Multiplicative vs. Additive Half-Quadratic Minimization for Robust Cost Optimization

Christopher Zach (christopher.m.zach@gmail.com) — Toshiba Research Europe, Cambrige, UK

Guillaume Bourmaud (guillaume.bourmaud@u-bordeaux.fr) — University of Bordeaux, France

1) Introduction

Problem statement

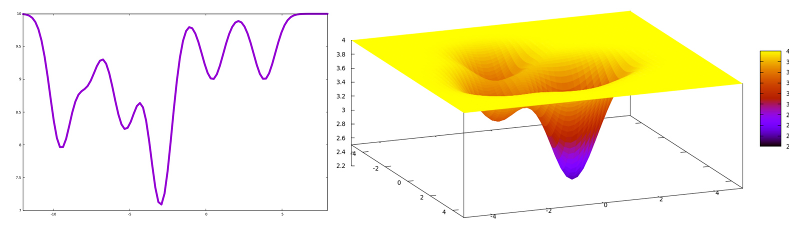
- Minimize a cost function involving robust data terms

$$\min_{\mathbf{x}} \Psi(\mathbf{x}) \quad \text{with} \quad \Psi(\mathbf{x}) = \sum_i \psi(\|\mathbf{f}_i(\mathbf{x})\|)$$

where $\mathbf{f}_i(\mathbf{x}) : \mathbb{R}^p \to \mathbb{R}^d$ and $\psi(\cdot)$ is a robust kernel function.

Challenges

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



How to obtain an efficient algorithm able to escape poor local minima?

2) Multiplicative Lifting

Introduce a Multiplicative Lifting Variable (M-LV) $v_i \in [0,1]$ to lift $\psi(\|\mathbf{f}_i(\mathbf{x})\|)$:

$$\psi(\|\mathbf{f}_i(\mathbf{x})\|) = \min_{v_i \in [0,1]} \tfrac{1}{2} v_i \|\mathbf{f}_i(\mathbf{x})\|^2 + \gamma(v_i)$$

and rewrite $\min_{\mathbf{x}} \Psi(\mathbf{x})$ to obtain an NLLS problem:

$$\min_{\mathbf{x}} \Psi(\mathbf{x}) = \min_{\mathbf{x},\{v_i\}_i} \Psi^{\text{M-HQ}}(\mathbf{x}, \{v_i\}_i)$$

where

$$\Psi^{\text{M-HQ}}(\mathbf{x}, \{v_i\}_i) = \frac{1}{2} \sum_i \left\| \begin{matrix} \sqrt{v_i}\, \mathbf{f}_i(\mathbf{x}) \\ \sqrt{2\gamma(v_i)} \end{matrix} \right\|^2$$

Reparametrize the M-LV $v_i = w(u_i)$ to avoid the constraint $v \geq 0$, e.g. $w(u) = u^2$ ($v_i \leq 1$ can usually be ignored)

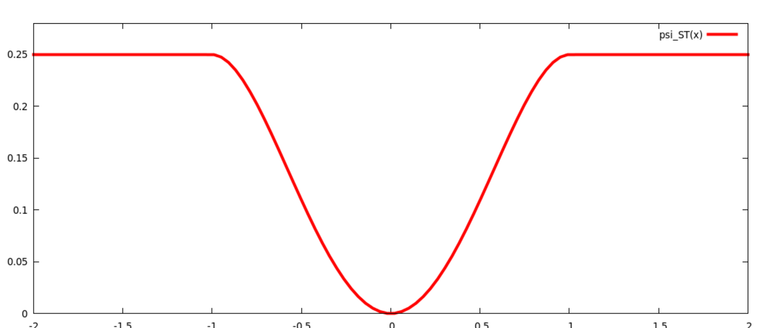
$$\Psi^{\text{M-HQ}}(\mathbf{x}, \{v_i\}_i) = \frac{1}{2} \sum_i \left\| \begin{matrix} \sqrt{w(u_i)}\, \mathbf{f}_i(\mathbf{x}) \\ \sqrt{2\gamma(w(u_i))} \end{matrix} \right\|^2$$

Levenberg-Marquardt solver for underlying NLLS

- $d+1$-dimensional residuals
- Schur complement trick applies

Example Choose $w(u) = u^2$ and $\gamma(v) = \frac{\tau^2}{4}(v-1)^2$ to obtain a smooth truncated kernel [1]

$$\psi_{ST}(\|\mathbf{f}\|) = \frac{\tau^2}{4}\left(1 - \left[1 - \frac{\|\mathbf{f}\|^2}{\tau^2}\right]_+^2\right)$$



