Pareto meets Huber: Efficiently avoiding poor minima in robust estimation

CHALMERS UNIVERSITY OF TECHNOLOGY — université de BORDEAUX

Christopher Zach (christopher.m.zach@gmail.com) — Chalmers University of Technology, Sweden
Guillaume Bourmaud (guillaume.bourmaud@u-bordeaux.fr) — University of Bordeaux, France

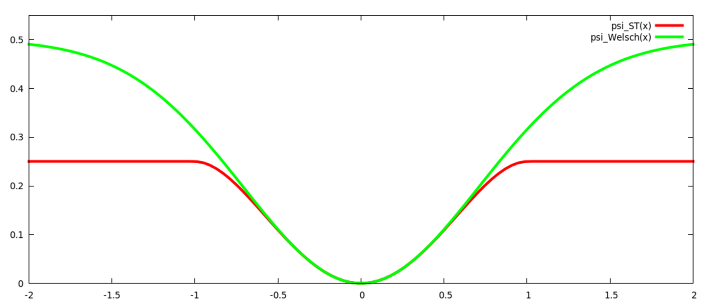
1) Introduction

Problem statement

Minimize a cost function involving robust data terms

$$\min_{\mathbf{x}} \Psi(\mathbf{x}) \quad \text{with} \quad \Psi(\mathbf{x}) = \sum_{i=1}^{N} \psi(\|\mathbf{r}_i(\mathbf{x})\|)$$

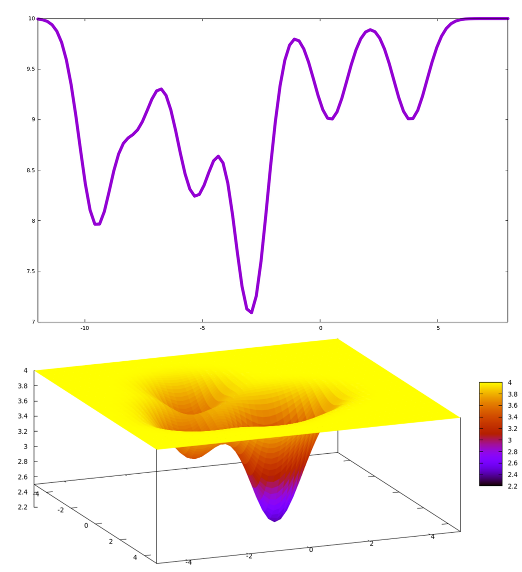
where $\mathbf{r}_i : \mathbb{R}^p \to \mathbb{R}^n$ is the vectorial residual function and $\psi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a *robust kernel* function.



Challenges

- large number of local minima
- large number of parameters to estimate

How to obtain an algorithm able to quickly decrease $\Psi(\mathbf{x})$ while avoiding poor local minima?



2) Contributions

1) we propose to use a Multi-Objective Optimization (MOO) approach to obtain an algorithm able of both avoiding poor local minima and quickly decreasing the target objective,

2) we derive an efficient Levenberg-Marquardt-MOO (LM-MOO) method yielding cooperative minimization steps.

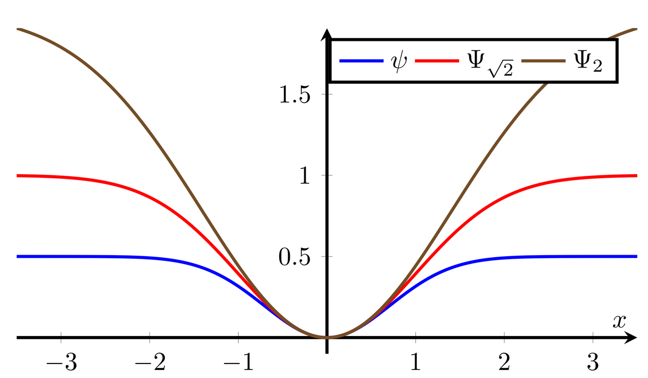
	IRLS [1], Triggs [2], $\sqrt{\psi}$ [3]	HQ [4], k-HQ [5]	GOM [6]	GOM+ [7]	LM-MOO (ours)
Quickly decreases target cost*	▮▮▮ (3/3)	▮▮ (2/3)	▯ (0/3)	▮ (1/3)	▮▮ (2/3)
Avoids poor local minima*	▯ (0/3)	▮ (1/3)	▮▮ (2/3)	▮▮ (2/3)	▮▮ (2/3)
Never ignores target cost	✓	×	×	×	✓
No extra variables	✓	×	✓	✓	✓

State of the art NLLS-based robust estimation algorithms and their corresponding properties.
(*) These rankings are observed experimentally on several computer vision problems.

