{V}} f_u(x_u) + \sum_{(u,v) \in \mathcal{E}} f_{(u,v)}(x_u, x_v) \right\}$$



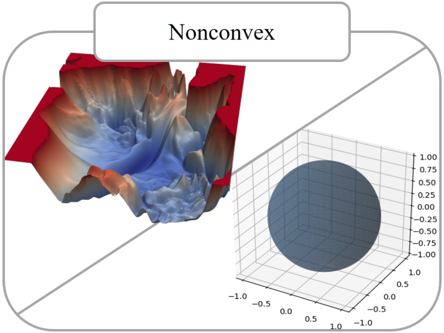


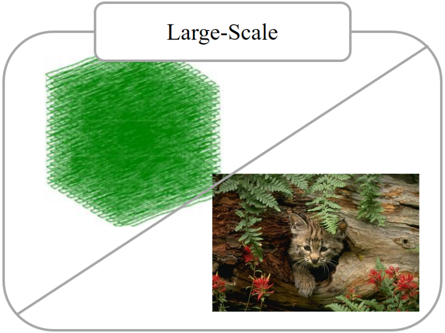
Maximum A Posteriori Inference for Markov Random Fields

MAP-MRF Problems are Nonconvex and Large-Scale

MAP-MRF

$$\underset{x \in \Omega^{\mathcal{V}}}{\text{minimize}} \left\{ F(x) = \sum_{u \in \mathcal{V}} f_u(x_u) + \sum_{(u,v) \in \mathcal{E}} f_{(u,v)}(x_u, x_v) \right\}$$





Maximum A Posteriori Inference for Markov Random Fields

The Local Marginal Polytope Relaxation is the Key to Scalable Convex Lifting

MAP-MRF

$$\underset{x \in \Omega^{\mathcal{V}}}{\text{minimize}} \left\{ F(x) = \sum_{u \in \mathcal{V}} f_u(x_u) + \sum_{(u,v) \in \mathcal{E}} f_{(u,v)}(x_u, x_v) \right\}$$

Global Marginal Polytope (GMP)

$$\min_{\mu \in \mathcal{P}(\Omega^{\mathcal{V}})} \langle F, \mu \rangle \quad = \quad \min_{x \in \Omega^{\mathcal{V}}} F(x)$$

Maximum A Posteriori Inference for Markov Random Fields

The Local Marginal Polytope Relaxation is the Key to Scalable Convex Lifting

MAP-MRF

$$\underset{x \in \Omega^{\mathcal{V}}}{\text{minimize}} \left\{ F(x) = \sum_{u \in \mathcal{V}} f_u(x_u) + \sum_{(u,v) \in \mathcal{E}} f_{(u,v)}(x_u, x_v) \right\}$$

Global Marginal Polytope (GMP)

$$\min_{\mu \in \mathcal{P}(\Omega^{\mathcal{V}})} \langle F, \mu \rangle \quad = \quad \min_{x \in \Omega^{\mathcal{V}}} F(x)$$

$$\text{IV}$$

Local Marginal Polytope (LMP)

$$\min_{\mu \in \mathcal{P}(\Omega)^{\mathcal{V}}} \sum_{u \in \mathcal{V}} \langle f_u, \mu_u \rangle + \sum_{(u,v) \in \mathcal{E}} \mathrm{OT}_{f_{(u,v)}}(\mu_u, \mu_v)$$

Maximum A Posteriori Inference for Markov Random Fields

The Local Marginal Polytope Relaxation is the Key to Scalable Convex Lifting

MAP-MRF

$$\underset{x\in\Omega^{\mathcal{V}}}{\text{minimize}}\left\{F(x)=\sum_{u\in\mathcal{V}} f_u(x_u)+\sum_{(u,v)\in\mathcal{E}} f_{(u,v)}(x_u,x_v)\right\}$$

