# Iterated Lifting for Robust Cost Optimization
# Supplementary material

Christopher Zach* and Guillaume Bourmaud†

## Contents



# 1 Proof of Prop. 3

## 1.1 Lifting of the Welsch kernel w.r.t. a scaled Welsch kernel

In the following we use Eqs. 10 and 11 from the main text and find the inverse function of $\alpha(t)$. Recall that the Welsch kernel is given by

$$\phi_{\text{Wel},\tau}(x) = \frac{\tau^2}{2}\left(1 - e^{-x^2/\tau^2}\right) \qquad \text{and} \qquad \omega_{\text{Wel},\tau}(x) = e^{-x^2/\tau^2}. \tag{1}$$

If it is lifted against a scaled version of itself, we obtain

$$\alpha(t) = \frac{\omega_{\text{Wel},\tau}(t/\lambda)}{\omega_{\text{Wel},\tau}(t/\mu)} = \exp\left(-\frac{(1/\lambda^2 - 1/\mu^2)\, t^2}{\tau^2}\right) = \exp\left(-\frac{\mu^2-\lambda^2}{\lambda^2\mu^2}\cdot\frac{t^2}{\tau^2}\right). \tag{2}$$

We can invert the mapping $\alpha(t)$, which yields

$$\alpha^{-1}(w) = \lambda\tau\mu\sqrt{\frac{-\log w}{\mu^2-\lambda^2}}. \tag{3}$$

Note that $\log w < 0$ and $\mu^2 - \lambda^2 > 0$, since $w \in (0, 1]$ and $\mu > \lambda$ by assumption. By observing that for $\nu \neq 0$

$$e^{-\frac{t^2}{\tau^2\nu^2}} = \alpha(t)^{\frac{\lambda^2\mu^2}{\nu^2(\mu^2-\lambda^2)}}$$

we can simplify $\phi(x/\nu)$ (with $\nu \in \{\lambda, \mu\}$) to

$$\begin{aligned}
\phi(t/\lambda) &= \frac{\tau^2}{2}\left(1 - e^{-\frac{t^2}{\lambda^2\tau^2}}\right) = \frac{\tau^2}{2}\left(1 - \alpha(t)^{\mu^2/(\mu^2-\lambda^2)}\right) \\
\phi(t/\mu) &= \frac{\tau^2}{2}\left(1 - e^{-\frac{t^2}{\mu^2\tau^2}}\right) = \frac{\tau^2}{2}\left(1 - \alpha(t)^{\lambda^2/(\mu^2-\lambda^2)}\right).
\end{aligned}$$

*Toshiba Research Europe, christopher.zach@crl.toshiba.co.uk
†University of Bordeaux, guillaume.bourmaud@u-bordeaux.fr

Consequently $\gamma(w)$ (cf. Eq. 12 in the main text) simplifies to:

$$\begin{aligned}\gamma(w) &= \frac{\tau^2}{2}\left(\lambda^2\left(1 - w^{\frac{\mu^2}{\mu^2-\lambda^2}}\right) - \mu^2 w\left(1 - w^{\frac{\lambda^2}{\mu^2-\lambda^2}}\right)\right) \\ &= \frac{\tau^2}{2}\left(\lambda^2 + \left(\mu^2 - \lambda^2\right) w^{\frac{\mu^2}{\mu^2-\lambda^2}} - \mu^2 w\right). \end{aligned} \tag{4}$$

We assume $\mu = s\lambda$ for $s > 1$, and therefore Eq. 4 further simplifies to

$$\gamma_{s\times}(w) = \frac{\lambda^2\tau^2}{2}\left(1 + w\left(\left(s^2 - 1\right) w^{\frac{1}{s^2-1}} - s^2\right)\right), \tag{5}$$

which shows item 1 of the proposition.

