Sublabel–Accurate Relaxation of Nonconvex Energies
Supplementary Material

Thomas Möllenhoff* TU München moellenh@in.tum.de
Emanuel Laude* TU München laudee@in.tum.de
Michael Moeller TU München moellerm@in.tum.de
Jan Lellmann University of Lübeck lellmann@mic.uni-luebeck.de
Daniel Cremers TU München cremers@tum.de

Proof of Proposition 1. The proof follows from a direct calculation. We start with the definition of the biconjugate:

$$\begin{aligned}\boldsymbol{\rho}^{**}(\boldsymbol{u}) &= \sup_{\boldsymbol{v}\in\mathbb{R}^k} \langle \boldsymbol{u},\boldsymbol{v}\rangle - \left(\min_{1\le i\le k} \boldsymbol{\rho}_i(\boldsymbol{u})\right)^* \\ &= \sup_{\boldsymbol{v}\in\mathbb{R}^k} \langle \boldsymbol{u},\boldsymbol{v}\rangle - \max_{1\le i\le k} \boldsymbol{\rho}_i^*(\boldsymbol{u}).\end{aligned} \tag{1}$$

This shows the first equation inside the proposition. For the individual $\boldsymbol{\rho}_i^*$ we again start with the definition of the convex conjugate:

$$\begin{aligned}\boldsymbol{\rho}_i^*(\boldsymbol{v}) &= \sup_{\alpha\in[0,1]} \langle \alpha \mathbf{1}_i + (1-\alpha)\mathbf{1}_{i-1}, \boldsymbol{v}\rangle - \\ &\qquad \rho(\alpha\gamma_{i+1} + (1-\alpha)\gamma_i) \\ &= \sup_{\alpha\in[0,1]} \langle \mathbf{1}_{i-1}, \boldsymbol{v}\rangle + \alpha \boldsymbol{v}_i - \rho(\gamma_i^\alpha).\end{aligned} \tag{2}$$

Applying the substitution $\gamma_i^\alpha = \alpha\gamma_{i+1} + (1-\alpha)\gamma_i$ and consequently $\alpha = \frac{\gamma_i^\alpha - \gamma_i}{\gamma_{i+1}-\gamma_i}$ yields:

$$\begin{aligned}\boldsymbol{\rho}_i^*(\boldsymbol{v}) &= \sup_{\gamma_i^\alpha\in\Gamma_i} \langle \mathbf{1}_{i-1}, \boldsymbol{v}\rangle + \frac{\gamma_i^\alpha - \gamma_i}{\gamma_{i+1}-\gamma_i}\boldsymbol{v}_i - \rho(\gamma_i^\alpha) \\ &= \langle \mathbf{1}_{i-1}, \boldsymbol{v}\rangle - \frac{\gamma_i}{\gamma_{i+1}-\gamma_i}\boldsymbol{v}_i + \sup_{\gamma_i^\alpha\in\Gamma_i} \gamma_i^\alpha \frac{\boldsymbol{v}_i}{\gamma_{i+1}-\gamma_i} - \rho(\gamma_i^\alpha) \\ &= \langle \mathbf{1}_{i-1}, \boldsymbol{v}\rangle - \frac{\gamma_i}{\gamma_{i+1}-\gamma_i}\boldsymbol{v}_i + (\rho + \delta_{\Gamma_i})^* \left(\frac{\boldsymbol{v}_i}{\gamma_{i+1}-\gamma_i}\right) \\ &=: c_i(\boldsymbol{v}) + \rho_i^*\left(\frac{\boldsymbol{v}_i}{\gamma_{i+1}-\gamma_i}\right).\end{aligned} \tag{3}$$

□

*Those authors contributed equally.