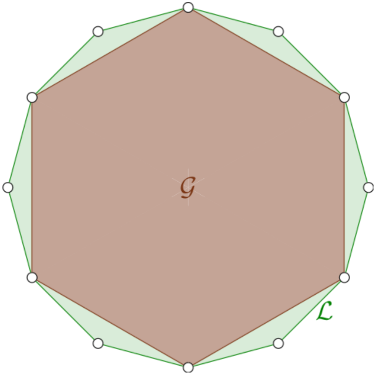
Global Marginal Polytope (GMP)

$$\min_{\mu\in\mathcal{P}(\Omega^{\mathcal{V}})} \langle F,\mu\rangle \quad = \quad \min_{x\in\Omega^{\mathcal{V}}} F(x)$$

$$\mathbin{\rotatebox[origin=c]{90}{$\geq$}}$$

Local Marginal Polytope (LMP)

$$\min_{\mu\in\mathcal{P}(\Omega)^{\mathcal{V}}} \sum_{u\in\mathcal{V}}\langle f_u,\mu_u\rangle + \sum_{(u,v)\in\mathcal{E}} \mathrm{OT}_{f_{(u,v)}}(\mu_u,\mu_v)$$



From Optimal Transport Theory to Polynomial Optimization

We rewrite the relaxation as an optimization problem over continuous functions

Local Marginal Polytope - Geodesic Penalty

$$\min_{\mu \in \mathcal{P}(\Omega)^{\mathcal{V}}} \left\{ \quad \mathcal{P}(\mu) = \sum_{u \in \mathcal{V}} \langle f_u, \mu_u \rangle + \sum_{(u,v) \in \mathcal{E}} \mathcal{W}_1(\mu_u, \mu_v) \quad \right\}$$

Kantorovich-Rubinstein Theorem

$$\mathcal{W}_1(\mu_u, \mu_v) = \max_{\phi_{(u,v)} \in \text{lip}(\Omega)} \langle \phi_{(u,v)}, \mu_v - \mu_u \rangle$$

From Optimal Transport Theory to Polynomial Optimization

We rewrite the relaxation as an optimization problem over continuous functions

Local Marginal Polytope - Geodesic Penalty

$$\min_{\mu \in \mathcal{P}(\Omega)^{\mathcal{V}}} \max_{\phi \in \mathrm{lip}(\Omega)^{\mathcal{E}}} \left\{ \mathcal{L}(\mu, \phi) = \sum_{u \in \mathcal{V}} \langle f_u, \mu_u \rangle + \sum_{(u,v) \in \mathcal{E}} \langle \phi_{(u,v)}, \mu_v - \mu_u \rangle \right\}$$

Kantorovich-Rubinstein Theorem

$$\mathcal{W}_1(\mu_u, \mu_v) = \max_{\phi_{(u,v)} \in \mathrm{lip}(\Omega)} \langle \phi_{(u,v)}, \mu_v - \mu_u \rangle$$

From Optimal Transport Theory to Polynomial Optimization

We rewrite the relaxation as an optimization problem over continuous functions

Local Marginal Polytope - Geodesic Penalty

$$\max_{\phi \in \mathrm{lip}(\Omega)^{\mathcal{E}}} \left\{ \quad \mathcal{D}(\phi) = -\sum_{u \in \mathcal{V}} \sigma_{\mathcal{P}(\Omega)} \left(\nabla_{\mathcal{G}} \cdot \phi - f_u \right) \quad \right\}$$

Kantorovich-Rubinstein Theorem

$$\mathcal{W}_1(\mu_u, \mu_v) = \max_{\phi_{(u,v)} \in \mathrm{lip}(\Omega)} \langle \phi_{(u,v)}, \mu_v - \mu_u \rangle$$

From Optimal Transport Theory to Polynomial Optimization

We rewrite the relaxation as an optimization problem over continuous functions

Local Marginal Polytope - Geodesic Penalty

$$\max_{\phi \in \mathrm{lip}(\Omega)^{\mathcal{E}}} \left\{ \quad \mathcal{D}(\phi) = -\sum_{u \in \mathcal{V}} \sigma_{\mathcal{P}(\Omega)} \left(\nabla_{\mathcal{G}} \cdot \phi - f_u \right) \quad \right\}$$

