Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 16, No. 2, pp. 289–313, 2020

FORMULAE OF EPIGRAPHICAL PROJECTION FOR SOLVING MINIMAX LOCATION PROBLEMS*

Gert Wanka and Oleg Wilfer

Abstract: We are interested in a numerical method for solving extended multifacility minimax location problems introduced by Drezner in 1991. For this purpose, we present some formulae of projections onto the epigraphs of the sum of powers of weighted norms and onto the epigraphs of gauges. By bringing the extended multifacility location problem into a form of an unconstrained optimization problem where its objective function is a sum of functions allows us then to use the parallel splitting algorithm in combination with the introduced projection formulae to solve this kind of location problems. Numerical experiments document the usefulness of our approach for the discussed location problems.

Key words: *gauges, continuous minimax multifacility location problems, epigraphical projection, projection operators*

Mathematics Subject Classification: *49J35, 65K10, 90B85, 90C47*

1 Introduction and Preliminaries

As argued in a large number of papers, the proximal method is an excellent tool for solving in an efficient way optimization problems of the form

$$\min_{x\in\mathcal{H}} \left\{ \sum_{i=1}^{n} f_i(x) \right\}, \tag{1.1}$$

where $\mathcal{H}$ is a real Hilbert space equipped with the scalar product $\langle\cdot,\cdot\rangle_{\mathcal{H}}$, where the associated norm $\|\cdot\|_{\mathcal{H}}$ is defined by $\|y\|_{\mathcal{H}} := \sqrt{\langle y,y\rangle_{\mathcal{H}}}$ for all $y \in \mathcal{H}$ and $f_i : \mathcal{H} \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ is a proper, lower semicontinuous and convex function, $i = 1,...,n$. At this point let us recall that for a given function $f : \mathcal{H} \to \overline{\mathbb{R}}$, its *effective domain* is $\operatorname{dom} f = \{x \in \mathcal{H} : f(x) \leq +\infty\}$ and its *epigraph* $\operatorname{epi} f = \{(x,r) \in \mathcal{H}\times\mathbb{R} : f(x) \leq r\}$. We call the function f *proper* when $f(x) > -\infty$ for all $x \in \mathcal{H}$ and $\operatorname{dom} f \neq \emptyset$, *lower semicontinuous at* $\overline{x} \in X$ if $\liminf_{x\to\overline{x}} f(x) \geq f(\overline{x})$ and when the function f is lower semicontinuous at all $x \in X$, then we call it *lower semicontinuous* (l.s.c. for short).
Optimization problems of the form (1.1) occur for instance in areas like image processing [2,8,9,12], portfolio optimization [4,17], cluster analysis [3,11], statistical learning theory [10], machine learning [6] and location theory [4, 7, 14, 16]. In the main step of the proximal method it is necessary to determine the proximity operators of the functions involved in the

*This work was supported by the DFG (the German Research Foundation) under Project WA922/9-1.

formulation of the associated optimization problem. The *proximity operator* (a.k.a. *proximal mapping*) of a proper, lower semicontinuous and convex function $f:\mathcal{H}\to\overline{\mathbb{R}}$ denoted by prox_f is defined by

$$\operatorname{prox}_f x:\mathcal{H}\to\mathcal{H},\ \ \operatorname{prox}_f x:=\operatorname*{arg\,min}_{y\in\mathcal{H}}\left\{f(y)+\frac{1}{2}\|x-y\|_{\mathcal{H}}^2\right\}\ \forall x\in\mathcal{H}. \tag{1.2}$$

The proximity operator can be understood as a generalization of the projection onto a convex set, as for a non-empty, closed and convex set $A\subseteq\mathcal{H}$ we have

$$\operatorname{prox}_{\delta_A} x=\mathrm{P}_A\, x\ \forall x\in\mathcal{H}, \tag{1.3}$$

where $\delta_A:\mathcal{H}\to\overline{\mathbb{R}}$ defined by

$$\delta_A(x):=\begin{cases}0, & \text{if } x\in A,\\ +\infty, & \text{otherwise,}\end{cases} \tag{1.4}$$

is a proper, convex and lower semicontinuous indicator function and P_A is the projection operator which maps every point x in $\mathcal{H}$ to its unique projection onto the set A (see [1]).

From (1.2) follows that the determination of the proximity operators of the functions f_i, $i=1,...,n$, of (1.1) requires the solving of n subproblems, where a favorable situation exists, when a closed formula of a proximity operator can be given. This in turn has a positive effect on the solving of optimization problems from the numerical point of view.

Motivated by this background, our aim is to solve numerically extended multifacility minimax location problems (see [15]) given by

$$(EP_N^{M,\beta})\qquad \min_{(x_1,\dots,x_m)\in\mathbb{R}^d\times\cdots\times\mathbb{R}^d}\ \max_{1\le i\le n}\left\{\sum_{j=1}^m w_{ij}\|x_j-p_i\|^{\beta_i}\right\}, \tag{1.5}$$

where $w_{ij}>0,\ \beta_i\ge 1$ and $p_i\in\mathbb{R}^d$ are distinct points, $j=1,...,m,\ i=1,...,n$. In this framework we first need to rewrite this kind of location problems into the form of (1.1) where the objective function is a sum of lower semicontinuous convex functions. For this purpose we introduce an additional variable and obtain for $(EP_N^{M,\beta})$ the following formulation

$$\begin{aligned}(EP_N^{M,\beta})\quad &\min_{\substack{(x_1,\dots,x_m,t)\in\mathbb{R}^d\times\cdots\times\mathbb{R}^d\times\mathbb{R},\\ \sum_{j=1}^m w_{ij}\|x_j-p_i\|^{\beta_i}\le t,\ i=1,\dots,n}} t = \min_{\substack{(x_1,\dots,x_m,t)\in\mathbb{R}^d\times\cdots\times\mathbb{R}^d\times\mathbb{R},\\ (x_1,\dots,x_m,t)\in\operatorname{epi}\left(\sum_{j=1}^m w_{ij}\|\cdot-p_i\|^{\beta_i}\right),\ i=1,\dots,n}} t\\ &= \min_{(x_1,\dots,x_m,t)\in\mathbb{R}^d\times\cdots\times\mathbb{R}^d\times\mathbb{R}}\left\{t+\sum_{i=1}^n\delta_{\operatorname{epi}\left(\sum_{j=1}^m w_{ij}\|\cdot-p_i\|^{\beta_i}\right)}(x_1,...,x_m,t)\right\}.\end{aligned} \tag{1.6}$$

Now, to apply the proximal method to $(EP_N^{M,\beta})$ one needs to calculate the proximity operators of the functions involved in the objective function of (1.6). For this reason and especially in the context of (1.3), we give in Section 2 formulae for the projection onto the epigraph of the sum of powers of weighted norms. As the power of norm in (1.6) can be replaced by a gauge function, we present also formulae of projections onto the epigraphs of gauges.

To point out the benefits of the presented formulae we consider then examples of location problems in different settings and compare the numerical results with a method proposed by

Cornejo and Michelot in [14]. The difference between these two methods is that the one given by Cornejo and Michelot splits the sum of powers of weighted norms by introducing $n \cdot m$ additional variables. In this situation one gets the following presentation of the extended multifacility minimax location problem

$$(EP_N^{M,\beta}) \quad \min_{\substack{t,\ t_{ij}\in\mathbb{R},\ x_j\in\mathbb{R}^d,\\ j=1,\dots,m,i=1,\dots,n}} \left\{ t + \sum_{j=1}^{m}\sum_{i=1}^{n} \delta_{\operatorname{epi}\left(w_{ij}\|\cdot-p_i\|^{\beta_i}\right)}(x_j,t_{ij}) + \sum_{i=1}^{n} \delta_{\operatorname{epi}\tau_i}(t_{i1},\dots,t_{im},t) \right\}, \tag{1.7}$$

where $\tau_i(t_{i1},...,t_{im}) := \sum_{j=1}^{m} t_{ij},\ i = 1,...,n$. In Section 3 we show that this concept makes the solving process for the considered examples of location problems very slow and the advantage of our approach more clear. The numerical tests are based on the parallel splitting algorithm, which can be found for instance in [1].

Finally, we collect some properties of Hilbert spaces, which can be found with proofs for instance in [1] and [13].

If for a function $f : \mathcal{H} \to \mathbb{R}$ we take an arbitrary $x \in \mathcal{H}$ such that $f(x) \in \mathbb{R}$, then we call the set

$$\partial f(x) := \{x^* \in \mathcal{H} : f(y) - f(x) \geq \langle x^*, y - x\rangle\ \forall y \in \mathcal{H}\}$$

the (convex) *subdifferential* of f at x, where the elements are called the *subgradients* of f at x. Moreover, if $\partial f(x) \neq \emptyset$, then we say that f is *subdifferentiable at* x and if $f(x) \notin \mathbb{R}$, then we make the convention that $\partial f(x) := \emptyset$. If f is Gâteaux-differentiable at $x \in \mathcal{H}$, then $\partial f(x) = \{\nabla f(x)\}$. The set of global minimizers of the function f is denoted by $\operatorname{Argmin} f$ and if f has a unique minimizer, it is denoted by $\arg\min_{x\in\mathcal{H}} f(x)$. It holds

$$x \in \operatorname{Argmin} f \Leftrightarrow 0_{\mathcal{H}} \in \partial f(x)\ \forall x \in \mathcal{H}. \tag{1.8}$$

It holds

$$y = \operatorname{prox}_f x \Leftrightarrow x - y \in \partial f(y)\ \forall x \in \mathcal{H},\ \forall y \in \mathcal{H}. \tag{1.9}$$

In addition, we make for the rest of this paper the convention that $\frac{0}{0} = 0$ and $\frac{1}{0}\cdot 0_{\mathcal{H}} = 0_{\mathcal{H}}$.

