澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

Estimation Contracts for Outlier-Robust Geometric Perception

MAO Yuanwen
mc35295@umac.mo

2024.1.19

catalogue

1、Geometric perception

2、Numerical Experiments

3、Algorithm

4、Variables in geometric perception

01

Geometric perception

1.1 The concept of geometric perception

- Geometric perception： Estimating **unknown geometric models** (e.g., poses, rotations, 3D structure) from **sensor data** (e.g., camera images, lidar scans, inertial data, wheel odometry).
- Measurement models：

$$\boldsymbol{y}_i = f_i(\boldsymbol{x}^\circ) + \boldsymbol{\epsilon}, \quad \text{with} \quad \boldsymbol{y}_i \in \mathbb{R}^{d_y} \quad \text{and} \quad \boldsymbol{x}^\circ \in \mathbb{X} \subseteq \mathbb{R}^{d_x},$$

已知测量值 目标未知值 测量得到的噪声 向量值 x的域，比如3D姿态集

1.1 The concept of geometric perception

- ***The perception frontend**感知前端: y_i are the result of a pre-processing of the raw sensor data, from raw image pixels (typically using a neuralnetwork). prone to errors.

- **The perception back-end**感知后端: the estimation algorithms that compute x from the y_i, computes the object pose given the yi's.

1.2 Outlier-robust estimation

- **Consensus Maximization(MC)**: searches for the largest set of inliers such that the measurements selected as inliers have a low error with respect to some estimate

$$\boldsymbol{x}_{\mathrm{MC}} = \underset{\substack{\boldsymbol{\omega} \in \{0;1\}^n \\ \boldsymbol{x} \in \mathbb{X}}}{\arg\max} \sum_{i=1}^{n} \omega_i \,, \quad \text{subject to} \quad \omega_i \cdot \|\boldsymbol{y}_i - f_i(\boldsymbol{x})\|_2^2 \leq \bar{c}^2 \,,$$

The value of ω_i when the formula reaches the maximum value

the maximum error for a measurement to be considered an inlier

1.2 Outlier-robust estimation

- **The Least Trimmed Squares (LTS)**: statistical technique for estimation of unknown parameters of a linear regression model

$$
\begin{aligned}
\boldsymbol{x}_{\text{TLS}} &= \underset{\boldsymbol{x}\in\mathbb{X}}{\arg\min} \sum_{i=1}^{n} \min\left(\|\boldsymbol{y}_i - f_i(\boldsymbol{x})\|_2^2 \ , \ \bar{c}^2 \right) \\
&= \underset{\substack{\omega\in\{0;1\}^n \\ \boldsymbol{x}\in\mathbb{X}}}{\arg\min} \sum_{i=1}^{n} \omega_i \cdot \|\boldsymbol{y}_i - f_i(\boldsymbol{x})\|_2^2 + (1-\omega_i)\cdot \bar{c}^2 \ ,
\end{aligned}
$$

1.3 Moment Relaxations

- **Relaxation Method**松弛算法: Gradually approach the solution of the problem through repeated iterative calculations. 1:determining initial values. 2:apply update rules repeatedly. 3:setting conditions to a certain threshold

- polynomial optimization problem (POP)多项式优化问题

- **Lasserre** proposed a global optimization algorithm全局优化算法 for solving POP, which is based on semidefinite programming (SDP半定规划) relaxation and can infinitely approximate the global optimal value of the problem.