3) Additive Lifting

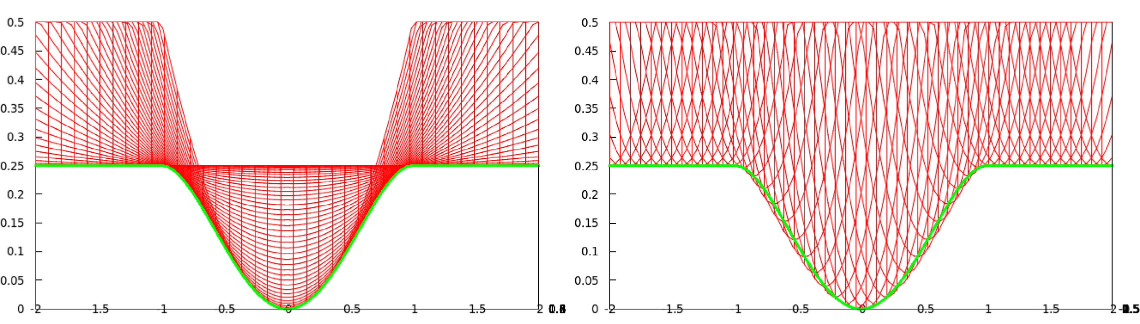
Introduce an Additive Lifting Variable (A-LV) $\mathbf{p}_i \in \mathbb{R}^d$ to lift $\psi(\|\mathbf{f}_i(\mathbf{x})\|)$:

$$\psi(\|\mathbf{f}_i(\mathbf{x})\|) = \min_{\mathbf{p}_i} \tfrac{\alpha}{2} \|\mathbf{f}_i(\mathbf{x}) - \mathbf{p}_i\|^2 + \rho(\|\mathbf{p}_i\|)$$

Problem: no closed-form expr. of $\rho \to$ relaxation

$$\tilde{\psi}^{\text{A-HQ}}(\mathbf{x}, \mathbf{p}_i) = \tfrac{\alpha}{2} \|\mathbf{f}_i(\mathbf{x}) - \mathbf{p}_i\|^2 + \psi(\|\mathbf{p}_i\|)$$

$\frac{\alpha}{2}\|\mathbf{f}_i(\mathbf{x}) - \mathbf{p}_i\|^2$ can be seen as quadratic penalizer



Insert into Ψ

$$\Psi^{\text{A-HQ}}(\mathbf{x}, \{\mathbf{p}_i\}_i) = \sum_i \left(\tfrac{\alpha}{2} \|\mathbf{f}_i(\mathbf{x}) - \mathbf{p}_i\|^2 + \psi(\|\mathbf{p}_i\|) \right)$$

Majorize-minimize (IRLS) applied on last term

$$\psi(\|\bar{\mathbf{p}}_i + \Delta\mathbf{p}_i\|) \leq \frac{\omega(\|\bar{\mathbf{p}}_i\|)}{2} \left(\|\bar{\mathbf{p}}_i + \Delta\mathbf{p}_i\|^2 - \|\bar{\mathbf{p}}_i\|^2\right) + \psi(\|\bar{\mathbf{p}}_i\|)$$

Levenberg-Marquardt-like solver

- Majorizer plus Gauss-Newton approximation
- $2d$-dimensional residuals
- Schur complement trick applies
- Approaches IRLS for $\alpha \to \infty$
  A-HQ is "relaxation" of IRLS

4) Double Lifting

Idea: apply A-HQ *and* M-HQ to avoid majorization step
Combine two steps:

- Additive lifting on $\psi(\|\mathbf{f}_i(\mathbf{x})\|)$

$$\psi(\|\mathbf{f}_i(\mathbf{x})\|) \rightsquigarrow \tfrac{\alpha}{2} \|\mathbf{f}_i(\mathbf{x}) - \mathbf{p}_i\|^2 + \psi(\|\mathbf{p}_i\|)$$

- Multiplicative lifting on $\psi(\|\mathbf{p}_i\|)$

$$\psi(\|\mathbf{p}_i\|) \rightsquigarrow \tfrac{1}{2} \left\| \sqrt{w(u_i)}\, \mathbf{p}_i \right\|^2 + \sqrt{\gamma(w(u_i))}^2$$