Kantorovich-Rubinstein Theorem

$$\mathcal{W}_1(\mu_u, \mu_v) = \max_{\phi_{(u,v)} \in \mathrm{lip}(\Omega)} \langle \phi_{(u,v)}, \mu_v - \mu_u \rangle$$

$$\sup_{\varphi \in \mathbb{R}^n} \mathcal{D}(\varphi) \leq \sup_{\varphi \in \mathbb{R}[x]_{n-1}} \mathcal{D}(\varphi) \leq \sup_{\varphi \in \mathbb{R}[x]_n} \mathcal{D}(\varphi) \leq \sup_{\varphi \in \mathbb{R}[x]} \mathcal{D}(\varphi) \leq \sup_{\varphi \in C} \mathcal{D}(\varphi)$$

$$\text{subject to} \quad \varphi \in \mathrm{lip}_d(\Omega)^{\mathcal{E}}$$

From Optimal Transport Theory to Polynomial Optimization

We rewrite the relaxation as an optimization problem over continuous functions

Local Marginal Polytope - Geodesic Penalty

$$\max_{\phi \in \mathrm{lip}(\Omega)^{\mathcal{E}}} \left\{ \quad \mathcal{D}(\phi) = -\sum_{u \in \mathcal{V}} \sigma_{\mathcal{P}(\Omega)} \left(\nabla_{\mathcal{G}} \cdot \phi - f_u \right) \quad \right\}$$

Kantorovich-Rubinstein Theorem

$$\mathcal{W}_1(\mu_u, \mu_v) = \max_{\phi_{(u,v)} \in \mathrm{lip}(\Omega)} \langle \phi_{(u,v)}, \mu_v - \mu_u \rangle$$

$$\sup_{\varphi \in \mathbb{R}^n} \mathcal{D}(\varphi) \leq \sup_{\varphi \in \mathbb{R}[x]_{n-1}} \mathcal{D}(\varphi) \leq \sup_{\varphi \in \mathbb{R}[x]_n} \mathcal{D}(\varphi) \leq \sup_{\varphi \in \mathbb{R}[x]} \mathcal{D}(\varphi) \leq \sup_{\varphi \in C} \mathcal{D}(\varphi)$$

$$\text{subject to} \quad \varphi \in \mathrm{lip}_d(\Omega)^{\mathcal{E}}$$

✓ Finite Dimensional

From Optimal Transport Theory to Polynomial Optimization

We rewrite the relaxation as an optimization problem over continuous functions

Local Marginal Polytope - Geodesic Penalty

$$\max_{\phi \in \mathrm{lip}(\Omega)^{\mathcal{E}}} \left\{ \quad \mathcal{D}(\phi) = -\sum_{u \in \mathcal{V}} \sigma_{\mathcal{P}(\Omega)} \left(\nabla_{\mathcal{G}} \cdot \phi - f_u \right) \quad \right\}$$

Kantorovich-Rubinstein Theorem

$$\mathcal{W}_1(\mu_u, \mu_v) = \max_{\phi_{(u,v)} \in \mathrm{lip}(\Omega)} \langle \phi_{(u,v)}, \mu_v - \mu_u \rangle$$

$$\sup_{\varphi \in \mathbb{R}^n} \mathcal{D}(\varphi) \leq \sup_{\varphi \in \mathbb{R}[x]_{n-1}} \mathcal{D}(\varphi) \leq \sup_{\varphi \in \mathbb{R}[x]_n} \mathcal{D}(\varphi) \leq \sup_{\varphi \in \mathbb{R}[x]} \mathcal{D}(\varphi) \leq \sup_{\varphi \in C} \mathcal{D}(\varphi)$$

$$\text{subject to} \quad \varphi \in \mathrm{lip}_d(\Omega)^{\mathcal{E}}$$

- ✓ Finite Dimensional
- ✓ Increasingly tight

From Optimal Transport Theory to Polynomial Optimization

We rewrite the relaxation as an optimization problem over continuous functions

Local Marginal Polytope - Geodesic Penalty

$$\max_{\phi \in \mathrm{lip}(\Omega)^{\mathcal{E}}} \left\{ \quad \mathcal{D}(\phi) = -\sum_{u \in \mathcal{V}} \sigma_{\mathcal{P}(\Omega)}\left(\nabla_{\mathcal{G}} \cdot \phi - f_u\right) \quad \right\}$$