1.3 Moment Relaxations

- **Lasserre method**: (i) rewrite (POP) using the moment matrix X, (ii) relax the (non-convex) rank-1 constraint on X (and only enforce X>0), which is a convex constraint), (iii) add redundant constraints that are trivially satisfied in (POP) but still contribute to improving the quality of the relaxation.

$$m^\star \triangleq \min_{\boldsymbol{X} \in \mathbb{X}_{\mathrm{sdp}}} \{\langle \boldsymbol{C}, \boldsymbol{X} \rangle \mid \mathcal{A}(\boldsymbol{X}) = \boldsymbol{b},\ \boldsymbol{X} \succeq 0\}.$$

$$(\mathrm{POP}) \xrightarrow{\text{moment relaxation}} (\mathrm{LAS}_r) \xrightarrow{\text{SDP solver}} (\boldsymbol{X}^\star, m^\star) \xrightarrow{\text{rounding}} (\hat{\boldsymbol{x}}, \hat{p}) \xrightarrow{\text{certification}} \hat{p} \overset{?}{=} m^\star,$$

02

Variables in geometric perception

Fact 1 (Variables in geometric perception). *The d-dimensional* Special Orthogonal group $\mathrm{SO}(d) \triangleq \{\boldsymbol{R} \in \mathbb{R}^{d\times d} \mid \boldsymbol{R}^{\mathsf{T}}\boldsymbol{R} = \mathbf{I}_d, \det(\boldsymbol{R}) = +1\}$ *(i.e., the group of rotations), and the* Special Euclidean group $\mathrm{SE}(d) \triangleq \{\left[\begin{smallmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \mathbf{0} & 1 \end{smallmatrix}\right] \in \mathbb{R}^{(d+1)\times(d+1)} \mid \boldsymbol{R} \in \mathrm{SO}(d), \boldsymbol{t} \in \mathbb{R}^d\}$ *(i.e., the group of poses and rigid transformations) are basic semi-algebraic sets.*[6] *Moreover, several geometric constraints (e.g., field-of-view or maximum distance constraints) can be written as basic semi-algebraic sets.*

- **Rotation search**: Estimate the rotation R ϵ SO(3) that aligns pairs of 3D points (a_i, b_i), i=1,2,.....,n. The measurement model for the (inlier) measurements is given by:

$$\boldsymbol{b}_i = \boldsymbol{R}\boldsymbol{a}_i + \boldsymbol{\epsilon} \quad \xrightarrow{\boldsymbol{x}^\circ \triangleq vec(\boldsymbol{R})} \quad \overbrace{\boldsymbol{b}_i = (\boldsymbol{a}_i^{\mathsf{T}} \otimes \mathbf{I}_3)\boldsymbol{x}^\circ + \boldsymbol{\epsilon},}^{same\ form\ as\ (3)}$$

where ϵ is the measurement noise. we used the vectorization operator vec(·) to **transform a 3D matrix into a vector** and manipulated the expression using standard vectorization properties. Rotation search arises, for instance, in satellite attitude estimation and image stitching.

- **3D point cloud registration**:

Example 2 (3D point cloud registration). *Estimate the rigid transformation* $(\boldsymbol{R}, \boldsymbol{t})$, *with* $\boldsymbol{R} \in \mathrm{SO}(3)$ *and* $\boldsymbol{t} \in \mathbb{R}^3$, *that aligns pairs of 3D points* $(\boldsymbol{a}_i, \boldsymbol{b}_i)$, $i = 1, \ldots, n$. *The measurement model for the (inlier) measurements is:*

$$\boldsymbol{b}_i = \boldsymbol{R}\boldsymbol{a}_i + \boldsymbol{t} + \boldsymbol{\epsilon} \quad \xRightarrow{\boldsymbol{x}^\circ \triangleq \left[\begin{smallmatrix} \mathrm{vec}(\boldsymbol{R}) \\ \boldsymbol{t} \end{smallmatrix}\right]} \quad \overbrace{\boldsymbol{b}_i = \left[\begin{array}{cc} \boldsymbol{a}_i^\mathsf{T} \otimes \mathbf{I}_3 & \mathbf{I}_3 \end{array}\right] \boldsymbol{x}^\circ + \boldsymbol{\epsilon}}^{\textit{same form as } (3)},$$

Point-to-plane 3D registration can be similarly formulated using a linear model involving a rigid transformation [30]. Registration problems are commonly encountered in instance-level object pose estimation, scan-matching for 3D reconstruction, and (stereo or RGB-D) visual odometry [1].