Combined:

$$\psi(\|\mathbf{f}_i(\mathbf{x})\|) = \min_{\mathbf{p}_i, u_i} \frac{\alpha}{2} \|\mathbf{f}_i(\mathbf{x}) - \mathbf{p}_i\|^2 + \frac{1}{2} \left\| \sqrt{w(u_i)}\, \mathbf{p}_i \right\|^2 + \sqrt{\gamma(w(u_i))}^2$$

$$= \min_{\mathbf{p}_i, u_i} \frac{1}{2} \left\| \begin{matrix} \sqrt{\alpha}(\mathbf{f}_i(\mathbf{x}) - \mathbf{p}_i) \\ \sqrt{w(u_i)}\, \mathbf{p}_i \\ \sqrt{2\gamma(w(u_i))} \end{matrix} \right\|^2$$

Minimize jointly over $\mathbf{x}$, $\{\mathbf{p}_i\}_i$ and $\{u_i\}_i$

$$\Psi^{\text{DL}}(\mathbf{x}, \{\mathbf{p}_i, u_i\}_i) = \frac{1}{2} \sum_i \left\| \begin{matrix} \sqrt{\alpha}(\mathbf{f}_i(\mathbf{x}) - \mathbf{p}_i) \\ \sqrt{w(u_i)}\, \mathbf{p}_i \\ \sqrt{2\gamma(w(u_i))} \end{matrix} \right\|^2$$

Levenberg-Marquardt solver

- $2d+1$-dimensional residuals
- Schur complement trick applies
- Closed-form inverse of $(d+1) \times (d+1)$ matrix

Run-time comparison

Relative iteration times

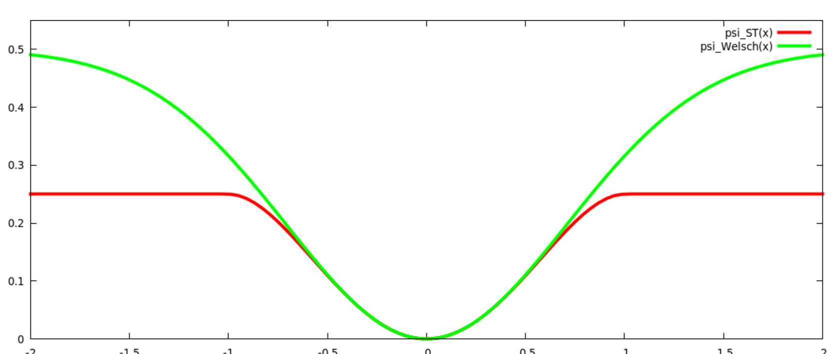
IRLS	M-HQ	A-HQ	DL
1	1.6	1.45	1.72

5) Results

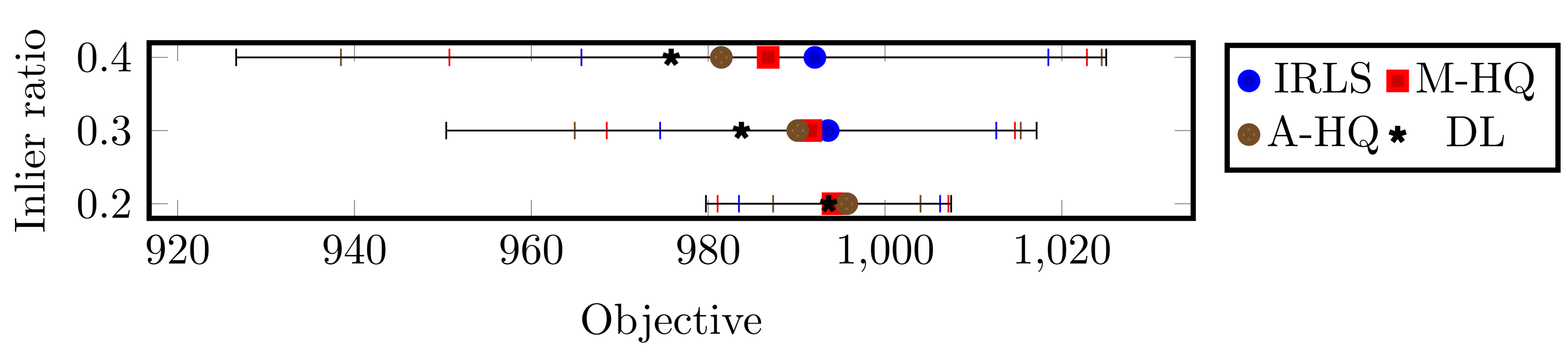
Tested kernels: Welsch and smooth truncated kernel

