Outlier-Robust Spatial Perception: Hardness, Algorithms, Guarantees

Vasileios Tzoumas
Research Scientist
Laboratory of Decision and Information Systems
Massachusetts Institute of Technology

Outlier-robust spatial perception…

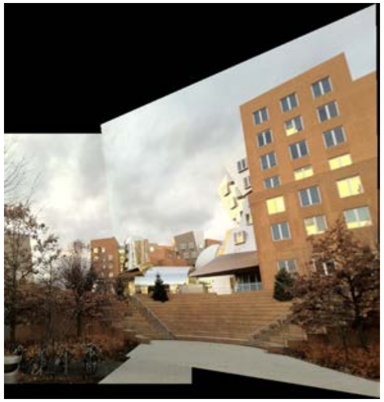
… aka protecting spatial perception from misinformative data, called outliers

In previous lectures, we saw need for RANSAC to protect from misinformative correspondences:

Some correspondences are wrong (outliers)

without RANSAC

with RANSAC



Outlier-robust spatial perception…

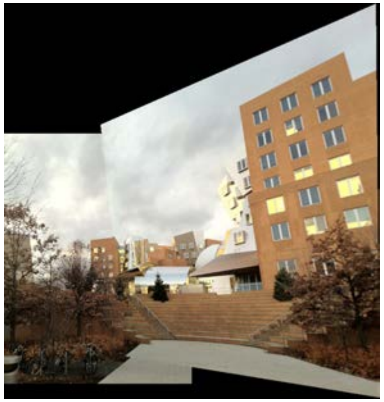
… aka protecting spatial perception against misinformative data, called outliers

In this and the following two lectures, we'll learn more about:

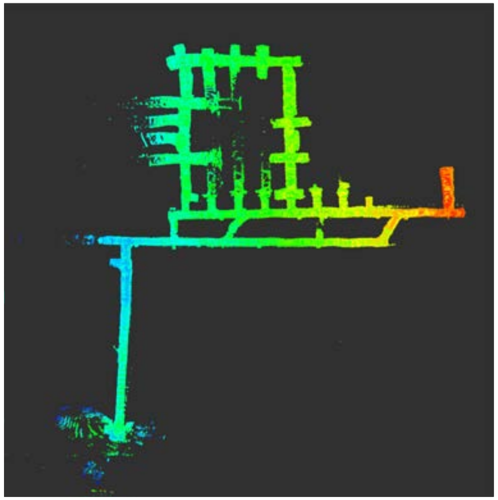
- **Rigorous formulations for outlier-robust perception**
- **Hardness**: How easy is it to detect outliers?
- **Algorithms**: How to remove outliers?
- **Guarantees**: How do the algorithms perform?

…from here:

so we go here…



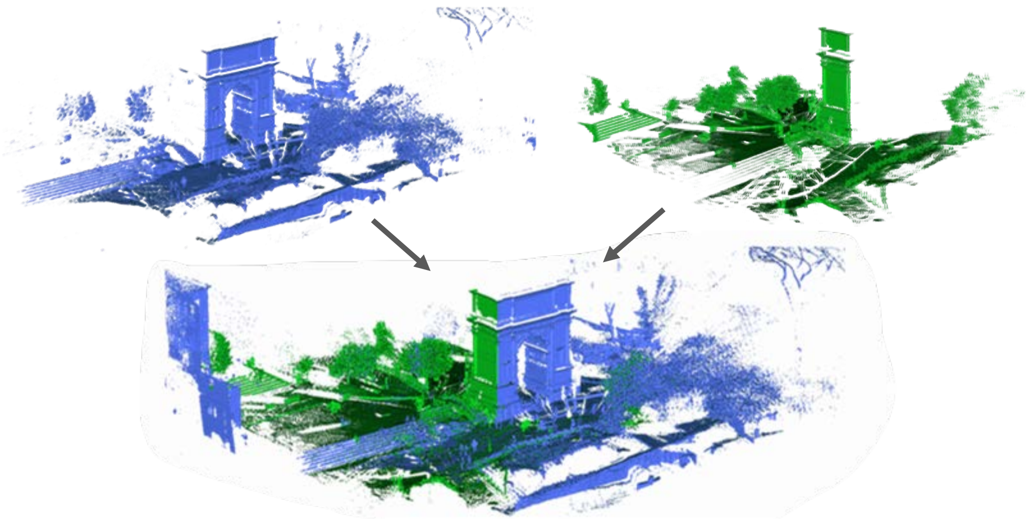
Successful image stitching / Failed image stitching