Kantorovich-Rubinstein Theorem

$$\mathcal{W}_1(\mu_u, \mu_v) = \max_{\phi_{(u,v)} \in \mathrm{lip}(\Omega)} \langle \phi_{(u,v)}, \mu_v - \mu_u \rangle$$

$$\sup_{\varphi \in \mathbb{R}^n} \mathcal{D}(\varphi) \leq \sup_{\varphi \in \mathbb{R}[x]_{n-1}} \mathcal{D}(\varphi) \leq \sup_{\varphi \in \mathbb{R}[x]_n} \mathcal{D}(\varphi) \leq \sup_{\varphi \in \mathbb{R}[x]} \mathcal{D}(\varphi) \overset{?}{=} \sup_{\varphi \in C} \mathcal{D}(\varphi)$$

$$\text{subject to} \quad \varphi \in \mathrm{lip}_d(\Omega)^{\mathcal{E}}$$

- ✓ Finite Dimensional
- ✓ Increasingly tight
- ✗ No convergence guarantee

From Optimal Transport Theory to Polynomial Optimization

We rewrite the relaxation as an optimization problem over continuous functions

Local Marginal Polytope - Geodesic Penalty

$$\max_{\phi \in \mathrm{lip}(\Omega)^{\mathcal{E}}} \left\{ \mathcal{D}(\phi) = -\sum_{u \in \mathcal{V}} \sigma_{\mathcal{P}(\Omega)} \left(\nabla_{\mathcal{G}} \cdot \phi - f_u \right) \right\}$$

Kantorovich-Rubinstein Theorem

$$\mathcal{W}_1(\mu_u, \mu_v) = \max_{\phi_{(u,v)} \in \mathrm{lip}(\Omega)} \langle \phi_{(u,v)}, \mu_v - \mu_u \rangle$$

$$\sup_{\varphi \in \mathbb{R}^n} \mathcal{D}(\varphi) \leq \sup_{\varphi \in \mathbb{R}[x]_{n-1}} \mathcal{D}(\varphi) \leq \sup_{\varphi \in \mathbb{R}[x]_n} \mathcal{D}(\varphi) \leq \sup_{\varphi \in \mathbb{R}[x]} \mathcal{D}(\varphi) \leq \sup_{\varphi \in C} \mathcal{D}(\varphi)$$

$$\text{subject to} \quad \varphi \in \mathrm{lip}_d(\Omega)^{\mathcal{E}}$$

- ✓ Finite Dimensional
- ✓ Increasingly tight
- ✗ No convergence guarantee

Convergence:

Let $\Omega = \mathcal{S}^m$, then

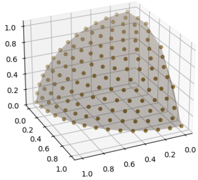
$$\lim_{n \to \infty} \max_{\mathbb{R}[x]_n} \mathcal{D}(\phi) = \max_{\phi \in C(\Omega)} \mathcal{D}(\phi) \qquad \text{subject to} \quad \phi \in \mathrm{lip}(\Omega)^{\mathcal{E}}.$$

Results Using the Semidefinite Relaxation

Our Convex Relaxation may yield Globally Optimal Results even for Low Stages of the Hierarchy

Geodesic TV denoising – Baseline

- Baseline: discretize Ω and solve the discrete MAP-MRF problem using specialized solvers [1, 2]



V. Kolmogorov.
Convergent tree-reweighted message passing for energy minimization.
In *International Workshop on Artificial Intelligence and Statistics*, pages 182–189. PMLR, 2005.

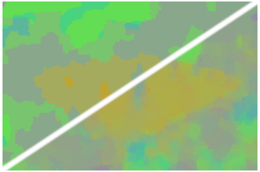
V. Kolmogorov.
Solving relaxations of map-mrf problems: Combinatorial in-face frank-wolfe directions.
In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11980–11989, 2023.