Proof of Proposition 2. It is easy to see that

$$\boldsymbol{\sigma}^*(\boldsymbol{v}) = \max_{i\in\{1,\ldots,L\}} \left(\sum_{l=1}^{i-1} \boldsymbol{v}_l - \rho(\gamma_i)\right).$$

To compute the biconjugate, we write any input argument $\boldsymbol{u} = \sum_{i=1}^k \mu_i \mathbf{1}_{i+1}$, and use $\boldsymbol{\sigma}^{**} = \boldsymbol{\rho}^{**}$ to obtain

$$\begin{aligned}\boldsymbol{\rho}^{**}(\boldsymbol{u}) &= \sup_{\boldsymbol{v}} \langle \boldsymbol{u},\boldsymbol{v}\rangle - \max_{i\in\{1,\ldots,L\}} \left(\sum_{l=1}^{i-1} \boldsymbol{v}_l - \rho(\gamma_i)\right) \\ &= \sup_{\boldsymbol{v}} \sum_{i=1}^k \mu_i \sum_{l=1}^i \boldsymbol{v}_l - \max_{i\in\{1,\ldots,L\}} \left(\sum_{l=1}^{i-1} \boldsymbol{v}_l - \rho(\gamma_i)\right).\end{aligned}$$

Instead of taking the supremum of all $\boldsymbol{v}$, we might as well take the supremum over all vectors $\mathbf{p}$ with $\mathbf{p}_i = \sum_{l=1}^i \boldsymbol{v}_l$. Care has to be taken of the first summand in the second term of the above formulation. We obtain

$$\begin{aligned}&\sup_{\boldsymbol{v}} \sum_{i=1}^k \mu_i \sum_{l=1}^i \boldsymbol{v}_l - \max_{i\in\{1,\ldots,L\}} \left(\sum_{l=1}^{i-1} \boldsymbol{v}_l - \rho(\gamma_i)\right), \\ &= \sup_{\mathbf{p}} \sum_{i=1}^k \mu_i \mathbf{p}_i - \max_{i\in\{2,\ldots,L\}} \max(\mathbf{p}_{i-1} - \rho(\gamma_i), -\rho(\gamma_1)), \\ &= \sup_{\mathbf{p}} \sum_{i=1}^k \mu_i \mathbf{p}_i - \max_{i\in\{1,\ldots,k\}} \max(\mathbf{p}_i - \rho(\gamma_{i+1}), -\rho(\gamma_1)), \\ &= \sum_{i=1}^k \mu_i \; \rho(\gamma_{i+1}) \\ &\quad + \sup_{\mathbf{p}} \sum_{i=1}^k \mu_i \mathbf{p}_i - \max_{i\in\{1,\ldots,k\}} \max(\mathbf{p}_i, -\rho(\gamma_1)),\end{aligned}$$

Note that for any μ_i being negative, the supremum immediately yields infinity by taking $\mathbf{p}_i \to -\infty$. Similarly, if $\sum_{i=1}^k \mu_i > 1$ yields infinity by taking all $\mathbf{p}_i \to \infty$. For $\mu_i \geq 0$ for all i, and $\sum_{i=1}^k \mu_i \leq 1$, we know that $\sum_{i=1}^k \mu_i \mathbf{p}_i \leq (\max_i \mathbf{p}_i) \sum_{i=1}^k \mu_i$. Since equality can be obtained by choosing $\mathbf{p}_l = \max_i \mathbf{p}_i$ for all l, we can reduce the above supremum to

$$\sup_z \left(z \sum_{i=1}^k \mu_i - \max(z, -\rho(\gamma_1)) \right) = \left(1 - \sum_{i=1}^k \mu_i \right) \rho(\gamma_1),$$

where we used that the supremum over z is attained at $z = -\rho(\gamma_1)$ (still assuming that $\sum_{i=1}^k \mu_i \leq 1$). Let us now undo our change of variable. It is easy to see that $\mu_k = \boldsymbol{u}_k$, and $\mu_i = \boldsymbol{u}_i - \boldsymbol{u}_{i+1}$ for $i = 1, ..., k-1$. The latter leads to