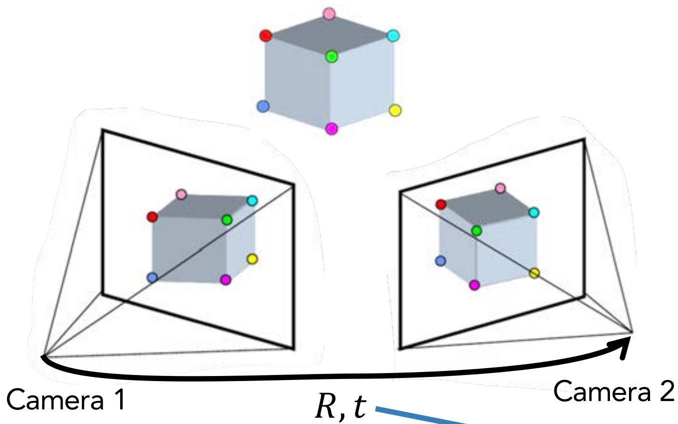
Successful SLAM / Failed SLAM

Successful shape alignment / Failed shape alignment

Perception in robotics and computer vision

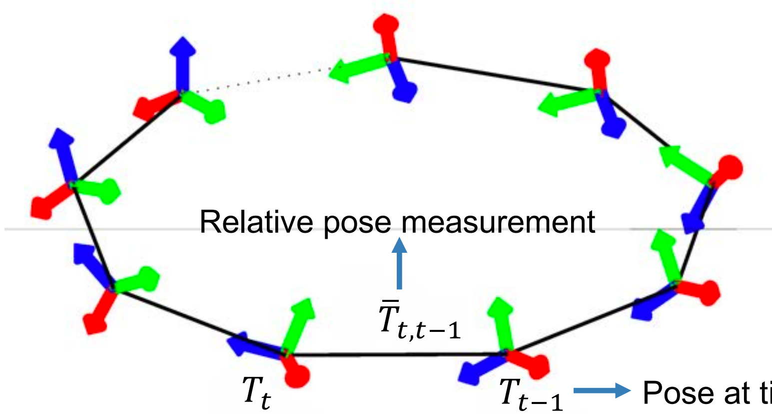


Point cloud registration



Two-view geometry

R: rotation
t: translation



Pose graph optimization

Perception as least squares optimization

When Gaussian measurement noise, maximum likelihood estimation (MLE) gives:

$$\min_{\boldsymbol{x} \in \mathcal{X}} \sum_{i=1}^{N} r^2(\boldsymbol{y}_i, \boldsymbol{x})$$

Estimate

Measurements/data

Residual

Examples:

Point cloud registration:

$$\min_{\substack{R \in SO(3) \\ t \in \mathbb{R}^3}} \sum_{(i,j) \in \mathcal{M}} \| R p_i + t - p_j' \|^2$$

Rotation + translation

Point clouds (data)

Correspondences between p_i, p'_j

SLAM:

$$\min_{\substack{T_i \in \mathrm{SE}(3) \\ i=1,\ldots,n}} \sum_{(i,j) \in \mathcal{M}} \| T_j - T_i \bar{T}_{ij} \|_F^2$$