In the following let $\mathcal{H}_1 \times \cdots \times \mathcal{H}_n$ be a real Hilbert space endowed with inner product and norm, respectively, defined by

$$\begin{aligned} \langle (x_1,...,x_n),(y_1,...,y_n)\rangle_{\mathcal{H}_1\times\cdots\times\mathcal{H}_n} &= \sum_{i=1}^{n}\langle x_i,y_i\rangle_{\mathcal{H}_i} \text{ and} \\ \|(x_1,...,x_n)\|_{\mathcal{H}_1\times\cdots\times\mathcal{H}_n} &= \sqrt{\sum_{i=1}^{n}\|x_i\|^2_{\mathcal{H}_i}}, \end{aligned}$$

where $(x_1,...,x_n) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n$ and $(y_1,...,y_n) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n$.

We close this section with a lemma, which presents a formula for the projection onto a unit ball generated by the weighted sum of norms and generalizes the results given in [18] to real Hilbert spaces $\mathcal{H}_i,\ i = 1,...,n$. Let $w_i > 0,\ i = 1,...,n$, and $C := \{(x_1,...,x_n) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n : \sum_{i=1}^{n} w_i\|x_i\|_{\mathcal{H}_i} \leq 1\}$, then the following statement holds.

Lemma 1.1. *For all* $(x_1,...,x_n) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n$ *it holds*

$$\mathrm{P}_C(x_1,...,x_n) = \begin{cases} (x_1,...,x_n), & \text{if } \sum_{i=1}^{n} w_i\|x_i\|_{\mathcal{H}_i} \leq 1, \\ (\overline{y}_1,...,\overline{y}_n), & \text{otherwise,} \end{cases}$$

where

$$\overline{y}_i = \frac{\max\{\|x_i\|_{\mathcal{H}_i} - \overline{\lambda} w_i, 0\}}{\|x_i\|_{\mathcal{H}_i}} x_i,\ i = 1, ..., n,$$

with

$$\overline{\lambda} = \frac{\sum\limits_{i=k+1}^{n} w_i^2 \tau_i - 1}{\sum\limits_{i=k+1}^{n} w_i^2}$$

and $k \in \{0, 1, ..., n-1\}$ *is the unique integer such that* $\tau_k \leq \overline{\lambda} \leq \tau_{k+1}$, *where the values* $\tau_0, ..., \tau_n$ *are defined by* $\tau_0 := 0$ *and* $\tau_i := \|x_i\|_{\mathcal{H}_i}/w_i$, $i = 1, ..., n$, *and in ascending order.*

Proof. In order to determine the projection onto the set C, we consider for fixed $(x_1, ..., x_n) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n$ the following optimization problem

$$\min_{\substack{(y_1,...,y_n) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n, \\ \sum\limits_{i=1}^{n} w_i \|y_i\|_{\mathcal{H}_i} \leq 1}} \left\{ \sum_{i=1}^{n} \frac{1}{2} \|y_i - x_i\|_{\mathcal{H}_i}^2 \right\}. \tag{1.10}$$

Obviously, if $\sum_{i=1}^{n} w_i \|x_i\|_{\mathcal{H}_i} \leq 1$, i.e. $(x_1, ..., x_n) \in C$, then the unique solution is $\overline{y}_i = x_i$, $i = 1, ..., n$. In the following we consider the non-trivial situation where $\sum_{i=1}^{n} w_i \|x_i\|_{\mathcal{H}_i} > 1$, i.e. $(x_1, ..., x_n) \notin C$ and define the function $f : \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \to \mathbb{R}$ by $f(y_1, ..., y_n) := \sum_{i=1}^{n} (1/2) \|y_i - x_i\|_{\mathcal{H}_i}^2$ and the function $g : \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \to \mathbb{R}$ by $g(y_1, ..., y_n) := \sum_{i=1}^{n} w_i \|y_i\|_{\mathcal{H}_i} - 1$. Hence, by [1, Proposition 26.18] it holds for the unique solution $(\overline{y}_1, ..., \overline{y}_n)$ of (1.10) that

$$\nabla f(\overline{y}_1, ..., \overline{y}_n) \in -\overline{\lambda} \partial g(\overline{y}_1, ..., \overline{y}_n) \Leftrightarrow \overline{y}_i - x_i \in -\overline{\lambda} \partial \left(w_i \| \cdot \|_{\mathcal{H}_i} \right) (\overline{y}_i),\ i = 1, ..., n,$$

as well as

$$\overline{\lambda} \left(\sum_{i=1}^{n} w_i \|\overline{y}_i\|_{\mathcal{H}_i} - 1 \right) = 0 \text{ and } \sum_{i=1}^{n} w_i \|\overline{y}_i\|_{\mathcal{H}_i} \leq 1,$$

where $\overline{\lambda} \geq 0$ is the associated Lagrange multiplier of $(\overline{y}_1, ..., \overline{y}_n)$. If $\overline{\lambda} = 0$, then $\overline{y}_i = x_i$, $i = 1, ..., n$, and by the feasibility condition we obtain $\sum_{i=1}^{n} w_i \|x_i\|_{\mathcal{H}_i} \leq 1$, which contradicts our assumption. Therefore, $\overline{\lambda} > 0$ and we get by (1.9) that

$$\begin{aligned} \overline{y}_i - x_i \in -\overline{\lambda} \partial \left(w_i \| \cdot \|_{\mathcal{H}_i} \right) (\overline{y}_i) &\Leftrightarrow x_i - \overline{y}_i \in \partial (\overline{\lambda} w_i \| \cdot \|_{\mathcal{H}_i})(\overline{y}_i) \\ &\Leftrightarrow \overline{y}_i = \operatorname{prox}_{\overline{\lambda} w_i \| \cdot \|_{\mathcal{H}_i}} x_i,\ i = 1, ..., n. \end{aligned}$$

Using [13, Proposition 2.8] reveals that

$$\overline{y}_i = \begin{cases} x_i - \frac{\overline{\lambda} w_i}{\|x_i\|_{\mathcal{H}_i}} x_i, & \text{if } \|x_i\|_{\mathcal{H}_i} > \overline{\lambda} w_i, \\ 0_{\mathcal{H}_i}, & \text{if } \|x_i\|_{\mathcal{H}_i} \leq \overline{\lambda} w_i \end{cases} = \frac{\max\{\|x_i\|_{\mathcal{H}_i} - \overline{\lambda} w_i, 0\}}{\|x_i\|_{\mathcal{H}_i}} x_i,\ i = 1, ..., n,$$

and as $\sum_{i=1}^{n} w_i \|\overline{y}_i\|_{\mathcal{H}_i} = 1$, we conclude that

$$\sum_{i=1}^{n} w_i \max\left\{ \|x_i\|_{\mathcal{H}_i} - \overline{\lambda} w_i, 0 \right\} = 1. \tag{1.11}$$

Now, we define the function $\kappa : \mathbb{R} \to \mathbb{R}$ by $\kappa(\lambda) = \sum_{i=1}^{n} w_i^2 \max\{\tau_i - \lambda, 0\} - 1$. Note, that there exists $\widetilde{\lambda} \geq \tau_i$ for all $i = 1, ..., n$, such that $\kappa(\widetilde{\lambda}) = -1 < 0$. Moreover, κ is a piecewise linear function with $\kappa(0) = w_i^2\tau_i - 1$ and its slope changes at $\lambda = \tau_i$, $i = 1, ..., n$. To be more precise, at $\lambda = 0$ the slope of κ is $-\sum_{i=1}^{n} w_i^2$ and increases by w_1^2 when $\lambda = \tau_1$. If we continue in this matter for $i = 2, ..., n$, the slope keeps increasing and when $\lambda \geq \tau_n$, $\kappa(\lambda) = -1$ such that the slope is 0. In summary, to find the zero of κ one needs to determine the unique integer $k \in \{0, 1, ..., n-1\}$ such that $\kappa(\tau_k) \geq 0$ and $\kappa(\tau_{k+1}) \leq 0$. In the light of the above, it holds

$$\kappa(\lambda) = \sum_{i=k+1}^{n} w_i^2 \tau_i - \lambda \sum_{i=k+1}^{n} w_i^2 - 1,$$

where $\tau_k \leq \lambda \leq \tau_{k+1}$, and hence, one gets for $\overline{\lambda}$ such that $\kappa(\overline{\lambda}) = 0$,

$$\overline{\lambda} = \frac{\sum\limits_{i=k+1}^{n} w_i^2 \tau_i - 1}{\sum\limits_{i=k+1}^{n} w_i^2}.$$

□

2 Formulae of Epigraphical Projection

The first aim of this section is to give formulae for the projection operators onto the epigraph of the sum of powers of weighted norms. For this purpose, we give a general formula in our central theorem, from which we deduce special cases used in our numerical tests.