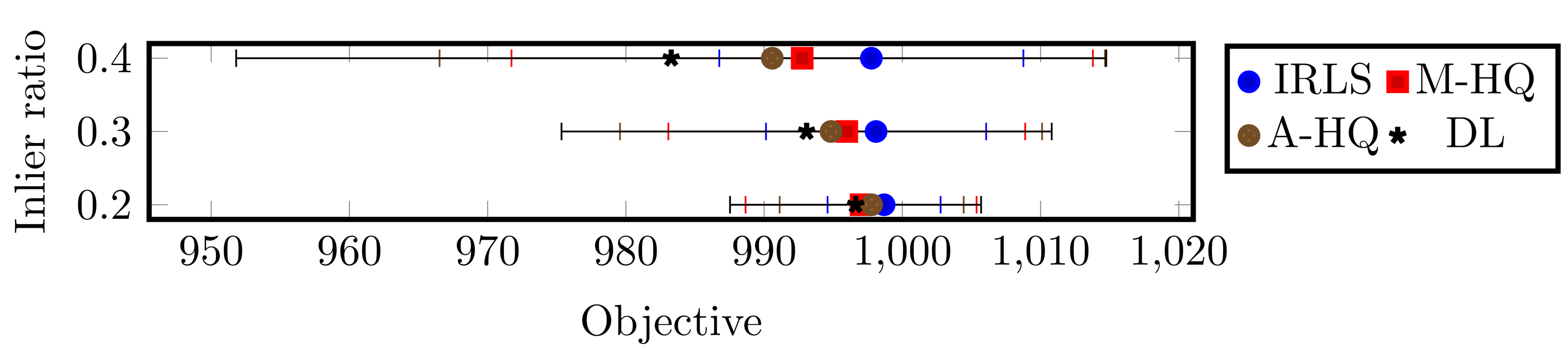
$$\psi_{ST}(z) = \tfrac{\tau^2}{4}\left(1 - \left[1 - \tfrac{z^2}{\tau^2}\right]_+^2\right)$$

$$\psi_{Welsch}(z) = \tfrac{\tau^2}{2}\left(1 - e^{-z^2/\tau^2}\right)$$



Robust mean: $\min_{\boldsymbol{\theta}} \sum_i \psi(\|\boldsymbol{\theta} - \mathbf{y}_i\|)$