$$\begin{aligned}
&\sum_{i=1}^k \mu_i \, \rho(\gamma_{i+1}) + \left(1 - \sum_{i=1}^k \mu_i \right) \rho(\gamma_1) \\
&= \rho(\gamma_{k+1}) \boldsymbol{u}_k + \sum_{i=1}^{k-1} (\boldsymbol{u}_i - \boldsymbol{u}_{i+1}) \, \rho(\gamma_{i+1}) + (1 - \boldsymbol{u}_1) \rho(\gamma_1) \\
&= \rho(\gamma_1) + \langle \boldsymbol{u}, \mathbf{r} \rangle,
\end{aligned}$$

for $\mathbf{r}_i = \rho(\gamma_{i+1}) - \rho(\gamma_i)$. Considering the aforementioned constraints of $\mu_i \geq 0$, and $\sum_{i=1}^k \mu_i \leq 1$, we finally find

$$\boldsymbol{\rho}^{**}(\boldsymbol{u}) = \begin{cases} \rho(\gamma_1) + \langle \boldsymbol{u}, \mathbf{r} \rangle & \text{if } 1 \geq \boldsymbol{u}_1 \geq ... \geq \boldsymbol{u}_k \geq 0, \\ \infty, & \text{else.} \end{cases}$$

□

Proof of Proposition 3. For the special case $k = 1$ the biconjugate from (1) is just:

$$\boldsymbol{\rho}^{**}(\boldsymbol{u}) = \sup_{\boldsymbol{v} \in \mathbb{R}} \boldsymbol{u}\boldsymbol{v} - \boldsymbol{\rho}_1^*(\boldsymbol{v}) = \boldsymbol{\rho}_1^{**}(\boldsymbol{u}). \tag{4}$$

Now using the first line in (3), $\boldsymbol{\rho}_1^{**}$ becomes:

$$\begin{aligned}
\boldsymbol{\rho}_1^{**}(\boldsymbol{u}) &= \sup_{\boldsymbol{v} \in \mathbb{R}} \boldsymbol{u}\boldsymbol{v} - \sup_{\gamma \in \Gamma} \frac{\gamma - \gamma_1}{\gamma_2 - \gamma_1} \boldsymbol{v} - \rho(\gamma) \\
&= \sup_{\boldsymbol{v} \in \mathbb{R}} \boldsymbol{v} \left(\boldsymbol{u} + \frac{\gamma_1}{\gamma_2 - \gamma_1} \right) - \sup_{\gamma \in \Gamma} \gamma \frac{\boldsymbol{v}}{\gamma_2 - \gamma_1} - \rho(\gamma) \\
&= \sup_{\boldsymbol{v} \in \mathbb{R}} \boldsymbol{v} \left(\boldsymbol{u} + \frac{\gamma_1}{\gamma_2 - \gamma_1} \right) - \rho^* \left(\frac{\boldsymbol{v}}{\gamma_2 - \gamma_1} \right) \\
&= \sup_{\tilde{\boldsymbol{v}} \in \mathbb{R}} \tilde{\boldsymbol{v}}(\gamma_1 + \boldsymbol{u}(\gamma_2 - \gamma_1)) - \rho^*(\tilde{\boldsymbol{v}}) \\
&= \rho^{**}(\gamma_1 + \boldsymbol{u}(\gamma_2 - \gamma_1)),
\end{aligned} \tag{5}$$

where we used $\text{dom}(\rho) = \Gamma$ as well as the substitution $\boldsymbol{v} = (\gamma_2 - \gamma_1)\tilde{\boldsymbol{v}}$. □