Results Using the Semidefinite Relaxation

Our Convex Relaxation may yield Globally Optimal Results even for Low Stages of the Hierarchy

Geodesic TV denoising – Baseline

- Baseline: discretize Ω and solve the discrete MAP-MRF problem using specialized solvers [1, 2]
- Baseline suffers from labeling bias



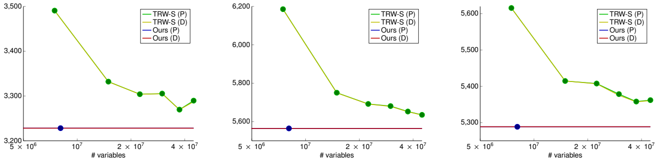
V. Kolmogorov.
Convergent tree-reweighted message passing for energy minimization.
In *International Workshop on Artificial Intelligence and Statistics*, pages 182–189. PMLR, 2005.

V. Kolmogorov.
Solving relaxations of map-mrf problems: Combinatorial in-face frank-wolfe directions.
In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11980–11989, 2023.

Results Using the Semidefinite Relaxation

Our Convex Relaxation may yield Globally Optimal Results even for Low Stages of the Hierarchy

Geodesic TV denoising – Problem Scaling



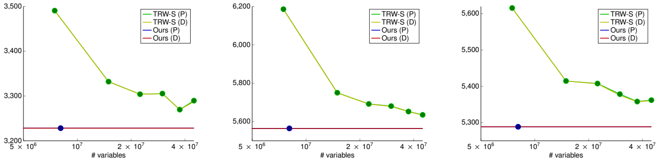
Geodesic TV Denoising – Timing Results

Time = 1000 s		Ours [2]	TRW-S [150]	TRW-S [200]	LR-Solver [150]	LR-Solver [200]
Exp. 2	P	3228.25	3288.05	3266.50	3287.44	3402.78
	D	3228.17	3287.37	3256.48	3287.44	2782.90
Exp. 2	P	5565.06	5635.22	5650.13	5634.27	5815.74
	D	5565.82	5634.19	5623.32	5634.24	5104.79
Exp. 3	P	5288.99	5362.19	5367.40	5360.76	5670.43
	D	5289.92	5360.68	5338.28	5360.75	3935.67

Results Using the Semidefinite Relaxation

Our Convex Relaxation may yield Globally Optimal Results even for Low Stages of the Hierarchy

Geodesic TV denoising – Problem Scaling



✓ Our method finds the global optimum even in the degree 2 case.

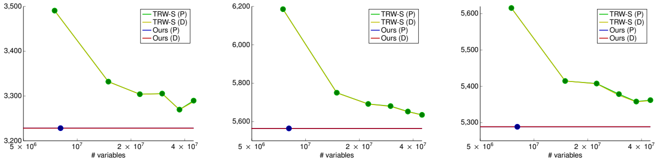
Geodesic TV Denoising – Timing Results

Time = 1000 s		Ours [2]	TRW-S [150]	TRW-S [200]	LR-Solver [150]	LR-Solver [200]
Exp. 2	*P*	3228.25	3288.05	3266.50	3287.44	3402.78
	D	3228.17	3287.37	3256.48	3287.44	2782.90
Exp. 2	*P*	5565.06	5635.22	5650.13	5634.27	5815.74
	D	5565.82	5634.19	5623.32	5634.24	5104.79
Exp. 3	*P*	5288.99	5362.19	5367.40	5360.76	5670.43
	D	5289.92	5360.68	5338.28	5360.75	3935.67

Results Using the Semidefinite Relaxation

Our Convex Relaxation may yield Globally Optimal Results even for Low Stages of the Hierarchy

Geodesic TV denoising – Problem Scaling



✓ Our method finds the global optimum even in the degree 2 case.