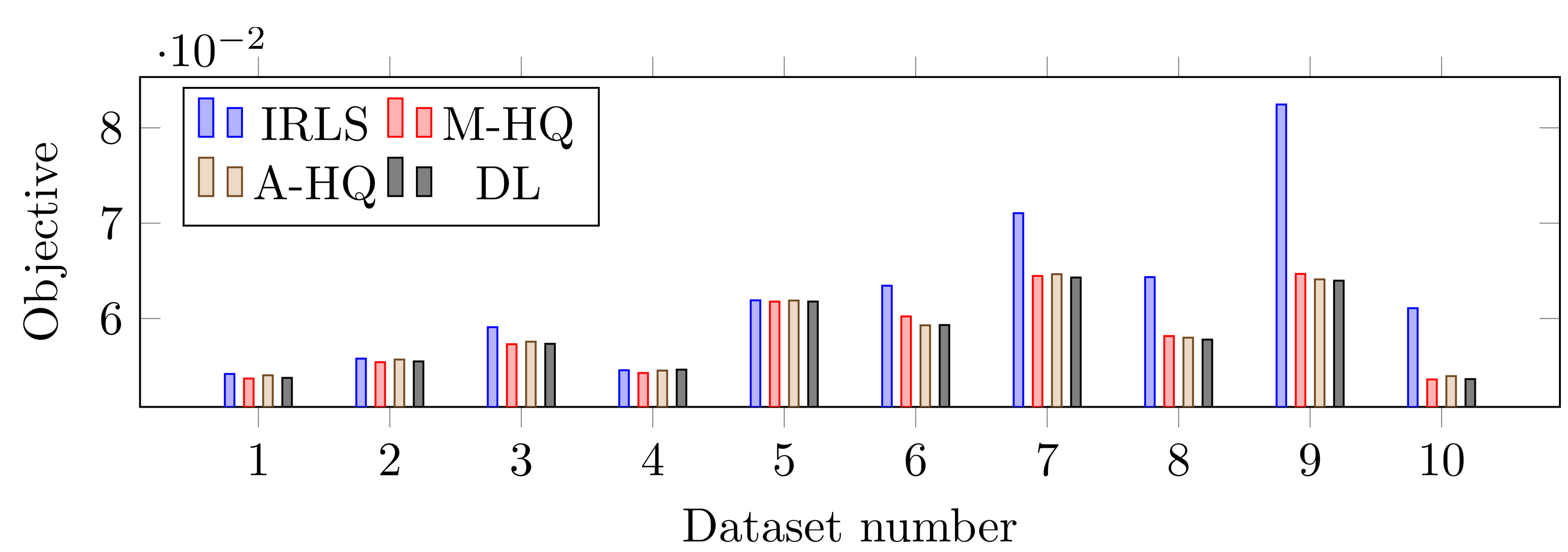


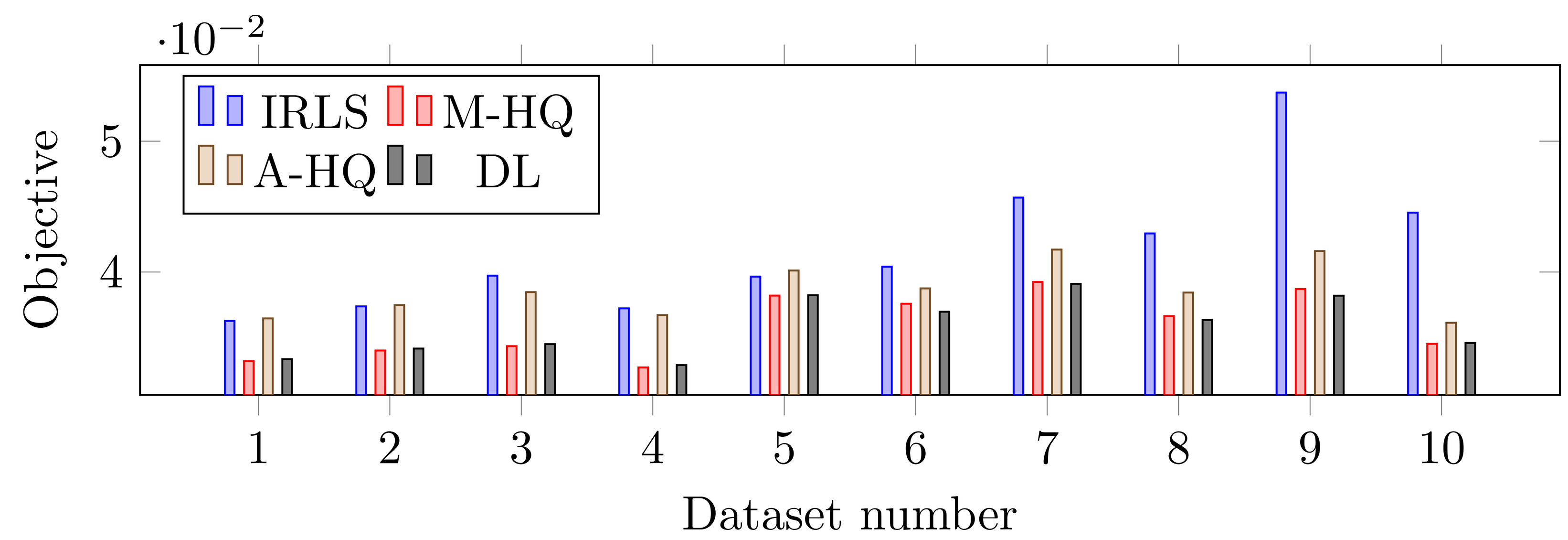


Avg. objective values (and std. deviations) reached by different methods
Top: 2D, bottom: 3D

Bundle adjustment

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





Objective values reached by the different methods for linearized bundle adjustment
Top: Welsch kernel, bottom: ST kernel

[1] Christopher Zach. Robust bundle adjustment revisited. ECCV 2014.