The second aim is to present formulae of the projection operators onto the epigraphs of gauges. By using the fact that the sum of gauges is again a gauge (see [21]), we also present a formula of the projector onto the epigraph of the sum of gauges.

2.1 Sum of weighted norms

Let us consider the following function $h : \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \to \mathbb{R}$ defined as

$$h(x_1, ..., x_n) := \sum_{i=1}^{n} w_i \|x_i\|_{\mathcal{H}_i}^{\beta_i}, \tag{2.1}$$

where $w_i > 0$ and $\beta_i \geq 1$, $i = 1, ..., n$. By defining the sets

$$L := \{l \in \{1, ..., n\} : \beta_l > 1\} \text{ and } R := \{r \in \{1, ..., n\} : \beta_r = 1\},$$

we can state the following formula for the projection onto the epigraph of the sum of powers of weighted norms, which generalizes the results given for instance in [1, 12, 13, 17].

Theorem 2.1. *Assume that h is given by (2.1). Then, for every $(x_1, ..., x_n, \xi) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \times \mathbb{R}$ one has*

$$\mathrm{P}_{\operatorname{epi} h}(x_1, ..., x_n, \xi) = \begin{cases} (x_1, ..., x_n, \xi), & \text{if } \sum\limits_{i=1}^{n} w_i \|x_i\|_{\mathcal{H}_i}^{\beta_i} \leq \xi, \\ (\overline{y}_1, ..., \overline{y}_n, \overline{\theta}), & \text{otherwise,} \end{cases} \tag{2.2}$$

with

$$\overline{y}_r = \frac{\max\{\|x_r\|_{\mathcal{H}_r} - \overline{\lambda} w_r, 0\}}{\|x_r\|_{\mathcal{H}_r}} x_r,\ r \in R,$$
$$\overline{y}_l = \frac{\|x_l\|_{\mathcal{H}_l} - \eta_l(\overline{\lambda})}{\|x_l\|_{\mathcal{H}_l}} x_l,\ l \in L,$$
$$\overline{\theta} = \xi + \overline{\lambda},$$

where $\eta_l(\overline{\lambda})$ *is the unique non-negative real number that solves the equation*

$$\eta_l(\overline{\lambda}) + \left(\frac{\eta_l(\overline{\lambda})}{\overline{\lambda} w_l \beta_l}\right)^{\frac{1}{\beta_l - 1}} = \|x_l\|_{\mathcal{H}_l},\ l \in L, \tag{2.3}$$

and $\overline{\lambda} > 0$ *is a solution of the equation*

$$\sum_{r \in R} w_r \max\{\|x_r\|_{\mathcal{H}_r} - \lambda w_r, 0\} + \sum_{l \in L} w_l (\|x_l\|_{\mathcal{H}_l} - \eta_l(\lambda))^{\beta_l} = \lambda + \xi. \tag{2.4}$$

Proof. For given $\xi \in \mathbb{R}$ and $(x_1, ..., x_n) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n$, let us consider the following optimization problem

$$\min_{\substack{(y_1,\ldots,y_n,\theta) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \times \mathbb{R} \\ \sum_{i=1}^n w_i \|y_i\|_{\mathcal{H}_i}^{\beta_i} \le \theta}} \left\{ \frac{1}{2}(\theta - \xi)^2 + \sum_{i=1}^n \frac{1}{2} \|y_i - x_i\|_{\mathcal{H}_i}^2 \right\}. \tag{2.5}$$

It is clear that in the situation when $\sum_{i=1}^n w_i \|x_i\|_{\mathcal{H}_i}^{\beta_i} \le \xi$, i.e. $(x_1, ..., x_n, \xi) \in \operatorname{epi} h$, the unique solution of (2.5) is $\overline{y}_i = x_i$, $i = 1, ..., n$, and $\overline{\theta} = \xi$. Therefore, we consider in the following the non-trivial case where $\sum_{i=1}^n w_i \|x_i\|_{\mathcal{H}_i}^{\beta_i} > \xi$, i.e. $(x_1, ..., x_n, \xi) \notin \operatorname{epi} h$.

Let us now define the function $f : \mathcal{H}_1 \times ... \times \mathcal{H}_n \times \mathbb{R} \to \mathbb{R}$ by $f(y_1, ..., y_n, \theta) := (1/2)(\theta - \xi)^2 + \sum_{i=1}^n (1/2) \|y_i - x_i\|_{\mathcal{H}_i}^2$ and the function $g : \mathcal{H}_1 \times ... \times \mathcal{H}_n \times \mathbb{R} \to \mathbb{R}$ by $g(y_1, ..., y_n, \theta) := \sum_{i=1}^n w_i \|y_i\|_{\mathcal{H}_i}^{\beta_i} - \theta$, then by [1, Proposition 26.18] there exists $\overline{\lambda} \ge 0$, such that for the unique solution $(\overline{y}_1, ..., \overline{y}_n, \overline{\theta})$ of (2.5) it holds

$$\nabla f(\overline{y}_1, ..., \overline{y}_n, \overline{\theta}) \in -\overline{\lambda} \partial g(\overline{y}_1, ..., \overline{y}_n, \overline{\theta}) \Leftrightarrow \begin{cases} \overline{y}_i - x_i \in -\overline{\lambda} \partial (w_i \| \cdot \|_{\mathcal{H}_i}^{\beta_i})(\overline{y}_i),\ i = 1, ...n, \\ \overline{\theta} - \xi = \overline{\lambda}, \end{cases} \tag{2.6}$$

where $\overline{\lambda}$ is the associated Lagrange multiplier of $(\overline{y}_1, ..., \overline{y}_n, \overline{\theta})$. If $\overline{\lambda} = 0$, then one gets by (2.6) that $\overline{y}_i = x_i$, $i = 1, ..., n$, and $\overline{\theta} = \xi$ and by the feasibility of the solution it follows that $\sum_{i=1}^n w_i \|x_i\|_{\mathcal{H}_i}^{\beta_i} \le \xi$, which contradicts our assumption. Hence, it holds $\overline{\lambda} > 0$ and by (1.9) and (2.6) we have

$$\begin{cases} x_i - \overline{y}_i \in \partial(\overline{\lambda} w_i \| \cdot \|_{\mathcal{H}_i}^{\beta_i})(\overline{y}_i),\ i = 1, ...n, \\ \overline{\theta} = \overline{\lambda} + \xi, \end{cases} \Leftrightarrow \begin{cases} \overline{y}_i = \operatorname{prox}_{\overline{\lambda} w_i \| \cdot \|_{\mathcal{H}_i}^{\beta_i}} x_i,\ i = 1, ...n, \\ \overline{\theta} = \overline{\lambda} + \xi. \end{cases}$$

Further, from [13, Proposition 2.8] it follows for the case $r \in R$, i.e. $\beta_r = 1$, that

$$\overline{y}_r = \begin{cases} x_r - \frac{\overline{\lambda} w_r}{\|x_r\|_{\mathcal{H}_r}} x_r, & \text{if } \|x_r\|_{\mathcal{H}_r} > \overline{\lambda} w_r, \\ 0_{\mathcal{H}_r}, & \text{if } \|x_r\|_{\mathcal{H}_r} \le \overline{\lambda} w_r \end{cases} = \frac{\max\{\|x_r\|_{\mathcal{H}_r} - \overline{\lambda} w_r, 0\}}{\|x_r\|_{\mathcal{H}_r}} x_r, \tag{2.7}$$

and for the case $l \in L$, i.e. $\beta_l > 0$, that

$$\overline{y}_l = x_l - \frac{\eta_l(\overline{\lambda})}{\|x_l\|_{\mathcal{H}_l}} x_l = \frac{\|x_l\|_{\mathcal{H}_l} - \eta_l(\overline{\lambda})}{\|x_l\|_{\mathcal{H}_l}} x_l, \tag{2.8}$$

where $\eta_l(\overline{\lambda})$ is the unique non-negative real number that solves the following equation

$$\eta_l(\overline{\lambda}) + \left(\frac{\eta_l(\overline{\lambda})}{\overline{\lambda} w_l \beta_l}\right)^{\frac{1}{\beta_l - 1}} = \|x_l\|_{\mathcal{H}_l} \tag{2.9}$$