03

Algorithm

Estimation Contracts for (LTS)

Algorithm 1: Moment relaxation for (LTS), version 1 [15].

Input: input data $(\boldsymbol{y}_i, \boldsymbol{A}_i)$, $i \in [n]$, inlier rate α, exponent k, relaxation order $r \geq k$.

Output: estimate of $\boldsymbol{x}^\circ$.

```
/* Algorithm solves a relaxation of the following (LTS) problem:    */
/*
```

$$\min_{\boldsymbol{\omega},\boldsymbol{x},\bar{\boldsymbol{y}}_i,\bar{\boldsymbol{A}}_i,i\in[n]} \left(\frac{1}{n}\sum_{i=1}^{n}\left\|\bar{\boldsymbol{y}}_i - \bar{\boldsymbol{A}}_i^{\mathsf{T}}\boldsymbol{x}\right\|_2^2\right)^{\frac{k}{2}} \quad \text{s.t.} \quad \mathcal{L}_{\boldsymbol{\omega},\boldsymbol{x}} \triangleq \left\{ \begin{array}{c} \omega_i^2 = \omega_i, \quad i \in [n] \\ \sum_{i=1}^{n} \omega_i = \alpha n \\ \omega_i \cdot (\bar{\boldsymbol{y}}_i - \boldsymbol{y}_i) = \boldsymbol{0} \quad i \in [n] \\ \omega_i \cdot (\bar{\boldsymbol{A}}_i - \boldsymbol{A}_i) = \boldsymbol{0} \quad i \in [n] \\ \boldsymbol{x} \in \mathbb{X} \end{array} \right\} \quad \text{(LTS1)}$$

```
/* Compute matrix X* by solving SDP resulting from moment relaxation    */
```

$\boldsymbol{X}^\star = \mathsf{solve\_moment\_relaxation\_at\_order\_}r\,\text{(LTS1)}$

```
/* Pick entries of X* corresponding to x    */
```

$\boldsymbol{x}_{\mathsf{lts-sdp1}} \triangleq \boldsymbol{X}^\star_{[\boldsymbol{x}]}$

return $\boldsymbol{x}_{\mathsf{lts-sdp1}}$.

Estimation Contract for (MC)

Algorithm 3: Moment relaxation for (MC).

Input: input data $(\boldsymbol{y}_i, \boldsymbol{A}_i)$, $i \in [n]$, relaxation order $r \geq 2$.

Output: estimate of $\boldsymbol{x}^\circ$.

`/* Algorithm solves a relaxation of the following` (MC) `problem: */`

`/*`

$$\max_{\boldsymbol{\omega}, \boldsymbol{x}} \sum_{i=1}^{n} \omega_i, \qquad \texttt{s.t.}\ \mathcal{M}_{\boldsymbol{\omega},\boldsymbol{x}} \triangleq \left\{ \begin{array}{c} \omega_i^2 = \omega_i, \quad i \in [n] \\ \omega_i \cdot \left\| \boldsymbol{y}_i - \boldsymbol{A}_i^{\mathsf{T}} \boldsymbol{x} \right\|_2^2 \leq \bar{c}^2 \quad i \in [n] \\ \boldsymbol{x} \in \mathbb{X} \end{array} \right\} \tag{MC1}$$

`/* Compute matrix` $\boldsymbol{X}^\star$ `by solving SDP resulting from moment relaxation */`

$\boldsymbol{X}^\star = \mathsf{solve\_moment\_relaxation\_at\_order\_}r\,(\mathrm{MC1})$

`/* Compute estimate */`

for each $i \in [n]$ set: $\boldsymbol{v}_i = \begin{cases} \dfrac{\boldsymbol{X}^\star_{[\omega_i \boldsymbol{x}]}}{\boldsymbol{X}^\star_{[\omega_i]}} & \text{if } \boldsymbol{X}^\star_{[\omega_i]} > 0 \\ \boldsymbol{0} & \text{otherwise} \end{cases}$