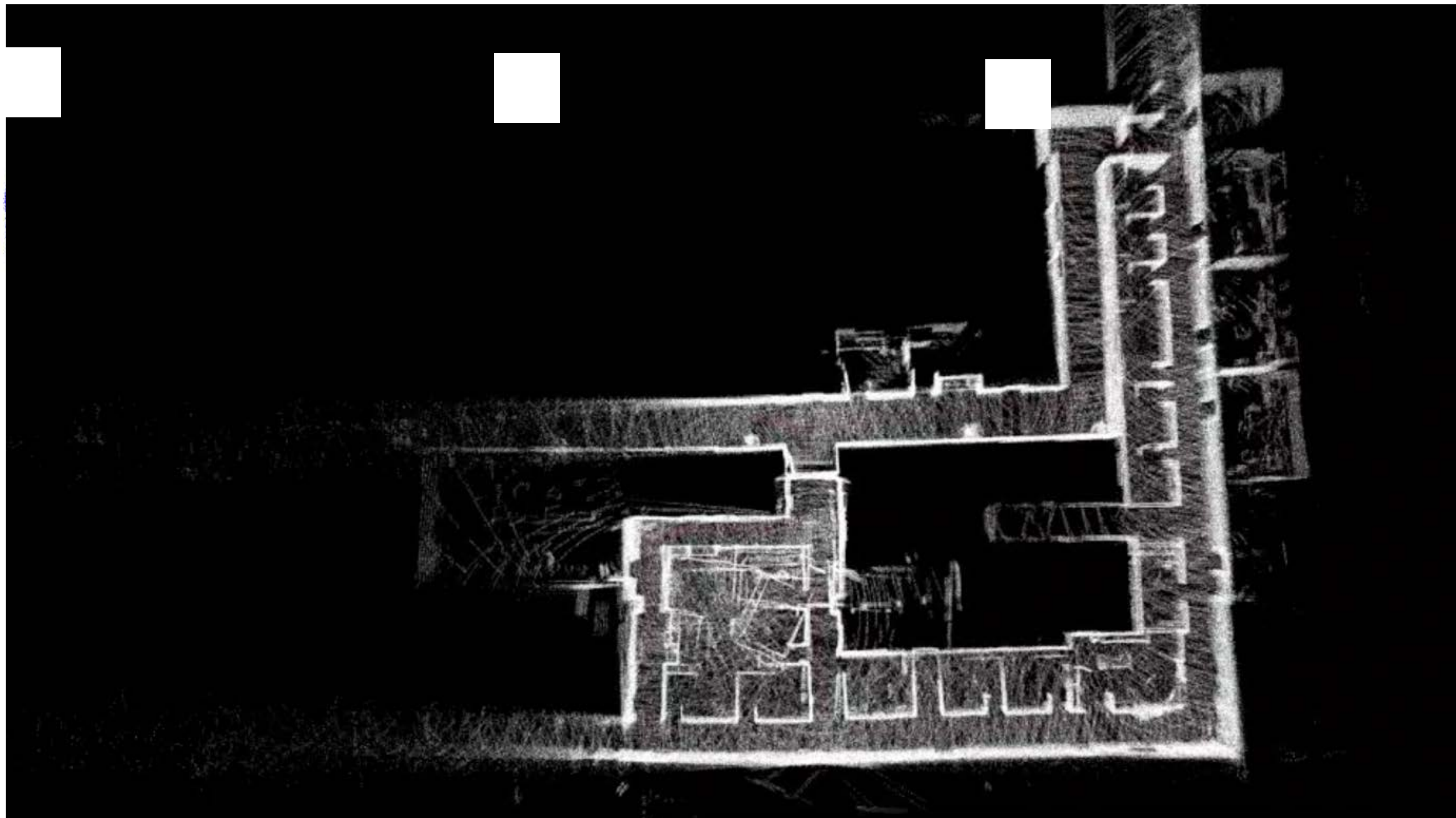
Pose

Relative pose measurement

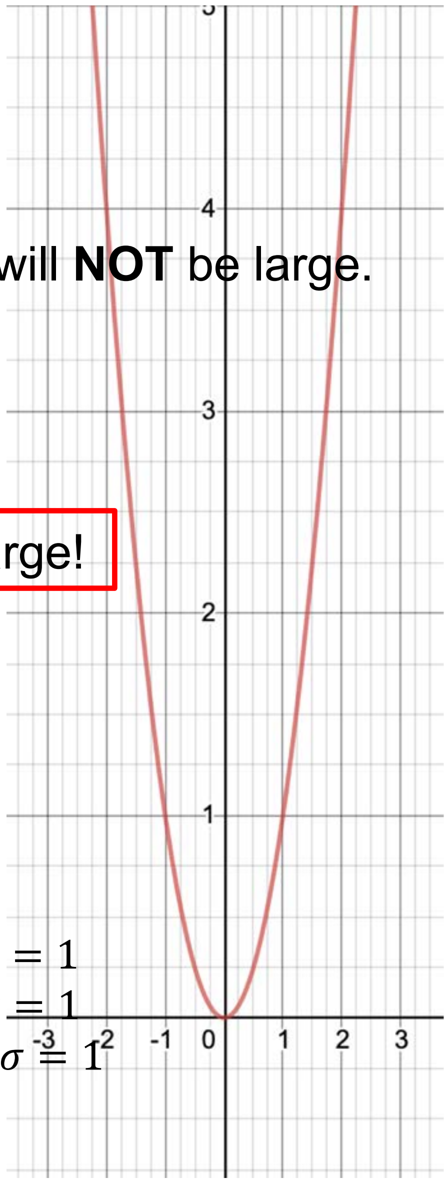
Loop closures between T_i, T_j

Outliers compromise least squares solutions

But if some y_i are outliers, solution of $\min_{\boldsymbol{x} \in \mathcal{X}} \sum_{i=1}^{N} r^2(\boldsymbol{y}_i, \boldsymbol{x})$ can be wrong:



Why least squares can fail?



Least squares penalizes large residuals a **LOT** (due to square).

➡ Least squares finds an estimate x where residuals will **NOT** be large.