Geodesic TV Denoising – Timing Results

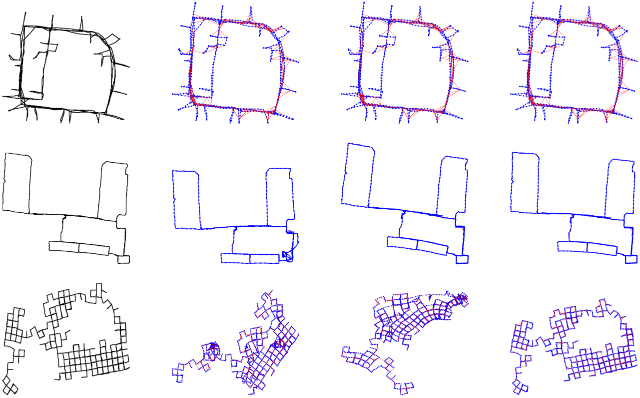
Time = 1000 s		Ours [2]	TRW-S [150]	TRW-S [200]	LR-Solver [150]	LR-Solver [200]
Exp. 2	*P*	3228.25	3288.05	3266.50	3287.44	3402.78
	D	3228.17	3287.37	3256.48	3287.44	2782.90
Exp. 2	*P*	5565.06	5635.22	5650.13	5634.27	5815.74
	D	5565.82	5634.19	5623.32	5634.24	5104.79
Exp. 3	*P*	5288.99	5362.19	5367.40	5360.76	5670.43
	D	5289.92	5360.68	5338.28	5360.75	3935.67

✓ Given a fixed time budget of 1000 seconds, our method significantly outperforms the baseline approach.

Results Using the Semidefinite Relaxation

Our Convex Relaxation Can Be Used to Obtain Good Initializers for Local Methods

Robotics – Pose Graph Optimization

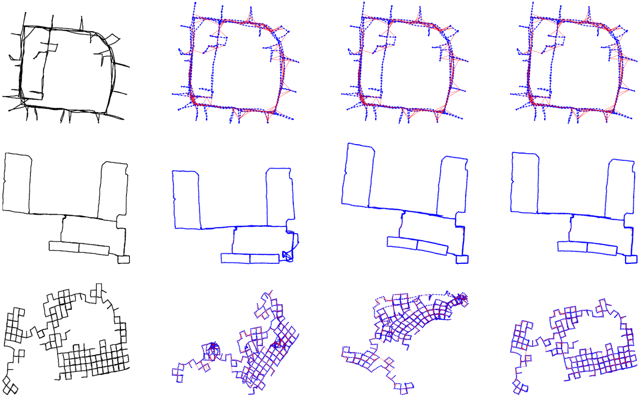


(a) Ground truth (b) Local (c) Ours (d) Ours+local

Results Using the Semidefinite Relaxation

Our Convex Relaxation Can Be Used to Obtain Good Initializers for Local Methods

Robotics – Pose Graph Optimization



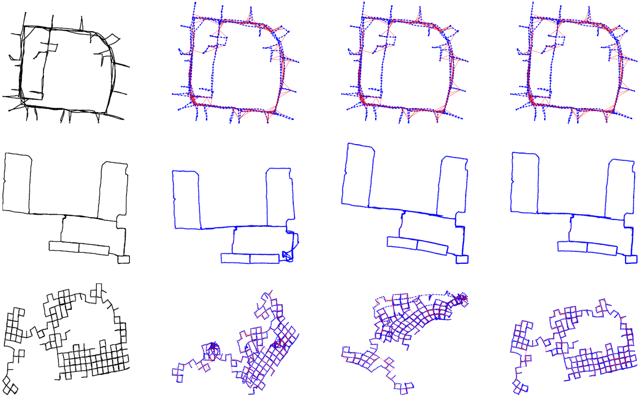
(a) Ground truth (b) Local (c) Ours (d) Ours+local

✓ Easy: Near optimal results without rounding

Results Using the Semidefinite Relaxation

Our Convex Relaxation Can Be Used to Obtain Good Initializers for Local Methods

Robotics – Pose Graph Optimization



(a) Ground truth (b) Local (c) Ours (d) Ours+local

- ✓ Easy: Near optimal results without rounding
- ✓ General: Optimal results after rounding

Thank you!