$\boldsymbol{x}_{\mathsf{mc-sdp}} = \sum_{i=1}^{n} \frac{\boldsymbol{X}^\star_{[\omega_i]}}{\sum_{j=1}^{n} \boldsymbol{X}^\star_{[\omega_j]}} \boldsymbol{v}_i$

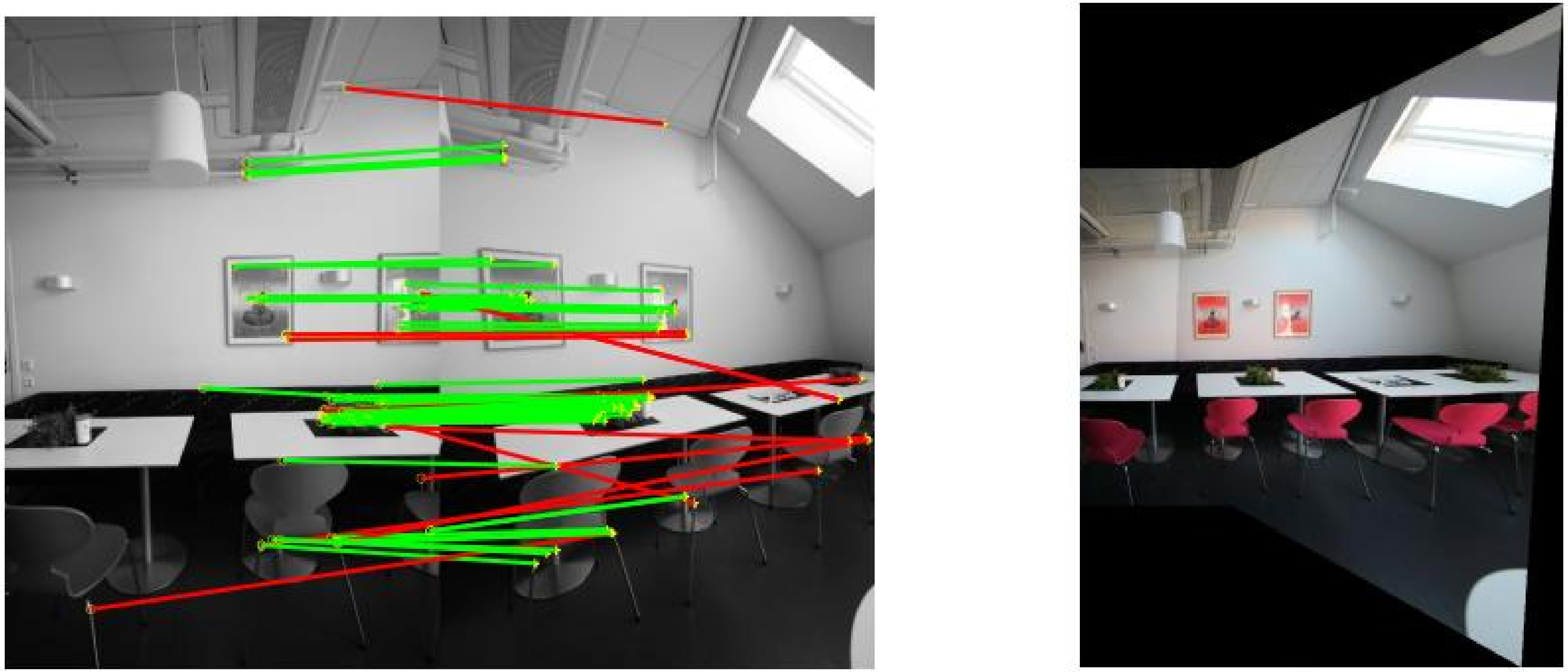
return $\boldsymbol{x}_{\mathsf{mc-sdp}}$.