But at x_{true} the residuals of the outliers will be large!

Example:

- $x_{true} = 0$
- Measurements' model: $y_1 = x + gaussian\ noise\ of\ \mu = 0, \sigma = 1$
  $y_2 = x + gaussian\ noise\ of\ \mu = 0, \sigma = 1$
  $y_3 = 2x + gaussian\ noise\ of\ \mu = 0, \sigma = 1$
- Observed measurements: $y_1 = y_2 = 0,\ y_3 = 10$

Least squares opt. solution is $x = 3.33 \neq x_{true} = 0$!

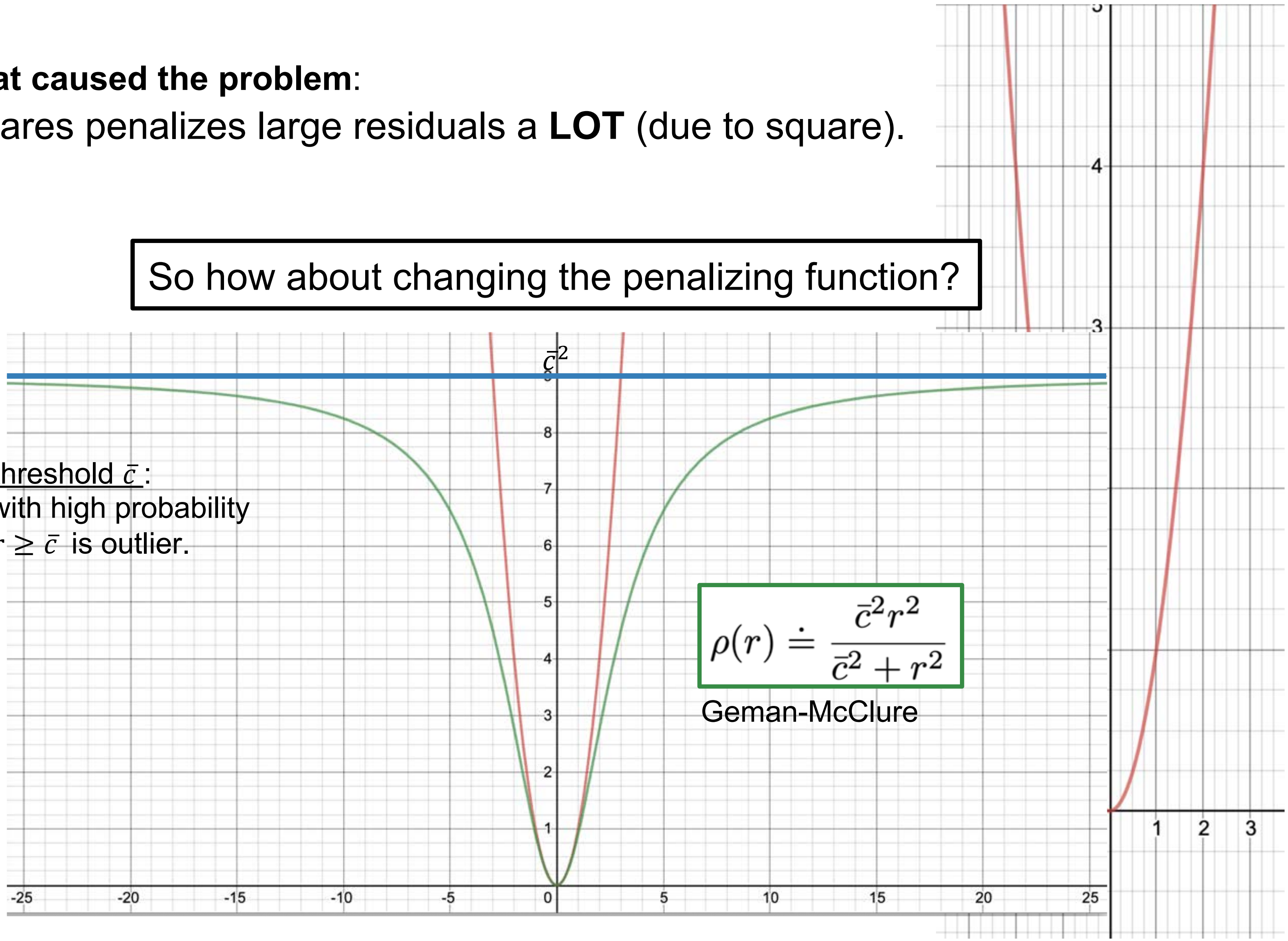
Outlier-robust least squares reformulations

This slide has been intentionally left blank.