(notice that by (2.9) follows that $\|x_l\|_{\mathcal{H}_l} - \eta_l(\overline{\lambda}) \geq 0$). Furthermore, the complementary slackness condition

$$\overline{\lambda}\left(\sum_{i=1}^{n} w_i \|\overline{y}_i\|_{\mathcal{H}_i}^{\beta_i} - \overline{\theta}\right) = 0 \tag{2.10}$$

implies that

$$\sum_{i=1}^{n} w_i \|\overline{y}_i\|_{\mathcal{H}_i}^{\beta_i} = \overline{\theta}, \tag{2.11}$$

and from here follows by (2.7) and (2.8) that

$$\sum_{i=1}^{n} w_i \|\overline{y}_i\|_{\mathcal{H}_i}^{\beta_i} = \sum_{r \in R} w_r \max\{\|x_r\|_{\mathcal{H}_r} - \overline{\lambda} w_r, 0\} + \sum_{l \in L} w_l (\|x_l\|_{\mathcal{H}_l} - \eta_l(\overline{\lambda}))^{\beta_l} = \overline{\lambda} + \xi. \tag{2.12}$$

□

Remark 2.2. In the situation when $\beta_i > 1$ for all i=1,...,n, we get by summarizing the formulae (2.3) and (2.4)

$$\begin{aligned} &\eta_i(\lambda) + \left(\frac{\eta_i(\lambda)}{w_i\beta_i\left(\sum_{j=1}^n w_j(\|x_j\|_{\mathcal{H}_j} - \eta_j(\lambda))^{\beta_j}\right) - w_i\beta_i\xi}\right)^{\frac{1}{\beta_i - 1}} = \|x_i\|_{\mathcal{H}_i} \\ \Leftrightarrow\ &\frac{\eta_i(\lambda)}{w_i\beta_i\left(\sum_{j=1}^n w_j(\|x_j\|_{\mathcal{H}_j} - \eta_j(\lambda))^{\beta_j}\right) - w_i\beta_i\xi} = (\|x_i\|_{\mathcal{H}_i} - \eta_i(\lambda))^{\beta_i - 1},\ i = 1, ..., n. \end{aligned} \tag{2.13}$$

By setting $\chi_i = \|x_i\|_{\mathcal{H}_i} - \eta_i(\lambda) \geq 0$, $i = 1, ..., n$, formula (2.13) can be expressed by

$$\begin{aligned} &\frac{\|x_i\|_{\mathcal{H}_i} - \chi_i}{w_i\beta_i\left(\sum_{j=1}^n w_j\chi_j^{\beta_j}\right) - w_i\beta_i\xi} = \chi_i^{\beta_i - 1} \\ \Leftrightarrow\quad & w_i\beta_i\chi_i^{\beta_i-1}\sum_{j=1}^n w_j\chi_j^{\beta_j} - \xi w_i\beta_i\chi_i^{\beta_i-1} + \chi_i = \|x_i\|_{\mathcal{H}_i} \\ \Leftrightarrow\quad & w_i^2\beta_i\chi_i^{2\beta_i-1} + w_i\beta_i\chi_i^{\beta_i-1}\sum_{\substack{j=1\\ j\neq i}}^n w_j\chi_j^{\beta_j} - \xi w_i\beta_i\chi_i^{\beta_i-1} + \chi_i = \|x_i\|_{\mathcal{H}_i},\ i = 1, ..., n. \end{aligned}$$

Hence, it holds for every $(x_1, ..., x_n, \xi) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \times \mathbb{R}$

$$\mathrm{P}_{\mathrm{epi}\, h}(x_1, ..., x_n, \xi) = \begin{cases} (x_1, ..., x_n, \xi), \text{ if } \sum\limits_{i=1}^n w_i\|x_i\|_{\mathcal{H}_i}^{\beta_i} \leq \xi, \\ (\overline{y}_1, ..., \overline{y}_n, \overline{\theta}), \text{ otherwise,} \end{cases}$$

with

$$\overline{y}_i = \frac{\overline{\chi}_i}{\|x_i\|_{\mathcal{H}_i}} x_i,\ i = 1, ...n,\ \text{and}\ \overline{\theta} = \sum_{i=1}^{n} w_i \overline{\chi}_i^{\beta_i},$$

where $\overline{\chi_i} \geq 0$, $i = 1, ..., n$, are the unique real numbers that solve a polynomial equation system of the form

$$w_i^2 \beta_i \chi_i^{2\beta_i - 1} + w_i \beta_i \chi_i^{\beta_i - 1} \sum_{\substack{j=1 \\ j \neq i}}^{n} w_j \chi_j^{\beta_j} - \xi w_i \beta_i \chi_i^{\beta_i - 1} + \chi_i = \|x_i\|_{\mathcal{H}_i},\ i = 1, ..., n.$$

Let us additionally mention that the case where $n = 1$ was considered for instance in [12].

An important consequence of Theorem 2.1 where $\beta_i = 1$ for all $i = 1, ..., n$, follows.

Corollary 2.3. *Let h be given by (2.1) where $\beta_i = 1$ for all $i = 1, ..., n$. Then for all $(x_1, ..., x_n, \xi) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \times \mathbb{R}$ it holds*

$$\mathrm{P}_{\mathrm{epi}\, h}(x_1, ..., x_n, \xi) = \begin{cases} (x_1, ..., x_n, \xi), & \text{if } \sum_{i=1}^{n} w_i \|x_i\|_{\mathcal{H}_i} \leq \xi, \\ (0_{\mathcal{H}_1}, ..., 0_{\mathcal{H}_n}, 0), & \text{if } \xi < 0 \text{ and } \|x_i\|_{\mathcal{H}_i} \leq -\xi w_i,\ i = 1, ..., n, \\ (\overline{y}_1, ..., \overline{y}_n, \overline{\theta}), & \text{otherwise}, \end{cases} \tag{2.14}$$

where

$$\overline{y}_i = \frac{\max\{\|x_i\|_{\mathcal{H}_i} - \overline{\lambda} w_i, 0\}}{\|x_i\|_{\mathcal{H}_i}} x_i,\ i = 1, ..., n,\ \text{and}\ \overline{\theta} = \xi + \overline{\lambda},$$

with

$$\overline{\lambda} = \frac{\sum_{i=k+1}^{n} w_i^2 \tau_i - \xi}{\sum_{i=k+1}^{n} w_i^2 + 1} \tag{2.15}$$

and $k \in \{0, 1, ..., n-1\}$ is the unique integer such that $\tau_k \leq \overline{\lambda} \leq \tau_{k+1}$, where the values $\tau_0, ..., \tau_n$ are defined by $\tau_0 := 0$ and $\tau_i := \|x_i\|_{\mathcal{H}_i}/w_i$, $i = 1, ..., n$ and in ascending order.

Proof. As $\beta_i = 1$ for all $i = 1, ..., n$, Theorem 2.1 yields

$$\mathrm{P}_{\mathrm{epi}\, h}(x_1, ..., x_n, \xi) = \begin{cases} (x_1, ..., x_n, \xi), & \text{if } \sum_{i=1}^{n} w_i \|x_i\|_{\mathcal{H}_i} \leq \xi, \\ (\overline{y}_1, ..., \overline{y}_n, \overline{\theta}), & \text{otherwise}, \end{cases}$$

with

$$\overline{y}_i = \frac{\max\{\|x_i\|_{\mathcal{H}_i} - \overline{\lambda} w_i, 0\}}{\|x_i\|_{\mathcal{H}_i}} x_i,\ i = 1, ..., n,\ \text{and}\ \overline{\theta} = \xi + \overline{\lambda},$$

where $\overline{\lambda} > 0$ is a solution of the equation

$$\sum_{i=1}^{n} w_i \max\{\|x_i\|_{\mathcal{H}_i} - \lambda w_i, 0\} = \lambda + \xi.$$

Now, we consider the case where $\sum_{i=1}^n w_i \|x_i\|_{\mathcal{H}_i} > \xi$ and distinguish two cases.

(a) Let $\xi < 0$. If $\|x_i\|_{\mathcal{H}_i} + \xi w_i \leq 0$ for all $i = 1, ..., n$, we have by $0 \leq \overline{\theta} = \xi + \overline{\lambda}$, i.e. $\xi \geq -\overline{\lambda}$, that

$$0 \geq \|x_i\|_{\mathcal{H}_i} + \xi w_i \geq \|x_i\|_{\mathcal{H}_i} - \overline{\lambda} w_i \;\forall i = 1, ..., n, \tag{2.16}$$

and from here follows that

$$\overline{\lambda} + \xi = \sum_{i=1}^n w_i \max\{\|x_i\|_{\mathcal{H}_i} - \overline{\lambda} w_i, 0\} = 0, \text{ i.e. } \overline{\lambda} = -\xi. \tag{2.17}$$

But this means that

$$(\overline{y}_1, ..., \overline{y}_n, \overline{\theta}) = (0_{\mathcal{H}_1}, ..., 0_{\mathcal{H}_n}, 0), \tag{2.18}$$

which verifies the second case of (2.14).