3) Creating a sequence of "guidance" costs

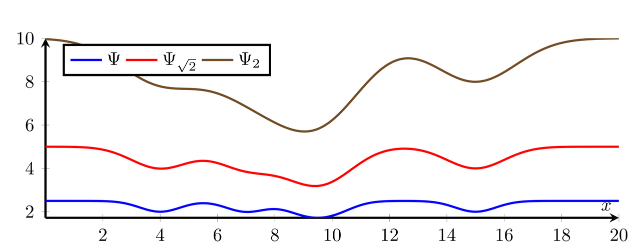
Scaled version of a robust kernel

$$\psi_\tau(x) = \tau^2 \psi(x/\tau)$$



Smoothed version of $\Psi(\mathbf{x})$

$$\Psi_\tau(\mathbf{x}) = \sum_{i=1}^{N} \psi_\tau(\|\mathbf{r}_i(\mathbf{x})\|)$$



Sequence of "guidance" costs provided as input of LM-MOO

$$(\Psi^1, \dots, \Psi^{K_{\max}}) \quad \text{where} \quad \Psi^i(\mathbf{x}) = \Psi_{2^{(i-1)}}(\mathbf{x})$$

4) Multi-objective Levenberg-Marquardt method (LM-MOO)

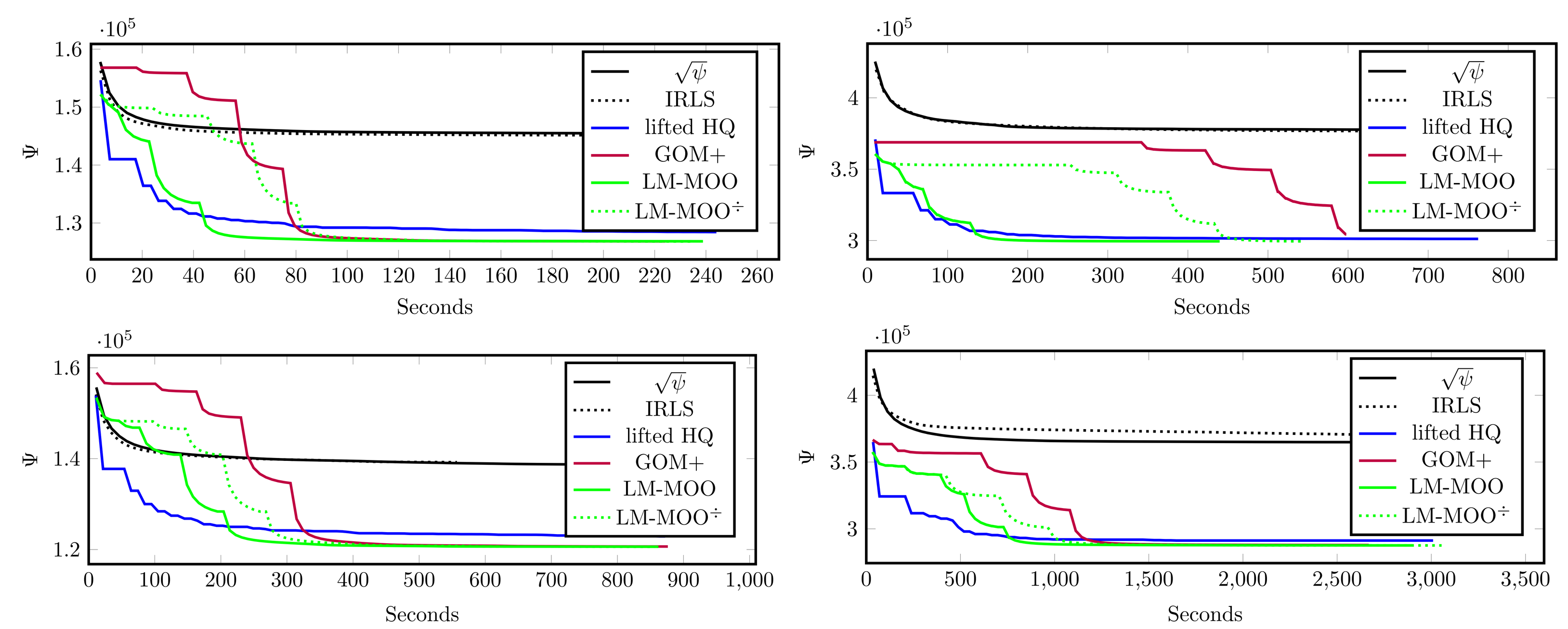
Require: Target Ψ and guidance costs $(\Psi^1, \dots, \Psi^{K_{\max}})$
Require: Initial solution $\mathbf{x}_0$, parameter, $\nu > 0$

```
k ← K_max
repeat
    μ ← ‖∇Ψ(x_0)‖ / (‖∇Ψ(x_0)‖ + ‖∇Ψ^k(x_0)‖)
    F^k ← (1 − μ)Ψ + μΨ^k
                                                ▷ Gauss-Newton / IRLS model
    g_F ← ∇F^k(x_0)      H_F ← ∇²F^k(x_0)
    v ← −(H_F + νI)^{-1} g_F                     ▷ Search direction
    x^+ ← x_0 + v
    if F^k(x^+) < F^k(x_0) then                  ▷ Success to reduce F^k
        strong ← Ψ(x^+) < Ψ(x_0) ∧ Ψ^k(x^+) < Ψ^k(x_0)
        stop ← TEST-STOPPING(Ψ, Ψ^k, x_0, x^+)
        if strong and not stop then
            x_0 ← x^+                            ▷ Update x_0
        else                                     ▷ Failure to reduce Ψ and Ψ^k
            k ← k − 1                            ▷ Go to next guidance function
        end if
        ν ← ν/10                                 ▷ Decrease the damping parameter
    else                                         ▷ Failure to reduce F^k
        ν ← 10ν                                  ▷ Increase the damping parameter
    end if
until k = 0
return the solution of a standard Levenberg-Marquardt method given current point x_0
```