Outlier-robust least squares reformulations

Recall what caused the problem:
Least squares penalizes large residuals a **LOT** (due to square).

So how about changing the penalizing function?



"Outlier-free" threshold $\bar{c}$:
Choose it so with high probability any residual $r \geq \bar{c}$ is outlier. (more later)

$$\rho(r) \doteq \frac{\bar{c}^2 r^2}{\bar{c}^2 + r^2}$$

Geman-McClure

Outlier-robust least squares reformulations

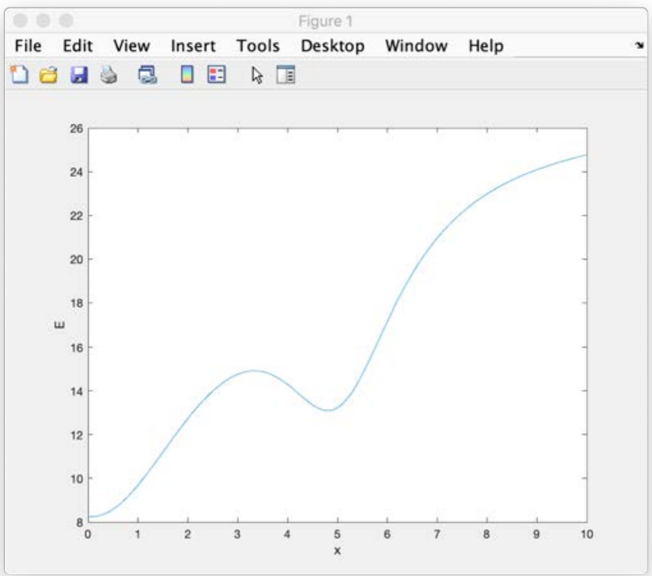
Example (revisited): Now instead of solving $\min_{\boldsymbol{x}\in\mathcal{X}} \sum_{i=1}^{N} r^2(\boldsymbol{y}_i, \boldsymbol{x})$ we solve $\min_{\boldsymbol{x}\in\mathcal{X}} \sum_{i=1}^{N} \rho(r(\boldsymbol{y}_i, \boldsymbol{x}))$.

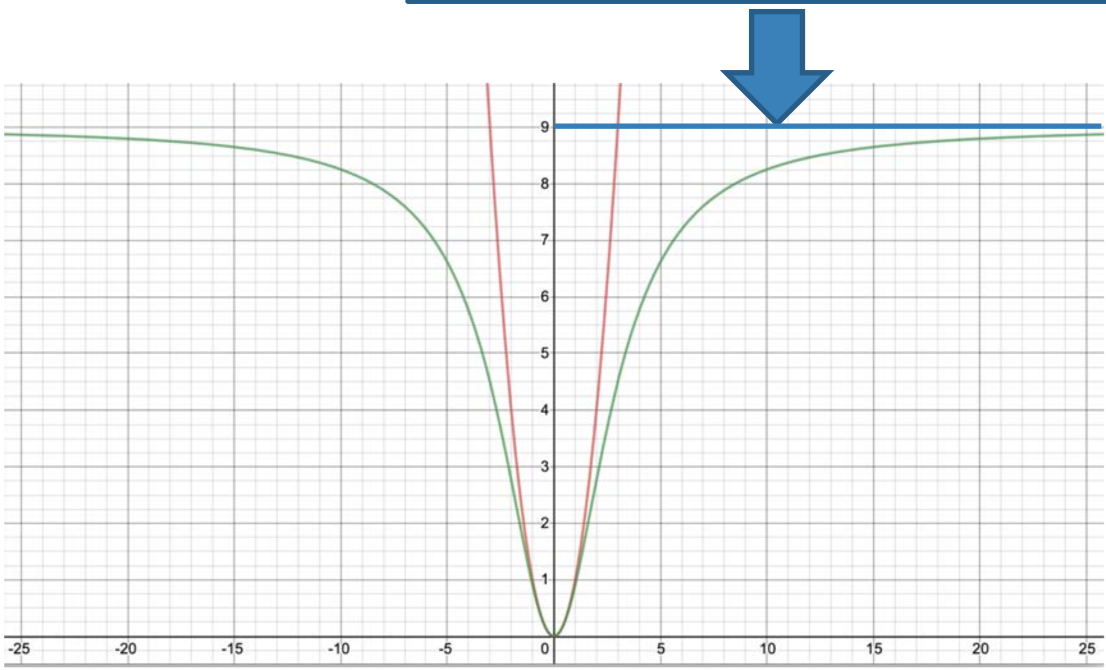
E

- $x_{true} = 0$
- Measurements' model: $y_1 = x + gaussian\ noise\ of\ \mu = 0, \sigma = 1$
  $y_2 = x + gaussian\ noise\ of\ \mu = 0, \sigma = 1$
  $y_3 = 2x + gaussian\ noise\ of\ \mu = 0, \sigma = 1$
- Observed measurements: $y_1 = y_2 = 0$, $y_3 = 10$