04

Numerical Experiments

Image stitching图像拼接

Feature matches between two images using the moment relaxation approach, the matches include many outliers, shown in red (inliers are shown in green).



(a) (b)

4.1 Datasets and steps

- We use the **PASSTA dataset** to test rotation search problems arising in image stitching applications. In order to generate the vector pairs a_i, $b_i \in R^3$, i=1,2,....,n, we first use **SURF** to detect and match point features between the two images.

- From the **SURF feature points**, we apply the inverse of the known camera intrinsic matrix K to obtain unit-norm bearing vectors $\{\boldsymbol{a}_i, \boldsymbol{b}_i\}_{i=1}^{70}$ observed in each camera frame.

- Estimate the rotation **R** between the two camera frames, using the outlier-corrupted pairs $\{\boldsymbol{a}_i, \boldsymbol{b}_i\}_{i=1}^{70}$ .Finally, using the estimated **R**, we can compute the homography matrix as H = KRK^{-1} to stitch the pair of images together.

4.1 Datasets and steps

- PASSTA dataset: a public dataset for evaluating image alignment
- SURF: Speeded Up Robust Features(加速稳健特征). It is a robust image recognition and description algorithm, first published at the ECCV Conference in 2006. This algorithm can be used for computer vision tasks such as object recognition and 3D reconstruction.
- SURF feature points: 兴趣点检测及描述子算法。其通过Hessian矩阵的行列式来确定兴趣点位置，再根据兴趣点邻域点的Haar小波响应来确定描述子，其描述子大小只有64维，是一种非常优秀的兴趣点检测算法。

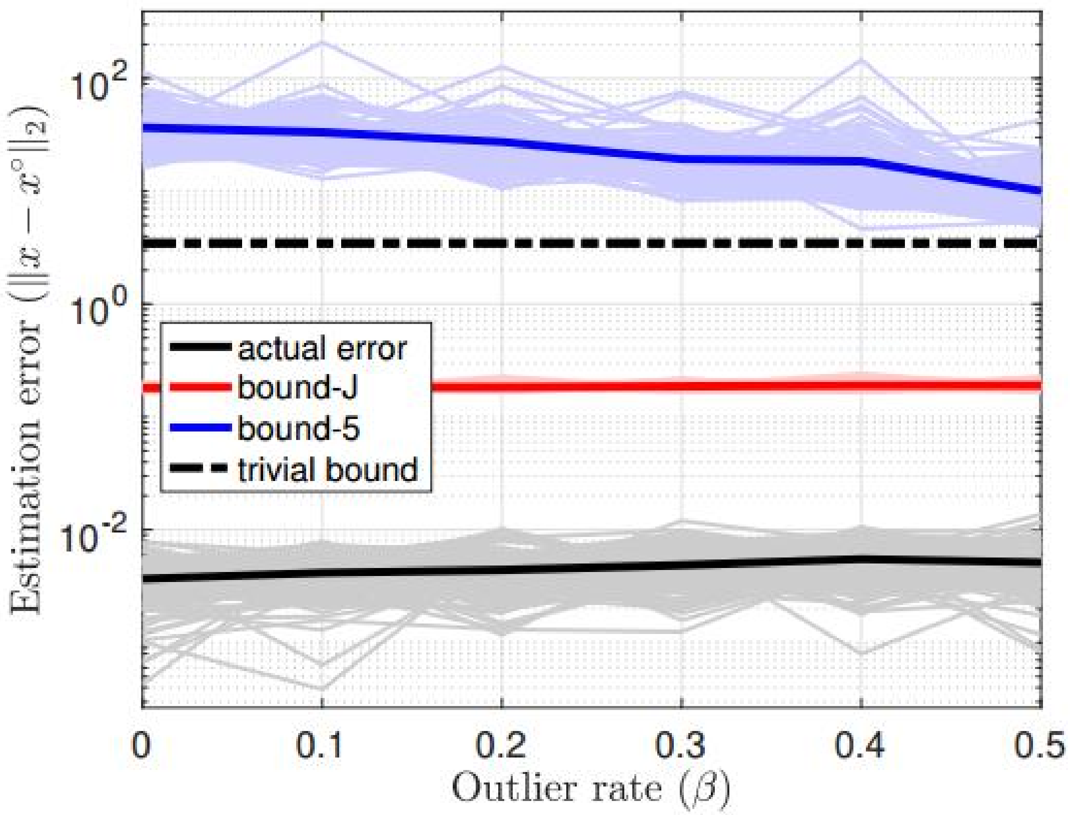
4.2 Algorithms and implementation details