5) Results

Bundle adjustment

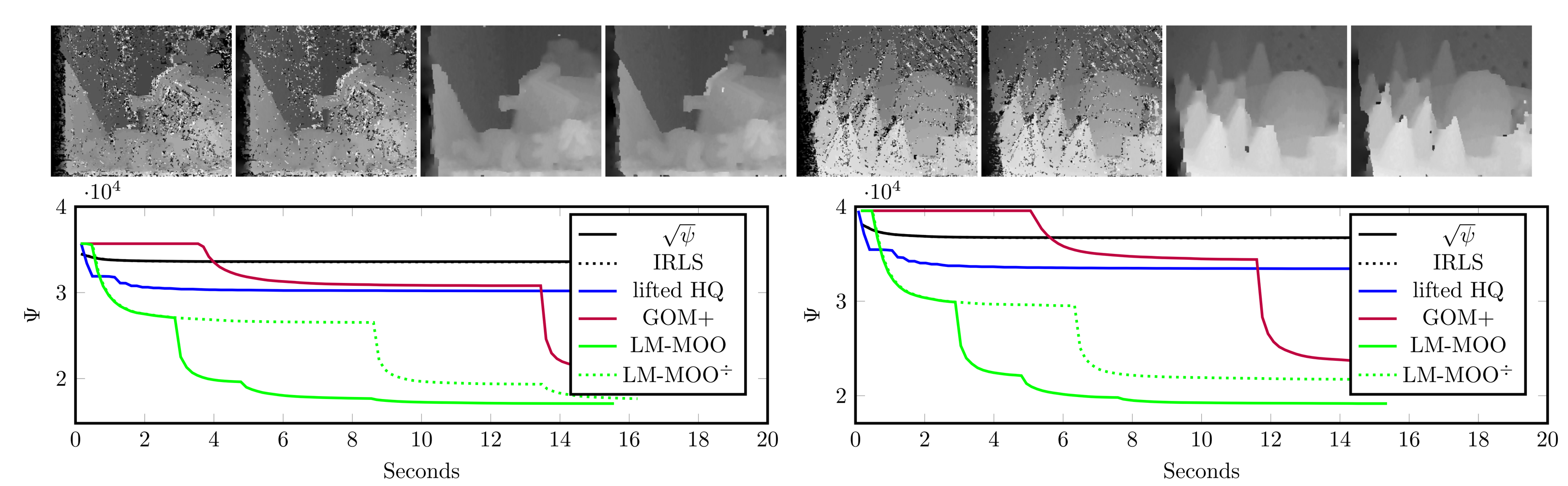
$$\min_{\{\mathtt{R}_i, \mathbf{t}_i\}_i, \{\mathbf{X}_j\}_j} \sum_{i,j} \psi\left(f_i \eta_i \left(\pi(\mathtt{R}_i \mathbf{X}_j + \mathbf{t}_i)\right) - \hat{\mathbf{p}}_{ij}\right)$$



Best encountered objective values obtained versus wall clock time as reported by different methods for linearized (top) and metric (bottom) bundle adjustment instances.

Dense correspondence

$$\min_{\mathbf{d}} \sum_{p \in \mathcal{V}} \left(\lambda \sum_{k=1}^{K} \psi_{\text{data}}(d_p - \hat{d}_{p,k}) + \sum_{q \in \mathcal{N}(p)} \psi_{\text{reg}}(d_p - d_q) \right)$$



Top: Initial best-cost depth and solutions of joint HQ, GOM+ and LM-MOO, respectively, for the "teddy" and "cones" stereo pair. Bottom: best objectives reached vs. runtime for different methods.

References

[1] Peter J Green. Iteratively reweighted least squares for maximum likelihood estimation, and some robust and resistant alternatives. 1984
[2] Bill Triggs, Philip McLauchlan, Richard Hartley, and Andrew Fitzgibbon. Bundle adjustment - A modern synthesis. 2000
[3] Chris Engels, Henrik Stewénius, and David Nistér. Bundle adjustment rules. 2006
[4] Christopher Zach. Robust bundle adjustment revisited. 2014
[5] Christopher Zach and Guillaume Bourmaud. Iterated lifting for robust cost optimization. 2017
[6] Andrew Blake and Andrew Zisserman. Visual reconstruction. 1987
[7] Christopher Zach and Guillaume Bourmaud. Descending, lifting or smoothing: Secrets of robust cost optimization. 2018