Proof of Proposition 4. We compute the individual conjugate as:

$$\begin{aligned}
\boldsymbol{\Phi}_{i,j}^*(\boldsymbol{q}) &= \sup_{\boldsymbol{g} \in \mathbb{R}^{d \times k}} \langle g, \boldsymbol{q} \rangle - \boldsymbol{\Phi}_{i,j}(\boldsymbol{q}) \\
&= \sup_{\alpha, \beta \in [0,1]} \sup_{\nu \in \mathbb{R}^d} \langle \boldsymbol{q}, (\mathbf{1}_i^\alpha - \mathbf{1}_j^\beta) \nu^\top \rangle - \left| \gamma_i^\alpha - \gamma_j^\beta \right| |\nu|_2 \\
&= \sup_{\alpha, \beta \in [0,1]} \sup_{\nu \in \mathbb{R}^d} \langle \boldsymbol{q}^\top (\mathbf{1}_i^\alpha - \mathbf{1}_j^\beta), \nu \rangle - \left| \gamma_i^\alpha - \gamma_j^\beta \right| |\nu|_2 \\
&= \sup_{\alpha, \beta \in [0,1]} \sup_{\nu \in \mathbb{R}^d} \langle \boldsymbol{q}^\top (\mathbf{1}_i^\alpha - \mathbf{1}_j^\beta), \nu \rangle - \left| \gamma_i^\alpha - \gamma_j^\beta \right| |\nu|_2.
\end{aligned} \tag{6}$$

The inner maximum over ν is the conjugate of the ℓ_2-norm scaled by $\left| \gamma_i^\alpha - \gamma_j^\beta \right|$ evaluated at $\boldsymbol{q}^\top \left(\mathbf{1}_i^\alpha - \mathbf{1}_j^\beta \right)$. This yields:

$$\boldsymbol{\Phi}_{i,j}^*(\boldsymbol{q}) = \begin{cases} 0, & \text{if } \left| \boldsymbol{q}^\top \left(\mathbf{1}_i^\alpha - \mathbf{1}_j^\beta \right) \right|_2 \leq \left| \gamma_i^\alpha - \gamma_j^\beta \right|, \\ & \qquad \forall \alpha, \beta \in [0,1], \\ \infty, & \text{else.} \end{cases} \tag{7}$$

For the overall biconjugate we have:

$$\begin{aligned}
\boldsymbol{\Phi}^{**}(\boldsymbol{g}) &= \sup_{\boldsymbol{q} \in \mathbb{R}^{k \times d}} \langle \boldsymbol{q}, \boldsymbol{g} \rangle - \max_{1 \leq i,j \leq k} \boldsymbol{\Phi}_{i,j}^*(\boldsymbol{q}) \\
&= \sup_{\boldsymbol{q} \in \mathcal{K}} \langle \boldsymbol{q}, \boldsymbol{g} \rangle.
\end{aligned} \tag{8}$$

Since we have the max over all $1 \leq i, j \leq k$ conjugates, the set $\mathcal{K}$ is given as the intersection of the sets described by the individual indicator functions $\boldsymbol{\Phi}_{i,j}$:

$$\begin{aligned}
\mathcal{K} = \Big\{ &\boldsymbol{q} \in \mathbb{R}^{k \times d} \mid \\
&\left| \boldsymbol{q}^\top (\mathbf{1}_i^\alpha - \mathbf{1}_j^\beta) \right|_2 \leq \left| \gamma_i^\alpha - \gamma_j^\beta \right|, \\
&\forall\, 1 \leq i \leq j \leq k,\ \forall \alpha, \beta \in [0,1] \Big\}.
\end{aligned} \tag{9}$$

□

Proof of Proposition 5. First we rewrite (9) by expanding the matrix-vector product into sums:

$$\begin{aligned}
&\left| \sum_{l=j}^{i-1} \boldsymbol{q}_l + \alpha \boldsymbol{q}_i - \beta \boldsymbol{q}_j \right|_2 \leq \left| \gamma_i^\alpha - \gamma_j^\beta \right|, \\
&\forall\, 1 \leq j \leq i \leq k,\ \forall \alpha, \beta \in [0,1].
\end{aligned} \tag{10}$$

Since the other cases for $1 \leq i \leq j \leq k$ in (9) are equivalent to (10), it is enough to consider (10) instead of (9).