If we now assume that there exists $j \in \{1, ..., n\}$ such that $\|x_j\|_{\mathcal{H}_j} + \xi w_j > 0$, then we define the function $g : \mathbb{R} \to \mathbb{R}$ by

$$g(\lambda) := \sum_{i=1}^n w_i^2 \max\{\tau_i - \lambda, 0\} - \lambda - \xi. \tag{2.19}$$

Moreover, this assumption yields

$$g(\lambda) = \sum_{i=1}^n w_i^2 \max\{\tau_i - \lambda, 0\} - \lambda - \xi < \sum_{i=1}^n w_i^2 \max\{\tau_i - \lambda, 0\} - \lambda + \frac{\|x_j\|_{\mathcal{H}_j}}{w_j}.$$

Now, we choose $\widetilde{\lambda} > 0$ such that $\|x_i\|_{\mathcal{H}_i} - w_i \widetilde{\lambda} < 0$ for all $i = 1, ..., n$, and get

$$g(\widetilde{\lambda}) < -\widetilde{\lambda} + \frac{\|x_j\|_{\mathcal{H}_j}}{w_j} < 0.$$

(b) Let $\xi \geq 0$. If there exists $j \in \{1, ..., n\}$ such that $\|x_j\|_{\mathcal{H}_j} + \xi w_j < 0$, we derive a contradiction. Therefore, it holds $\|x_i\|_{\mathcal{H}_i} + \xi w_i \geq 0$ for all $i = 1, ..., n$, and for the function g we have

$$g(\lambda) = \sum_{i=1}^n w_i^2 \max\{\tau_i - \lambda, 0\} - \lambda - \xi \leq \sum_{i=1}^n w_i^2 \max\{\tau_i - \lambda, 0\} - \lambda.$$

Now, we can take $\widetilde{\lambda} > 0$ such that $\|x_i\|_{\mathcal{H}_i} - w_i \widetilde{\lambda} < 0$ for all $i = 1, ..., n$, and derive that $g(\widetilde{\lambda}) \leq -\widetilde{\lambda} < 0$.

In summary, we can secure the existence of $\widetilde{\lambda} > 0$ such that $g(\widetilde{\lambda}) < 0$. Additionally, take note that, if $\lambda = 0$, then $g(0) = \sum_{i=1}^n w_i \|x_i\|_{\mathcal{H}_i} - \xi > 0$. The rest of the proof is oriented on the Algorithm I given in [18] to determine the projection onto an l_1-norm ball.

Since, the values $\tau_0, ..., \tau_n$ are in ascending order, g is a piecewise linear function in λ, where the slope of g changes at $\lambda = \tau_i$, $i = 0, ..., n$. More precisely, at $\lambda = 0$ the slope of g is $-(\sum_{i=1}^n w_i^2 + 1)$ and increases by w_1^2 when $\lambda = \tau_1$. If we proceed in this way, one may see that the slope keeps increasing when λ takes the values τ_k, $k = 2, ..., n$. In the case when $\lambda \geq \tau_n$ the slope of g is -1.

Hence, to determine λ such that $g(\lambda) = 0$, we have to locate the interval where g changes its sign from a positive to a negative value. In other words, we have to find the unique integer $k \in \{0, ..., n-1\}$ such that $g(\tau_k) \geq 0$ and $g(\tau_{k+1}) \leq 0$. Hence, we have

$$g(\lambda) = -\left(\sum_{i=k+1}^n w_i^2 + 1\right)\lambda + \sum_{i=k+1}^n w_i^2 \tau_i - \xi,$$

where $\tau_k \leq \lambda \leq \tau_{k+1}$. Finally, we can determine $\overline{\lambda}$ such that $g(\overline{\lambda}) = 0$:

$$\overline{\lambda} = \frac{\sum\limits_{i=k+1}^{n} w_i^2 \tau_i - \xi}{\sum\limits_{i=k+1}^{n} w_i^2 + 1}.$$

□

Remark 2.4. From the ideas of the previous proof, we can now construct an algorithm to determine $\overline{\lambda}$ of Corollary 2.3.

Algorithm:

1. If $\sum_{i=1}^{n} w_i \|x_i\|_{\mathcal{H}_i} \leq \xi$, then $\overline{\lambda} = 0$.
2. If $\xi < 0$ and $\|x_i\|_{\mathcal{H}_i} \leq -\xi w_i$ for all $i = 1, ..., n$, then $\overline{\lambda} = -\xi$.
3. Otherwise, define $\tau_0 := 0$, $\tau_i := \|x_i\|_{\mathcal{H}_i}/w_i$, $i = 1, ..., n$, and sort $\tau_0, ..., \tau_n$ in ascending order.
4. Determine the values of g defined in (2.19) at $\lambda = \tau_i$, $i = 0, ..., n$.
5. Find the unique $k \in \{0, ..., n-1\}$ such that $g(\tau_k) \geq 0$ and $g(\tau_{k+1}) \leq 0$.
6. Calculate $\overline{\lambda}$ by (2.15).

Corollary 2.5. *Let h be given by (2.1) where $\beta_i = 2$ and $w_i = 1$ for all $i = 1, ..., n$, then it holds*

$$\mathrm{P}_{\mathrm{epi}\, h}(x_1, ..., x_n, \xi) = \begin{cases} (x_1, ..., x_n, \xi), & \text{if } \sum\limits_{i=1}^{n} \|x_i\|_{\mathcal{H}_i}^2 \leq \xi, \\ (\overline{y}_1, ..., \overline{y}_n, \overline{\theta}), & \text{otherwise}, \end{cases}$$

where

$$\overline{y}_i = \frac{1}{2\overline{\lambda} + 1} x_i, \ \ i = 1, ..., n, \ \ \text{and } \overline{\theta} = \xi + \overline{\lambda},$$

and $\overline{\lambda} > 0$ is a solution of a cubic equation of the form

$$\lambda^3 + (1 + \xi)\lambda^2 + \frac{1}{4}(1 + 4\xi)\lambda + \frac{1}{4}\left(\xi - \sum_{i=1}^{n} \|x_i\|_{\mathcal{H}_i}^2\right) = 0. \tag{2.20}$$

Proof. By Theorem 2.1 we get that

$$\mathrm{P}_{\mathrm{epi}\, h}(x_1, ..., x_n, \xi) = \begin{cases} (x_1, ..., x_n, \xi), & \text{if } \sum\limits_{i=1}^{n} \|x_i\|_{\mathcal{H}_i}^2 \leq \xi, \\ (\overline{y}_1, ..., \overline{y}_n, \overline{\theta}), & \text{otherwise}, \end{cases}$$

with

$$\overline{y}_i = \frac{\|x_i\|_{\mathcal{H}_i} - \eta_i(\overline{\lambda})}{\|x_i\|_{\mathcal{H}_i}} x_i, \ \ i = 1, ..., n, \ \text{and } \overline{\theta} = \xi + \overline{\lambda}, \tag{2.21}$$

where $\eta_i(\overline{\lambda})$ is the unique non-negative real number that solves the equation

$$\eta_i(\overline{\lambda}) + \frac{\eta_i(\overline{\lambda})}{2\overline{\lambda}} = \|x_i\|_{\mathcal{H}_i},\ i = 1, ..., n, \tag{2.22}$$

and $\overline{\lambda} > 0$ is a solution of the equation

$$\sum_{i=1}^{n} (\|x_i\|_{\mathcal{H}_i} - \eta_i(\lambda))^2 = \lambda + \xi. \tag{2.23}$$

From (2.22) we get immediately

$$\eta_i(\overline{\lambda}) \left(1 + \frac{1}{2\overline{\lambda}}\right) = \|x_i\|_{\mathcal{H}_i} \Leftrightarrow \eta_i(\overline{\lambda}) = \frac{2\overline{\lambda}}{2\overline{\lambda}+1} \|x_i\|_{\mathcal{H}_i},\ i = 1, ..., n, \tag{2.24}$$

and in combination with (2.23) we derive

$$\begin{aligned} & \sum_{i=1}^{n} \left(\|x_i\|_{\mathcal{H}_i} - \frac{2\overline{\lambda}}{2\overline{\lambda}+1}\|x_i\|_{\mathcal{H}_i}\right)^2 = \overline{\lambda} + \xi \Leftrightarrow \frac{1}{(2\overline{\lambda}+1)^2} \sum_{i=1}^{n} \|x_i\|^2_{\mathcal{H}_i} = \overline{\lambda} + \xi \\ \Leftrightarrow \quad & (2\overline{\lambda}+1)^2(\overline{\lambda}+\xi) - \sum_{i=1}^{n} \|x_i\|^2_{\mathcal{H}_i} = 0 \\ \Leftrightarrow \quad & 4\overline{\lambda}^3 + 4(1+\xi)\overline{\lambda}^2 + (1+4\xi)\overline{\lambda} + \xi - \sum_{i=1}^{n} \|x_i\|^2_{\mathcal{H}_i} = 0. \end{aligned}$$

In the end, formula (2.24) implies that

$$\overline{y}_i = \frac{\|x_i\|_{\mathcal{H}_i} - \frac{2\overline{\lambda}}{2\overline{\lambda}+1}\|x_i\|_{\mathcal{H}_i}}{\|x_i\|_{\mathcal{H}_i}} x_i = \frac{1}{2\overline{\lambda}+1} x_i,\ i = 1, ..., n, \tag{2.25}$$

which completes the proof. □

The next remark discusses the question whether the solution $\overline{\lambda} > 0$ of Corollary 2.5 is unique.