It is interesting to analyse the robust kernel induced by $\gamma_{s\times}$: the minimizer $w^*(x)$ is given by solving $\min_{w\in[0,1]} wx^2/2 + \gamma_{s\times}(w)$

$$w^*(x) = \left[1 - \frac{x^2}{s^2\lambda^2\tau^2}\right]_+^{s^2-1} \tag{6}$$

and the induced robust kernel is

$$\phi_{s\times;\lambda\tau}(x) = \frac{\lambda^2\tau^2}{2}\left(1 - \left[1 - \frac{x^2}{s^2\lambda^2\tau^2}\right]_+^{s^2}\right). \tag{7}$$

If $s^2 = 2$, we identify $\phi_{\sqrt{2}\times}(\cdot)$ with $\phi_{\text{ST}}(\cdot; \sqrt{2}\lambda\tau)$, and $\phi_{\sqrt{3}\times}(\cdot)$ (i.e. $s^2 = 3$) coincides with $\phi_{\text{Tukey}}(\cdot; \sqrt{3}\lambda\tau)$. A natural choice for $s$ is $s = 2$, which leads to consecutive doubling of $\lambda\tau$.

The choice $\mu = \sqrt{2}\lambda$ is particularly convenient, since in this case $\gamma(w)$ reduces to

$$\gamma(w) = \frac{\tau^2\lambda^2}{2}(w-1)^2. \tag{8}$$

The expression above can be identified with $\gamma_{STQ}(w; \sqrt{2}\tau\lambda)$, hence a Welsch kernel with scale $\sqrt{2}^L\tau$ can be "boosted" to scale $\tau$ by iterating with the smooth truncated kernel.

A natural more aggressive scheme to increase the shape parameter $\tau$ is given by a doubling approach, $\mu = 2\lambda$, in Eq. 4. In this setting we obtain

$$\gamma_{2\times}(w) = \frac{\tau^2\lambda^2}{2}\left(1 + w\left(3\sqrt[3]{w} - 4\right)\right). \tag{9}$$

together with the corresponding robust kernel $\phi_{2\times}$,

$$\phi_{2\times}(x;\tau) = \frac{\lambda^2\tau^2}{2}\left(1 - \left[1 - \frac{x^2}{4\lambda^2\tau^2}\right]_+^4\right). \tag{10}$$

To our knowledge this kernel is not among the standard kernels known in the literature, but its graph is empirically very close to $\phi_{\text{Tukey}}(x; \sqrt{3/4})$.

## 1.2 Lifting of the smooth truncated kernel w.r.t. a scaled smooth truncated kernel

If we lift the smooth truncated kernel,

$$\phi_{\text{ST},\tau}(x) = \frac{\tau^2}{4}\left(1 - \left[1 - \frac{x^2}{\tau^2}\right]_+^2\right) \qquad \omega_{\text{ST},\tau}(x) = \left[1 - x^2/\tau^2\right]_+, \tag{11}$$

w.r.t. a scaled version of itself, we obtain

$$\alpha(t) = \frac{\omega_{\text{ST},\tau}(t)}{\omega_{\text{ST},\tau}(t/\mu)} = \frac{\tau^2 - t^2}{\tau^2 - t^2/\mu^2} \tag{12}$$

for $t \in [-\tau, \tau]$, and therefore

$$\alpha^{-1}(w) = \tau \sqrt{\frac{1-w}{1-w/\mu^2}}. \tag{13}$$

Plugging this expression into Eq. 12 in the main text yields a closed-form expression for $\gamma(w)$,

$$\gamma_{\mathrm{ST},\tau,s}(w) = \frac{s^2\tau^2(w-1)^2}{4(s^2-w)}, \tag{14}$$

which completes the proof of item 2.