Let $\gamma_1 < \gamma_2 < \ldots < \gamma_L$. In the following, we will show the equivalences:

(10)

$$\Leftrightarrow$$

$$\left|\sum_{l=j}^{i} \boldsymbol{q}_l\right|_2 \leq \gamma_{i+1} - \gamma_j,\ \forall\, 1 \leq j \leq i \leq k. \tag{11}$$

$$\Leftrightarrow$$

$$|\boldsymbol{q}_i|_2 \leq \gamma_{i+1} - \gamma_i,\ \forall\, 1 \leq i \leq k. \tag{12}$$

The direction "(10) $\Rightarrow$ (11)" follows by setting $\alpha = 1$ and $\beta = 0$ in (10), and "(11) $\Rightarrow$ (12)" follows by setting $i = j$ in (11).

The direction "(12) $\Rightarrow$ (11)" can be proven by a quick calculation:

$$\left|\sum_{l=j}^{i} \boldsymbol{q}_l\right|_2 \leq \sum_{l=j}^{i} |\boldsymbol{q}_l|_2 \leq \sum_{l=j}^{i} \gamma_{l+1} - \gamma_l = \gamma_{i+1} - \gamma_j. \tag{13}$$

It remains to show "(11) $\Rightarrow$ (10)". We start with the case $j = i$:

$$\begin{aligned} |\alpha \boldsymbol{q}_i - \beta \boldsymbol{q}_i|_2 &= |\alpha - \beta||\boldsymbol{q}_i|_2 \\ &\leq |\alpha - \beta|(\gamma_{i+1} - \gamma_i) \\ &= |(\gamma_{i+1} - \gamma_i)\alpha - (\gamma_{i+1} - \gamma_i)\beta| \\ &= |(\alpha - \beta)(\gamma_{i+1} - \gamma_i)| = |\gamma_i^\alpha - \gamma_i^\beta|. \end{aligned} \tag{14}$$

Now let $j < i$. Since $\gamma_j < \gamma_i$ it also holds that $\gamma_j^\beta \leq \gamma_i^\alpha$, thus it is equivalent to show (10) without the absolute value on the right hand side.

First we show that "(11) $\Rightarrow$ (10)" for $\beta \in \{0, 1\}$ and $\alpha \in [0, 1]$:

$$\begin{aligned} &\left|\sum_{l=j+1}^{i-1} \boldsymbol{q}_l + \alpha \boldsymbol{q}_i + (1-\beta)\boldsymbol{q}_j\right|_2 \\ &\leq \left|\sum_{l=j+1}^{i-1} \boldsymbol{q}_l + (1-\beta)\boldsymbol{q}_j\right|_2 + \alpha|\boldsymbol{q}_i|_2 \\ &\overset{\text{for } \beta=0 \text{ or } \beta=1}{\leq} \gamma_i - \gamma_j^\beta + \alpha(\gamma_{i+1} - \gamma_i) \\ &= \gamma_i^\alpha - \gamma_j^\beta. \end{aligned} \tag{15}$$

Using a similar argument we show that, using the above, "(11) $\Rightarrow$ (10)" for all $\alpha, \beta \in [0, 1]$.

$$\begin{aligned} &\left|\sum_{l=j+1}^{i-1} \boldsymbol{q}_l + \alpha \boldsymbol{q}_i + (1-\beta)\boldsymbol{q}_j\right|_2 \\ &\leq \left|\sum_{l=j+1}^{i-1} \boldsymbol{q}_l + \alpha \boldsymbol{q}_i\right|_2 + (1-\beta)|\boldsymbol{q}_j|_2 \\ &\overset{\text{using (15)}, \beta=1}{\leq} \gamma_i^\alpha - \gamma_{j+1} + (1-\beta)(\gamma_{j+1} - \gamma_j) \\ &= \gamma_i^\alpha - \gamma_j^\beta. \end{aligned} \tag{16}$$

$\square$