- We use the sparse moment relaxation to solve the (TLS) formulation of the rotation search problem, and use an SDP solver for the resulting semidefinite relaxation.
- In all results below we observe the sparse moment relaxation of (TLS) to be tight. Hence we do not differentiate between the two algorithms.
- The numerical results are obtained on a MacBook Pro with 2.8 GHz Quad-Core Intel Core i7 processor. T

4.3 Outlier rates and estimation error

- The figure compares the actual estimation error with the bounds from Proposition 6.
- The estimation error is computed as $\|x_{\text{TLS}} - x^{\circ}\|$, where x and x° are the vectorized representations of the estimated and the ground-truth rotation, respectively.

- Actual error of the (TLS) estimator compared to the a posteriori bounds from Proposition 6 and the trivial bound 2Mx = $2\sqrt{3}$, with n = 50 and increasing outlier rates β. We report a posteriori bounds for dJ = 5 ("bound-5") and for the set J chosen as the correctly selected inliers ("bound-J").



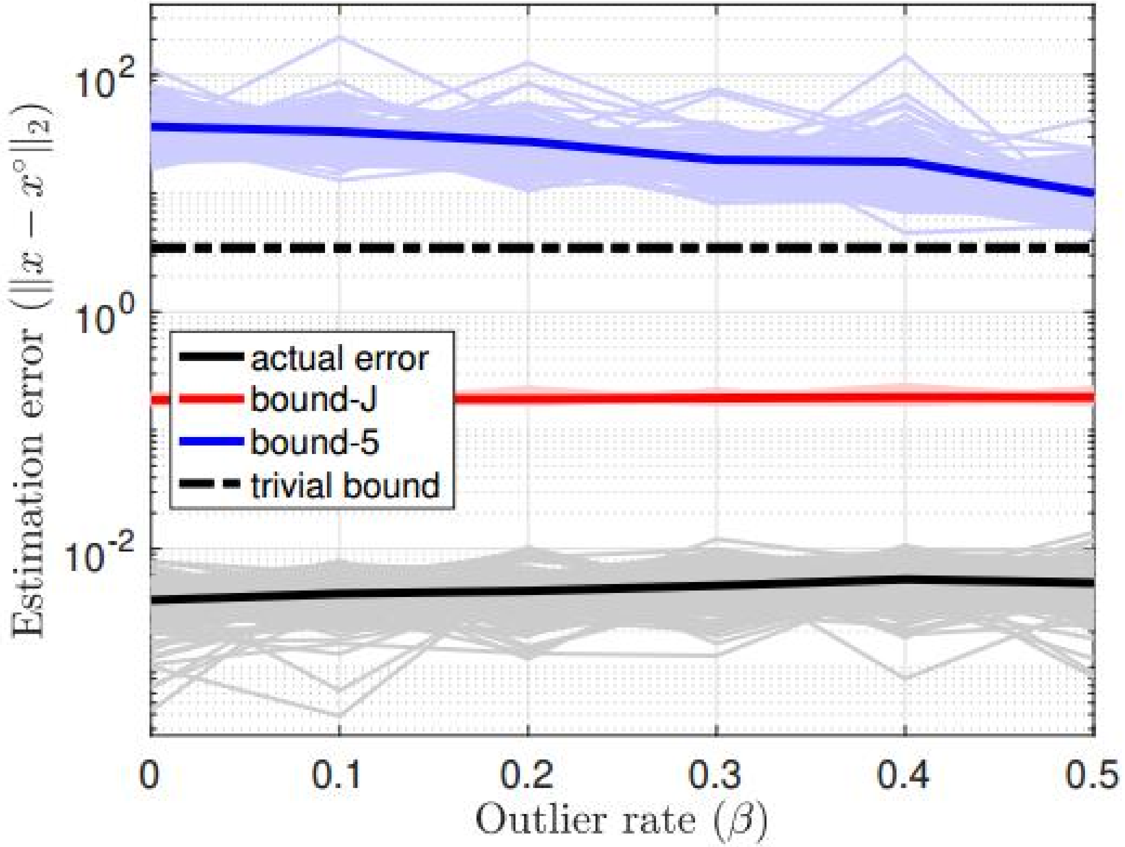
Proposition 6 (Low-outlier case: a posteriori estimation contract for (TLS)). *Consider Problem 1 with measurements* $(\boldsymbol{y}_i, \boldsymbol{A}_i)$, $i \in [n]$, *and denote with* γ° *the squared residual error of the ground truth* $\boldsymbol{x}^\circ$ *over the set of inliers* $\mathcal{I}$, i.e., $\gamma^\circ \triangleq \sum_{i \in \mathcal{I}} \|\boldsymbol{y}_i - \boldsymbol{A}_i^\top \boldsymbol{x}^\circ\|_2^2$. *Moreover, assume the measurement set contains at least* $\frac{n+\bar{d}}{2} + \frac{\gamma^\circ}{\bar{c}^2}$ *inliers, where* $\bar{d}$ *is the size of a minimal set, and that every subset of* $\bar{d}$ *inliers is nondegenerate. Then, for any integer* $d_{\mathcal{J}}$ *such that* $\bar{d} \leq d_{\mathcal{J}} \leq (2\alpha - 1)n - \frac{\gamma^\circ}{\bar{c}^2}$, *an optimal solution* $\boldsymbol{x}_{\mathsf{TLS}}$ *of* (TLS) *satisfies*