E with ρ as in figure has now opt. solution $x = x_{true} = 0$!

with probability .99 residuals ≤ 3 are inliers





Robust-cost "least squares"

Generally, instead of solving $\min_{\boldsymbol{x} \in \mathcal{X}} \sum_{i=1}^{N} r^2(\boldsymbol{y}_i, \boldsymbol{x})$ we solve, towards outlier-robustness,

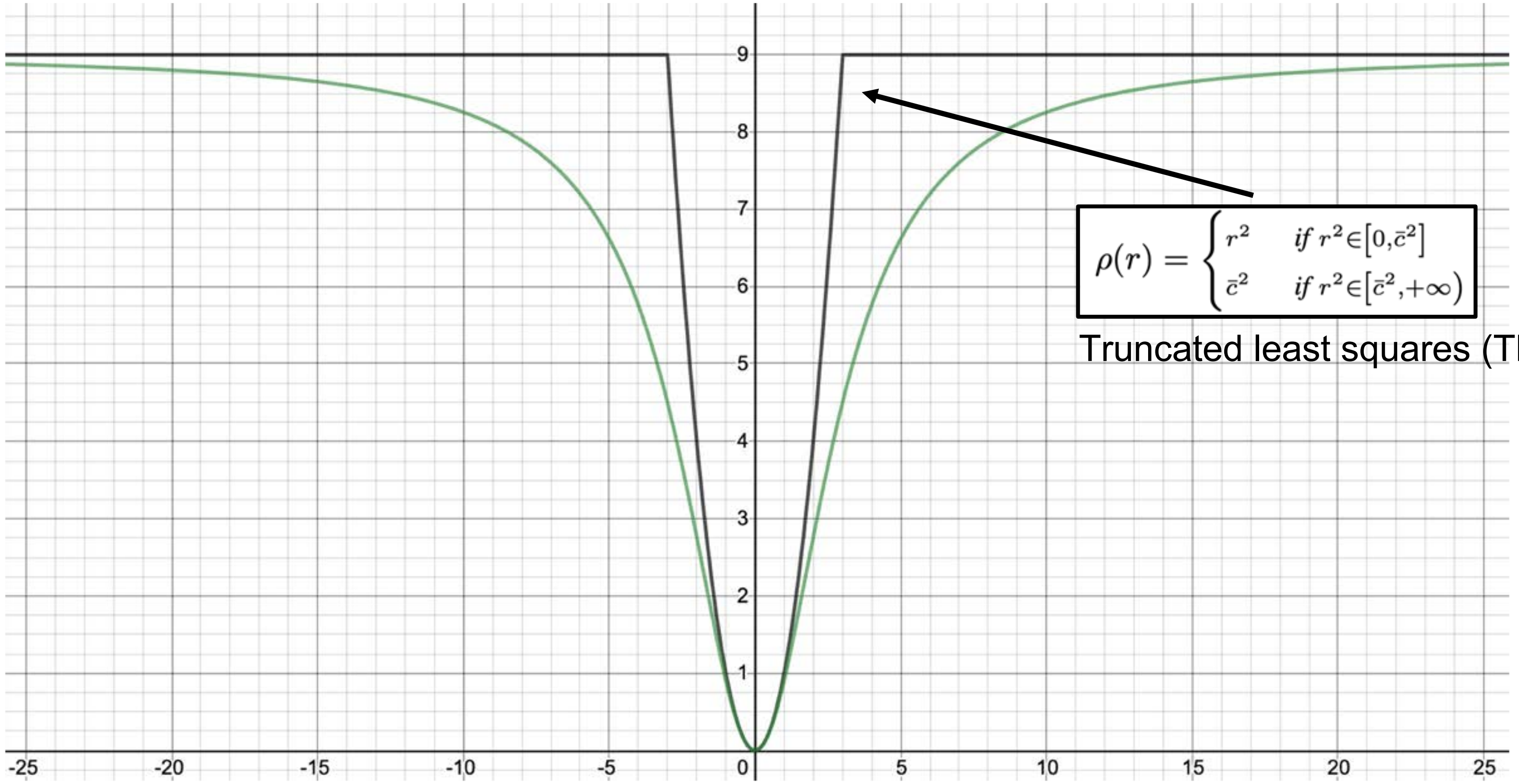
$$\boxed{\min_{\boldsymbol{x} \in \mathcal{X}} \sum_{i=1}^{N} \rho(r(\boldsymbol{y}_i, \boldsymbol{x}))}$$

for some robust cost function ρ, as the one before.