Remark 2.6. Let $(x_1, ..., x_n, \xi) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \times \mathbb{R}$ be such that $\sum_{i=1}^n \|x_i\|^2_{\mathcal{H}_i} > \xi$ and $g : \mathbb{R} \to \mathbb{R}$ be defined by $g(\lambda) := \lambda^3 + (1+\xi)\lambda^2 + (1/4)(1+4\xi)\lambda + (1/4)(\xi - \sum_{i=1}^n \|x_i\|^2_{\mathcal{H}_i})$, then $g'(\lambda) = 3\lambda^2 + 2(1+\xi)\lambda + (1/4)(1+4\xi)$ as well as $g''(\lambda) = 6\lambda + 2(1+\xi)$. From the zeros of g' we derive the local extrema of g as follows

$$\begin{aligned} \lambda_{1/2} &= -\frac{1}{3}(1+\xi) \pm \sqrt{\frac{(1+\xi)^2}{9} - \frac{1+4\xi}{12}} = -\frac{1}{3}(1+\xi) \pm \sqrt{\frac{4(1+2\xi+\xi^2) - 3(1+4\xi)}{36}} \\ &= -\frac{1}{3}(1+\xi) \pm \sqrt{\frac{1-4\xi+4\xi^2}{36}} = -\frac{1}{3}(1+\xi) \pm \frac{1}{6}(1-2\xi) \end{aligned}$$

and hence, $\lambda_1 = -(1/6)(1+4\xi)$ and $\lambda_2 = -(1/2)$.

Further, if $\xi > 1/2 \Leftrightarrow -1 + 2\xi > 0$, then g is strongly monotone increasing on $\mathbb{R}_+$, $g''(\lambda_1) = 1 - 2\xi < 0$ and $g''(\lambda_2) = -1 + 2\xi > 0$, which means that g has in λ_1 a local maximum and in λ_2 a local minimum. As $\lambda_1 < \lambda_2 < 0$ and $g(0) = (1/4)(\xi - \sum_{i=1}^n \|x_i\|^2_{\mathcal{H}_i}) < 0$, the function g has exactly one positive zero in this situation.

If $\xi < 1/2 \Leftrightarrow 1-2\xi > 0$, then $g''(\lambda_1) = 1-2\xi > 0$ and $g''(\lambda_2) = -1+2\xi < 0$ and we derive a local minimum in λ_1 and a local maximum in λ_2. From $g(0) < 0$ and $\lambda_2 < \lambda_1$ we conclude that g has also in this situation exactly one positive zero.

Finally, let us consider the case where $\xi = 1/2$, then g is strongly monotone increasing on $\mathbb{R}_+$, $\lambda_1 = \lambda_2 = -1/2$ and $g''(\lambda_1) = 0$, i.e. g has at the point $-(1/2)$ a saddle point. From the fact that $g''(\lambda) \leq 0$ for all $\lambda \in (-\infty, -(1/2)]$ and $g''(\lambda) > 0$ for all $\lambda \in (-(1/2), +\infty)$, it is clear that g has again exactly one positive zero.

In conclusion, the function g has in all situations exactly one positive zero, i.e. $\overline{\lambda} > 0$ is unique.

Remark 2.7. In the framework of Corollary 2.5, let us consider the case where $n = 1$. Then, by Remark 2.2 we have to find a real number $\overline{\chi} \geq 0$ that solves the equation

$$2\chi^3 + (1-2\xi)\chi - \|x\|_{\mathcal{H}} = 0, \tag{2.26}$$

to get a formula of the projection onto the epigraph of h.

As one may see by (2.20), the arithmetic effort for the case $n > 1$ is not much higher compared to the case $n = 1$. In both situations we have to solve a cubic equation to derive a formula for the projection onto the epigraph of h.

As a direct consequence of Corollary 2.3 one gets the following well-known statement (see for instance [1] or [12]).

Corollary 2.8. *Let h be given by (2.1) where $n = 1$, $w_1 = w \geq 1$ and $\beta_1 = 1$, i.e. $h(x) = w\|x\|_{\mathcal{H}}$. Then, for every $(x, \xi) \in \mathcal{H} \times \mathbb{R}$*

$$\mathrm{P}_{\operatorname{epi} w\|\cdot\|_{\mathcal{H}}}(x,\xi) = \begin{cases} (x,\xi), & \text{if } w\|x\|_{\mathcal{H}} \leq \xi, \\ (0,0), & \text{if } \|x\|_{\mathcal{H}} \leq -w\xi, \\ \left(\frac{\|x\|_{\mathcal{H}} + w\xi}{\|x\|_{\mathcal{H}}(w^2+1)} x, \frac{w\|x\|_{\mathcal{H}} + w^2\xi}{w^2+1}\right), & \text{otherwise.} \end{cases}$$

For our numerical tests we need two lemmas more.

Lemma 2.9. *For $p_i \in \mathcal{H}$, $i = 1, ..., n$, it holds*

$$\mathrm{P}_{\operatorname{epi}\left(\sum_{i=1}^n w_i\|\cdot - p_i\|_{\mathcal{H}_i}^{\beta_i}\right)}(x_1, ..., x_n, \xi) = \mathrm{P}_{\operatorname{epi}\left(\sum_{i=1}^n w_i\|\cdot\|_{\mathcal{H}_i}^{\beta_i}\right)}(x_1 - p_1, ..., x_n - p_n, \xi) + (p_1, ..., p_n, 0).$$

Proof. For $p_i \in \mathcal{H}_i$, $i = 1, ..., n$ one has

$$\begin{aligned} & (x_1, ..., x_n, \xi) \in \operatorname{epi}\left(\sum_{i=1}^n w_i\|\cdot - p_i\|_{\mathcal{H}_i}^{\beta_i}\right) \Leftrightarrow \sum_{i=1}^n w_i\|x_i - p_i\|_{\mathcal{H}_i}^{\beta_i} \leq \xi \\ \Leftrightarrow \quad & (x_1 - p_1, ..., x_n - p_n, \xi) \in \operatorname{epi}\left(\sum_{i=1}^n w_i\|\cdot\|_{\mathcal{H}_i}^{\beta_i}\right) \\ \Leftrightarrow \quad & (x_1, ..., x_n, \xi) \in \operatorname{epi}\left(\sum_{i=1}^n w_i\|\cdot\|_{\mathcal{H}_i}^{\beta_i}\right) + (p_1, ..., p_n, 0). \end{aligned}$$

Thus, by [1, Proposition 3.17] follows

$$\begin{aligned} & \mathrm{P}_{\operatorname{epi}\left(\sum_{i=1}^n w_i\|\cdot - p_i\|_{\mathcal{H}_i}^{\beta_i}\right)}(x_1, ..., x_n, \xi) = \mathrm{P}_{\operatorname{epi}\left(\sum_{i=1}^n w_i\|\cdot\|_{\mathcal{H}_i}^{\beta_i}\right) + (p_1, ..., p_n, 0)}(x_1, ..., x_n, \xi) \\ = \quad & \mathrm{P}_{\operatorname{epi}\left(\sum_{i=1}^n w_i\|\cdot\|_{\mathcal{H}_i}^{\beta_i}\right)}(x_1 - p_1, ..., x_n - p_n, \xi) + (p_1, ..., p_n, 0). \end{aligned}$$

□

Lemma 2.10. *Let $w > 0$ and $A : \mathcal{K} \to \mathcal{H}$ be a linear operator with $AA^* = \mu Id$, $\mu > 0$, where $\mathcal{K}$ is a real Hilbert space. Then,*

$$\mathrm{P}_{\mathrm{epi}\, w\|A\cdot\|_{\mathcal{H}}}(x,\xi) = (x,\xi) + \left(\frac{1}{\sqrt{\mu}}A^* \times \mathrm{Id}\right)\left(\mathrm{P}_{\mathrm{epi}\, w\sqrt{\mu}\|\cdot\|_{\mathcal{H}}}\left(\frac{1}{\sqrt{\mu}}Ax,\xi\right) - \left(\frac{1}{\sqrt{\mu}}Ax,\xi\right)\right),$$

where $\frac{1}{\sqrt{\mu}}A^* \times \mathrm{Id} : \mathcal{H}\times\mathbb{R} \to \mathcal{K}\times\mathbb{R}$ is defined as $\left(\frac{1}{\sqrt{\mu}}A^* \times \mathrm{Id}\right)(y,\zeta) = \left(\frac{1}{\sqrt{\mu}}A^* y,\zeta\right)$.