$$\|\boldsymbol{x}_{\mathsf{TLS}} - \boldsymbol{x}^\circ\|_2 \leq \frac{2\sqrt{d_{\mathcal{J}}}\;\bar{c}}{\min_{\mathcal{J} \subset \mathcal{I}_{\mathsf{TLS}}, |\mathcal{J}| = d_{\mathcal{J}}} \sigma_{\min(\boldsymbol{A}_{\mathcal{J}})}}, \tag{27}$$

where $\mathcal{I}_{\mathsf{TLS}}$ *is the set of inliers selected by* (TLS), $\boldsymbol{A}_{\mathcal{J}}$ *is the matrix obtained by horizontally stacking all submatrices* $\boldsymbol{A}_i$ *for all* $i \in \mathcal{J}$, *and* $\sigma_{\min}(\cdot)$ *denotes the smallest singular value of a matrix. Moreover, if the inliers are noiseless,* i.e., $\boldsymbol{\epsilon} = \boldsymbol{0}$ *in eq.* (17), *and for a sufficiently small* $\bar{c} > 0$, $\boldsymbol{x}_{\mathsf{TLS}} = \boldsymbol{x}^\circ$.

4.3 Outlier rates and estimation error

- Several comments are in order:
- First, the actual error increases with the outlier rate (the trend is slightly more difficult to see due to the log scale): this is expected since with increasing outlier rates the number of “useful” measurements decreases;
- However, the error remains very small (i.e., less than 1 degree) for all outlier rates.
- Second, bound-5 is unfortunately too loose and uninformative, since—while it improves for larger outlier $rates^{24}$ — it remains larger than the trivial bound.



Thank You!

Avenida da Universidade, Taipa, Macau, China
Tel : (853) 68910353 Fax : (853) 8822 8822
Email : mc35295@umac.mo Website : www.um.edu.mo