### 1.3 Lifting of the Geman-McClure kernel w.r.t. a scaled Geman-McClure kernel

In contrast to items 1 and 2 above we prove the claimed relation directly in the following. Let us define the following function

$$f(x,l) = l^2 \phi_{\mathrm{Gem},\tau_1}(x) + \frac{\tau_2^2}{2}(l-1)^2 = \frac{l^2}{2}\frac{\tau_1^2 x^2}{x^2+\tau_1^2} + \frac{\tau_2^2}{2}(l-1)^2 \tag{15}$$

Minimizing $f(x,l)$ w.r.t. $l$ we obtain:

$$0 = \frac{\partial f(x,l)}{\partial l} = l\frac{\tau_1^2 x^2}{x^2+\tau_1^2} + \tau_2^2(l-1) \tag{16}$$

Consequently

$$l^* = \operatorname*{argmin}_l f(x,l) = \frac{\tau_2^2}{\frac{\tau_1^2 x^2}{x^2+\tau_1^2} + \tau_2^2} \tag{17}$$

From this result we get

$$f(x,l^*) = \left(\frac{\tau_2^2}{\frac{\tau_1^2 x^2}{x^2+\tau_1^2} + \tau_2^2}\right)^2 \frac{1}{2}\frac{\tau_1^2 x^2}{x^2+\tau_1^2} + \frac{\tau_2^2}{2}\left(\frac{\tau_2^2}{\frac{\tau_1^2 x^2}{x^2+\tau_1^2} + \tau_2^2} - 1\right)^2 \tag{18}$$

However, we also have the following identity:

$$\left(\frac{\tau_2^2}{y+\tau_2^2}\right)^2 \frac{1}{2}y + \frac{\tau_2^2}{2}\left(\frac{\tau_2^2}{y+\tau_2^2} - 1\right)^2 = \left(\frac{\tau_2^2}{y+\tau_2^2}\right)^2 \frac{1}{2}y + \frac{\tau_2^2}{2}\left(\frac{y}{y+\tau_2^2}\right)^2 = \frac{1}{2}\frac{y\tau_2^2}{y+\tau_2^2} \tag{19}$$

Thus, defining $y = \frac{\tau_1^2 x^2}{x^2+\tau_1^2}$, we read

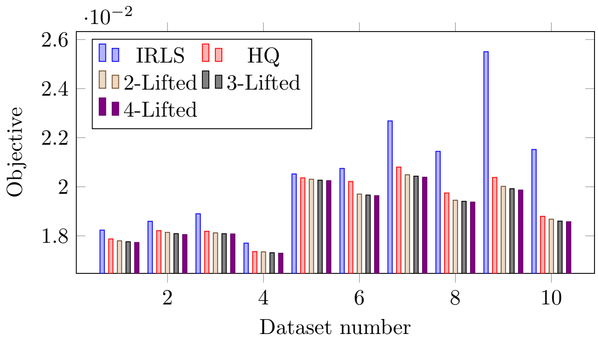
$$f(x,l^*) = \frac{1}{2}\frac{\frac{\tau_1^2 x^2}{x^2+\tau_1^2}\tau_2^2}{\frac{\tau_1^2 x^2}{x^2+\tau_1^2} + \tau_2^2} = \frac{1}{2}\frac{\frac{\tau_1^2\tau_2^2}{(\tau_1^2+\tau_2^2)}x^2}{x^2 + \frac{\tau_1^2\tau_2^2}{(\tau_1^2+\tau_2^2)}} = \phi_{\mathrm{Gem},\frac{\tau_1\tau_2}{\sqrt{\tau_1^2+\tau_2^2}}}(x) \tag{20}$$

Finally, defining $\tau$ and $s$ such that $\tau = \frac{\tau_1\tau_2}{\sqrt{\tau_1^2+\tau_2^2}}$ and $\tau_1 = s\tau$, we obtain $\tau_2 = \frac{s\tau}{\sqrt{s^2-1}}$ which shows that

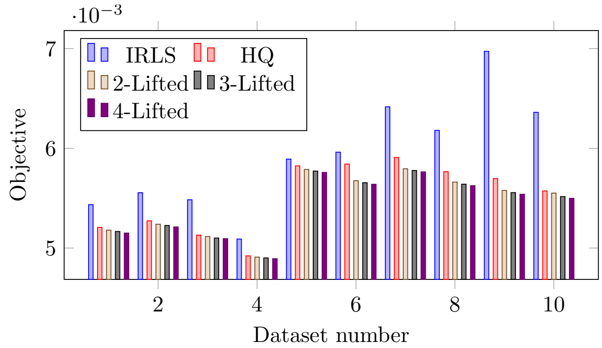
$$\phi_{\mathrm{Gem},\tau}(x) = \min_l l^2 \phi_{\mathrm{Gem},s\tau}(x) + \frac{s^2\tau^2}{2(s^2-1)}(l-1)^2 \tag{21}$$