Proof. We have

$$\delta_{\mathrm{epi}(w\|A\cdot\|_{\mathcal{H}})}(x,\xi) = \delta_{\mathrm{epi}(w\sqrt{\mu}\|\cdot\|_{\mathcal{H}})}\left(\frac{1}{\sqrt{\mu}}Ax,\xi\right) = \left(\delta_{\mathrm{epi}(w\sqrt{\mu}\|\cdot\|_{\mathcal{H}})} \circ \left(\frac{1}{\sqrt{\mu}}A \times \mathrm{Id}\right)\right)(x,\xi).$$

By [1, Proposition 23.32] (with $L = (1/\sqrt{\mu})A \times \mathrm{Id}$) it follows that

$$\begin{aligned}
&\mathrm{prox}_{\delta_{\mathrm{epi}\, w\|A\cdot\|_{\mathcal{H}}}}(x,\xi) = \mathrm{prox}_{\delta_{\mathrm{epi}(w\sqrt{\mu}\|\cdot\|_{\mathcal{H}})}\circ\left(\frac{1}{\sqrt{\mu}}A\times\mathrm{Id}\right)}(x,\xi)\\
=&(x,\xi) + \left(\frac{1}{\sqrt{\mu}}A \times \mathrm{Id}\right)^*\left(\mathrm{prox}_{\delta_{\mathrm{epi}\, w\sqrt{\mu}\|\cdot\|_{\mathcal{H}}}}\left(\frac{1}{\sqrt{\mu}}Ax,\xi\right) - \left(\frac{1}{\sqrt{\mu}}Ax,\xi\right)\right)\\
\Leftrightarrow& \mathrm{P}_{\mathrm{epi}\, w\|A\cdot\|_{\mathcal{H}}}(x,\xi) = (x,\xi) + \left(\frac{1}{\sqrt{\mu}}A^* \times \mathrm{Id}\right)\left(\mathrm{P}_{\mathrm{epi}\, w\sqrt{\mu}\|\cdot\|_{\mathcal{H}}}\left(\frac{1}{\sqrt{\mu}}Ax,\xi\right) - \left(\frac{1}{\sqrt{\mu}}Ax,\xi\right)\right).
\end{aligned}$$

□

2.2 Gauges

The next considerations are devoted to gauge functions (a.k.a. *Minkowski functional*) of a closed, convex and non-empty subset $C \subseteq \mathcal{H}$, $\gamma_C : \mathcal{H} \to \overline{\mathbb{R}}$ defined by

$$\gamma_C(x) := \begin{cases} \inf\{\lambda > 0 : x \in \lambda C\}, & \text{if } \{\lambda > 0 : x \in \lambda C\} \neq \emptyset,\\ +\infty, & \text{otherwise.}\end{cases}$$

Theorem 2.11. *Let C be a closed convex subset of $\mathcal{H}$ such that $0_{\mathcal{H}} \in C$, then it holds for every $(x,\xi) \in \mathcal{H}\times\mathbb{R}$*

$$\mathrm{P}_{\mathrm{epi}\,\gamma_C}(x,\xi) = \begin{cases} (x,\xi), & \text{if } \gamma_C(x) \leq \xi,\\ \left(P_{\mathrm{cl}(\mathrm{dom}\,\gamma_C)}(x),\xi\right), & \text{if } x \notin \mathrm{dom}\,\gamma_C \text{ and } \gamma_C\left(P_{\mathrm{cl}(\mathrm{dom}\,\gamma_C)}(x)\right) \leq \xi < \gamma_C(x),\\ (\overline{y},\overline{\theta}), & \text{otherwise,}\end{cases}$$

where

$$\overline{y} = x - \overline{\lambda}\,\mathrm{P}_{C^0}\left(\frac{1}{\overline{\lambda}}x\right) \text{ and } \overline{\theta} = \overline{\lambda} + \xi$$

and $\overline{\lambda} > 0$ is a solution of an equation of the form

$$\lambda + \xi = \left\langle x, \mathrm{P}_{C^0}\left(\frac{1}{\lambda}x\right)\right\rangle_{\mathcal{H}} - \lambda\left\|\mathrm{P}_{C^0}\left(\frac{1}{\lambda}x\right)\right\|^2_{\mathcal{H}}.$$

Proof. Let us consider for fixed $(x,\xi) \in \mathcal{H} \times \mathbb{R}$ the following optimization problem

$$\min_{\substack{(y,\theta)\in\mathcal{H}\times\mathbb{R},\\ \gamma_C(y)\leq\theta}} \left\{ \frac{1}{2}(\theta-\xi)^2 + \frac{1}{2}\|y-x\|_{\mathcal{H}}^2 \right\}. \tag{2.27}$$

If $\gamma_C(x) \leq \xi$, i.e. $(x,\xi) \in \operatorname{epi}\gamma_C$, then it is obvious that $(\overline{y},\overline{\theta}) = (x,\xi)$. In the following we consider the non-trivial situation where $\gamma_C(x) > \xi$.

We define the function $f : \mathcal{H} \times \mathbb{R} \to \mathbb{R}$ by $f(y,\theta) := (1/2)(\theta-\xi)^2 + (1/2)\|y-x\|_{\mathcal{H}}^2$ and the function $g : \mathcal{H} \times \mathbb{R} \to \overline{\mathbb{R}}$ by $g(y,\theta) = \gamma_C(y) - \theta$, then it is clear that f is continuous and strongly convex and g is proper, lower semicontinuous and convex by [22, Theorem 1]. As $\gamma_C(0) < 1$, it follows by [5, Theorem 3.3.16] (see also [5, Remark 3.3.8]) that

$$0 \in \partial(f + (\overline{\lambda}g))(\overline{y},\overline{\theta}) \tag{2.28}$$

and

$$\begin{cases} (\overline{\lambda}g)(\overline{y},\overline{\theta}) = 0, \\ g(\overline{y},\overline{\theta}) \leq 0, \end{cases} \Leftrightarrow \begin{cases} \overline{\lambda}(\gamma_C(\overline{y}) - \overline{\theta}) = 0, \\ \gamma_C(\overline{y}) \leq \overline{\theta}. \end{cases} \tag{2.29}$$

where $(\overline{y},\overline{\theta})$ is the unique solution of (2.27) and $\overline{\lambda} \geq 0$ the associated Lagrange multiplier. Furthermore, from [5, Theorem 3.5.13] one gets that

$$0 \in \partial(f + (\overline{\lambda}g))(\overline{y},\overline{\theta}) \Leftrightarrow 0 \in \partial f(\overline{y},\overline{\theta}) + \partial(\overline{\lambda}g)(\overline{y},\overline{\theta}). \tag{2.30}$$

If $\overline{\lambda} = 0$, then it follows by (1.9) and (1.3)

$$\begin{aligned} & 0 \in \partial f(\overline{y},\overline{\theta}) + \partial\delta_{\operatorname{dom} g}(\overline{y},\overline{\theta}) \Leftrightarrow 0 \in (\overline{y}-x,\overline{\theta}-\xi) + \partial\delta_{\operatorname{dom}\gamma_C\times\mathbb{R}}(\overline{y},\overline{\theta}) \\ \Leftrightarrow\quad & 0 \in (\overline{y}-x,\overline{\theta}-\xi) + \partial\delta_{\operatorname{cl}(\operatorname{dom}\gamma_C)\times\mathbb{R}}(\overline{y},\overline{\theta}) \Leftrightarrow (x-\overline{y},\xi-\overline{\theta}) \in \partial\delta_{\operatorname{cl}(\operatorname{dom}\gamma_C)\times\mathbb{R}}(\overline{y},\overline{\theta}) \\ \Leftrightarrow\quad & (\overline{y},\overline{\theta}) = \mathrm{P}_{\operatorname{cl}(\operatorname{dom}\gamma_C)\times\mathbb{R}}(x,\xi) \Leftrightarrow \begin{cases} \overline{y} = \mathrm{P}_{\operatorname{cl}(\operatorname{dom}\gamma_C)}(x), \\ \overline{\theta} = \xi, \end{cases} \end{aligned}$$

and thus, it holds by the feasibility condition (2.29) that $\gamma_C(\mathrm{P}_{\operatorname{cl}(\operatorname{dom}\gamma_C)}(x)) \leq \xi$, from which follows that $\mathrm{P}_{\operatorname{cl}(\operatorname{dom}\gamma_C)}(x) \in \operatorname{dom}\gamma_C$. If $x \in \operatorname{dom}\gamma_C$, this means that $\mathrm{P}_{\operatorname{cl}(\operatorname{dom}\gamma_C)}(x) = x$ and again by the feasibility condition (2.29) that $\gamma_C(x) \leq \xi$, which contradicts our assumption. Therefore, if $x \notin \operatorname{dom}\gamma_C$ and the inequalities $\gamma_C(\mathrm{P}_{\operatorname{cl}(\operatorname{dom}\gamma_C)(x)}) \leq \xi < \gamma_C(x)$ hold, then $(\overline{y},\overline{\theta}) = \big(P_{\operatorname{cl}(\operatorname{dom}\gamma_C)}(x),\xi\big)$.