Robust-cost "least squares"

Can we do better from ρ before?

Possibly: if any residual $r \geq \bar{c}$ can be treated as outlier, we can immediately penalize it with a $\bar{c}^2$ value:



$$\rho(r) = \begin{cases} r^2 & \text{if } r^2 \in [0, \bar{c}^2] \\ \bar{c}^2 & \text{if } r^2 \in [\bar{c}^2, +\infty) \end{cases}$$

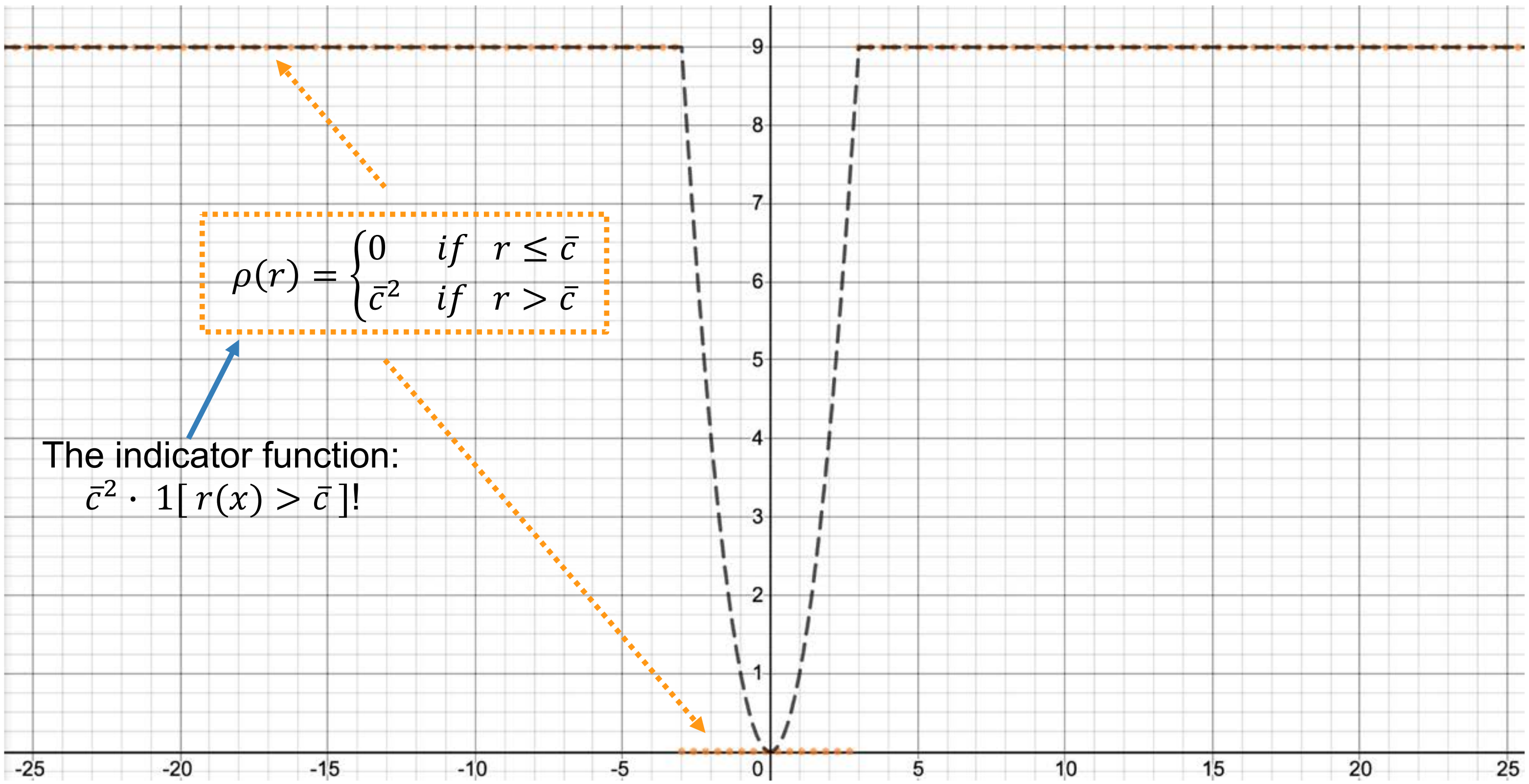
Truncated least squares (TLS)

Experimentally, TLS performs typically better, but no guarantees for superiority over green:
Issue algorithmic ability to find optimum solution (more later)

Robust-cost "least squares"

Other alternatives?