## 2 Additional experimental results

In Figs. 1-3 we depict the final objectives reached for bundle adjustment problems (the same instances as in the main text). In order to assess the influence of non-linearities and local minima of the original, non-robust bundle objective, we illustrate results for *linearized* residuals in Fig. 1, and increase the non-linearity of the underlying problem in Fig. 2 (metric bundle adjustment) and Fig. 3 (additionally optimize over camera calibration parameters).

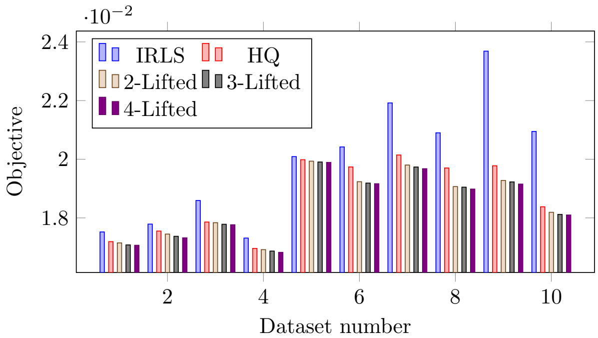



(a) $\tau = 1/4$

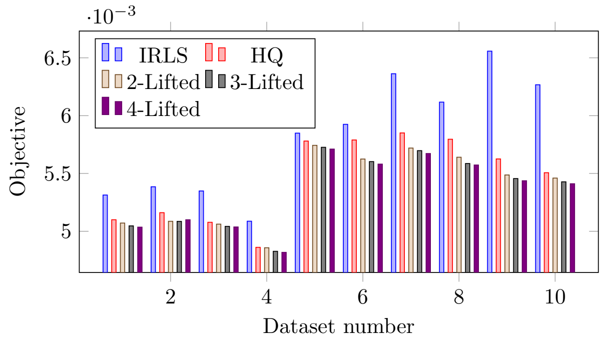



(b) $\tau = 1/8$

Figure 1: Linearized bundle objective

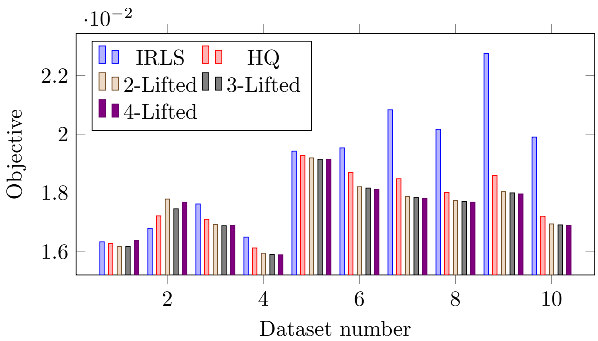



(a) $\tau = 1/4$

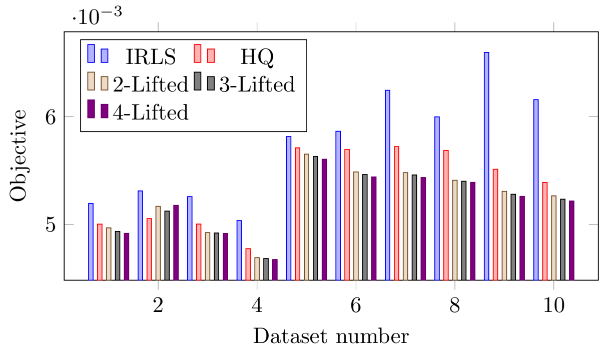



(b) $\tau = 1/8$

Figure 2: Metric bundle adjustment




(a) $\tau = 1/4$




(b) $\tau = 1/8$

Figure 3: Full bundle adjustment (including focal length and lens distortion parameters).