Sure: if any residual $\mathrm{r} \leq \bar{c}$ can be treated as inlier, why penalize it at all?



Robust-cost "least squares"

We ended up with a purely combinatorial problem, known as maximum consensus:

$$\min_{x \in \mathrm{X}} \sum_{i \in \mathcal{M}} 1[\, r(x, y_i) > \bar{c}\,] \;\equiv\; \max_{x \in \mathrm{X}} \sum_{i \in \mathcal{M}} 1[\, r(x, y_i) \leq \bar{c}\,]$$

Find x that maximizes number of explained measurements (as inliers)

Equivalently:

$$\min_{x \in \mathrm{X},\, \mathcal{O} \subseteq \mathcal{M}} |\mathcal{O}| \quad s.t. \quad r(x, y_i) \leq \bar{c}, \qquad \forall\, i \in \mathcal{M} \backslash \mathcal{O}$$

Minimize number of rejected measurements s.t. the rest are explained

Outlier rejection approach

Outlier rejection “least squares”

Generalizing, instead of using robust cost functions, we can still do least squares after rejecting outliers:

$$\min_{x \in \mathrm{X}, \mathcal{O} \subseteq \mathcal{M}} |\mathcal{O}| \quad s.t. \quad \left\| r\left(x, y_{\mathcal{M} \setminus \mathcal{O}}\right) \right\|^2 \leq \epsilon$$

Cumulative outlier-free threshold

Outlier-robust least squares reformulations

$$\min_{x \in \mathcal{X}} \sum_{i \in \mathcal{M}} \rho(r(x, y_i), \bar{c})$$

L

$$\min_{x \in \mathrm{X}, \mathcal{O} \subseteq \mathcal{M}} |\mathcal{O}| \quad s.t. \quad \left\| r\left(x, y_{\mathcal{M} \setminus \mathcal{O}}\right) \right\|^2 \leq \epsilon$$

R

- Two formulations can become equivalent (see maximum consensus case)
- In what follows we'll look into methods for solving L and R:
  - L inspires mainly non-linear/non-convex optimization approaches
  - R instead combinatorial approaches

But is it easy (computationally) to solve either of them?

Outlier-robust reformulations are harder than NP-hard

In the worst case, <u>even if true error is 0</u>, we will reject many more measurements (than the true outliers), and still incur larger than the true 0 error

Theorem (Chin et al. '18, Antonante et al. '19)

- Let $\mathcal{O}^\star$ be the true number of outliers, whose rejection leads to 0 residual error.
- Let $\epsilon = p_1(|\mathcal{M}|)$, where p_1 is a polynomial in number of measurements.
- Let $p_2(|\mathcal{M}|)$ another polynomial.

Then:

No quasi-polynomial algorithm can reject less than $p_2(|\mathcal{M}|)|\mathcal{O}^\star|$ measurements.

quasi-polynomial → **slower** than polynomial, **faster** than exponential

$p_2(|\mathcal{M}|)|\mathcal{O}^\star|$ → **no constant** approximation factor

Theorem applies to both L and R:

<u>Truncated least squares</u> (**L**): $\min_{x \in \mathcal{X}} \sum_{i \in \mathcal{M}} \rho\big(r(x, y_i)\big)$ with $\rho(r) = \begin{cases} r^2 & if \;\; r \in [0, \bar{c}] \\ \bar{c}^2 & if \;\; r \in [\bar{c}, +\infty] \end{cases}$

<u>Maximum consensus</u> (**R**): $\min_{x \in \mathrm{X}, \mathcal{O} \subseteq \mathcal{M}} |\mathcal{O}| \quad s.t. \quad r(x, y_i) \leq \bar{c}, \qquad \forall\, i \in \mathcal{M} \backslash \mathcal{O}$

But what if we could solve them optimally?

Even if $\bar{c}$ is picked correctly, true outlier may be impossible to detect **even if no noise**.

Reason: we have 2 knowns but 3 unknowns:

- $y_1 = x + o_1$
- $y_2 = x + o_2$

Known: $y_1 = 0, y_2 = 3$

All 3 unknown

Cannot uniquely determine x, o_1, o_2

Generally, for the noiseless case: we need to have redundant correct measurements, so if all outliers are rejected, remaining measurements can uniquely determine x (see [1])

(for the control oriented audience: observability)

[1] Stable Signal Recovery from Incomplete and Inaccurate Measurements, Candes and Tao, 2005

How to pick $\bar{c}$?

This slide has been intentionally left blank.