Now, let $\overline{\lambda} > 0$, then it follows from (2.30) and (1.9)

$$\begin{aligned} & 0 \in \partial(f + (\overline{\lambda}g))(\overline{y},\overline{\theta}) \Leftrightarrow 0 \in \partial f(\overline{y},\overline{\theta}) + \overline{\lambda}\partial g(\overline{y},\overline{\theta}) \\ \Leftrightarrow\quad & \nabla f(\overline{y},\overline{\theta}) \in -\overline{\lambda}\partial g(\overline{y},\overline{\theta}) \Leftrightarrow \begin{cases} \overline{y}-x \in -\overline{\lambda}\partial\gamma_C(\overline{y}), \\ \overline{\theta}-\xi = \overline{\lambda} \end{cases} \Leftrightarrow \begin{cases} \overline{y} = \operatorname{prox}_{\overline{\lambda}\gamma_C} x, \\ \overline{\theta} = \xi - \overline{\lambda}, \end{cases} \end{aligned} \tag{2.31}$$

by combining (2.31) and (2.29) we derive that $\gamma_C(\overline{y}) = \xi + \overline{\lambda}$. Finally, as by [22, Lemma 1] and [22, Remark 3] it holds that $\gamma_C^* = \delta_{C^0}$, one gets by [1, Theorem 14.3(iii)] the following equivalences

$$\begin{aligned} & \gamma_C(\overline{y}) = \xi + \overline{\lambda} \qquad\qquad (2.32) \\ \Leftrightarrow\quad & \xi + \overline{\lambda} = \gamma_C\left(\operatorname{prox}_{\overline{\lambda}\gamma_C} x\right) + \delta_{C^0}\left(\mathrm{P}_{C^0}\left(\frac{1}{\overline{\lambda}}x\right)\right) = \left\langle \operatorname{prox}_{\overline{\lambda}\gamma_C} x, \mathrm{P}_{C^0}\left(\frac{1}{\overline{\lambda}}x\right)\right\rangle_{\mathcal{H}} \end{aligned}$$

$$\Leftrightarrow \quad \xi + \overline{\lambda} = \left\langle x - \overline{\lambda} \operatorname{P}_{C^0}\left(\frac{1}{\overline{\lambda}}x\right), \operatorname{P}_{C^0}\left(\frac{1}{\overline{\lambda}}x\right)\right\rangle_{\mathcal{H}}.$$

□

Corollary 2.12. *Let $C \subseteq \mathcal{H}$ be a closed convex cone, then $\gamma_C = \delta_C$ and*

$$\operatorname{P}_{\operatorname{epi}\gamma_C}(x,\xi) = \operatorname{P}_{C\times\mathbb{R}_+}(x,\xi) = \begin{cases} (x,\xi), & \text{if } x \in C \text{ and } \xi \geq 0, \\ (\operatorname{P}_C x, \max\{0,\xi\}), & \text{otherwise.} \end{cases}$$

Proof. We use Theorem 2.11. Let $x \in \operatorname{dom}\gamma_C$ such that $\gamma_C(x) > \xi$, then one has from [1, Proposition 28.22] and [1, Theorem 6.29] that

$$\overline{y} = x - \overline{\lambda}\operatorname{P}_{C^0}\left(\frac{1}{\overline{\lambda}}x\right) = x - \operatorname{P}_{C^0} x = \operatorname{P}_C x. \tag{2.33}$$

Moreover, as $\gamma_C = \delta_C$ it holds that $\operatorname{dom}\gamma_C = C$ and by (2.29) we have

$$\gamma_C(\overline{y}) = \overline{\theta},$$

which yields

$$\operatorname{P}_{\operatorname{epi}\gamma_C}(x,\xi) = (\operatorname{P}_C x, \gamma_C(\operatorname{P}_C x)) = (\operatorname{P}_C x, 0).$$

If $x \notin \operatorname{dom}\gamma_C = C$, then

$$0 = \gamma_C\left(\operatorname{P}_{\operatorname{cl}(\operatorname{dom}\gamma_C)}(x)\right) \leq \xi < \gamma_C(x) = +\infty$$

and so, $\operatorname{P}_{\operatorname{epi}\gamma_C}(x,\xi) = (\operatorname{P}_{\operatorname{cl}(\operatorname{dom}\gamma_C)}(x),\xi) = (\operatorname{P}_C x, \xi)$, which implies the statement. □

Corollary 2.13. *Let C_i be a closed convex subset of $\mathcal{H}_i$ such that $0_{\mathcal{H}_i} \in \operatorname{int} C_i$, $i = 1,...,n$, and the gauge $\gamma_C : \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \to \overline{\mathbb{R}}$ be defined by $\gamma_C(x_1,...,x_n) = \sum_{i=1}^n \gamma_{C_i}(x_i)$. Then it holds for every $(x_1,...,x_n,\xi) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \times \mathbb{R}$*

$$\operatorname{P}_{\operatorname{epi}\gamma_C}(x_1,...,x_n,\xi) = \begin{cases} (x_1,...,x_n,\xi), & \text{if } \sum\limits_{i=1}^n \gamma_{C_i}(x_i) \leq \xi, \\ (\overline{y}_1,...,\overline{y}_n,\overline{\theta}), & \text{otherwise,} \end{cases}$$

where

$$\overline{y}_i = x_i - \overline{\lambda}\operatorname{P}_{C_i^0}\left(\frac{1}{\overline{\lambda}}x_i\right), \ i = 1,...,n, \ \text{and } \overline{\theta} = \overline{\lambda} + \xi \tag{2.34}$$

and $\overline{\lambda} > 0$ is a solution of an equation of the form

$$\lambda + \xi = \sum_{i=1}^n \left[\left\langle x_i, \operatorname{P}_{C_i^0}\left(\frac{1}{\lambda}x_i\right)\right\rangle_{\mathcal{H}_i} - \lambda\left\|\operatorname{P}_{C_i^0}\left(\frac{1}{\lambda}x_i\right)\right\|^2_{\mathcal{H}_i}\right]. \tag{2.35}$$

Proof. As $0_{\mathcal{H}_i} \in \operatorname{int} C_i$, $i = 1,...,n$, it is clear that the gauges are well-defined, i.e. $\operatorname{dom}\gamma_{C_i} = \mathcal{H}_i$, $i = 1,...,n$, and so, $\operatorname{dom}\gamma_C = \mathcal{H}_1 \times \cdots \times \mathcal{H}_n$. Further, let us recall that the polar set C^0 of the set C can be characterized by the dual gauge γ_{C^0} as

$$C^0 = \{x = (x_1,...,x_n) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n : \gamma_{C^0}(x) = \gamma_{C^0}(x_1,...,x_n) \leq 1\}. \tag{2.36}$$

This relation holds also for the polar set C_i^0 and its associated dual gauge $\gamma_{C_i^0}$, $i = 1, ..., n$. Moreover, in [21] it was shown that $\gamma_{C^0}(x) = \max_{1 \leq i \leq n}\{\gamma_{C_i^0}(x_i)\}$ and hence, the polar set in (2.36) can be written as

$$
\begin{aligned}
C^0 &= \left\{(x_1, ..., x_n) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n : \max_{1 \leq i \leq n}\{\gamma_{C_i^0}(x_i)\} \leq 1\right\} \\
&= \left\{(x_1, ..., x_n) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n : \gamma_{C_i^0}(x_i) \leq 1,\ i = 1, ..., n\right\} \\
&= \{x_1 \in \mathcal{H}_1 : \gamma_{C_1^0}(x_1) \leq 1\} \times \cdots \times \{x_n \in \mathcal{H}_n : \gamma_{C_n^0}(x_n) \leq 1\} = C_1^0 \times \cdots \times C_n^0.
\end{aligned}
$$

From here follows that

$$
\mathrm{P}_{C^0}(x) = \mathrm{P}_{C_1^0 \times \cdots \times C_n^0}(x_1, ..., x_n) = \mathrm{P}_{C_1^0}(x_1) \times \cdots \times \mathrm{P}_{C_n^0}(x_n),
$$

which by using Theorem 2.11 directly implies (2.34) and (2.35). □

Remark 2.14. Like in Lemma 2.10, one can give a formula for the projection onto the epigraph of a gauge composed with a linear operator $A : \mathcal{K} \to \mathcal{H}$ with $AA^* = \mu Id$, $\mu > 0$,

$$
\mathrm{P}_{\mathrm{epi}\,\gamma_C(A\cdot)}(x, \xi) = (x, \xi) + \left(\frac{1}{\sqrt{\mu}}A^* \times \mathrm{Id}\right)\left(\mathrm{P}_{\mathrm{epi}\,\sqrt{\mu}\gamma_C(\cdot)}\left(\frac{1}{\sqrt{\mu}}Ax, \xi\right) - \left(\frac{1}{\sqrt{\mu}}Ax, \xi\right)\right).
$$

Moreover, it can easily be observed that for $p \in \mathcal{H}$ holds (similar to the proof of Lemma 2.9)

$$
\mathrm{P}_{\mathrm{epi}\,\gamma_C(\cdot - p)}(x, \xi) = \mathrm{P}_{\mathrm{epi}\,\gamma_C}(x - p, \xi) + (p, 0).
$$

We close this section with a characterization of the subdifferential of a gauge function by the projection operator.

Remark 2.15. Let $C \subseteq \mathcal{H}$ be closed and convex such that $0_{\mathcal{H}} \in C$, then it holds by (1.3), (1.9), [22, Lemma 1], [22, Remark 3] and [1, Theorem 14.3(ii)] for all $x, y \in \mathcal{H}$ that

$$
\begin{aligned}
& x \in \partial\gamma_C(y) \Leftrightarrow x + y - y \in \partial\gamma_C(y) \Leftrightarrow y = \mathrm{prox}_{\gamma_C}(x + y) \\
\Leftrightarrow\ & y = x + y - \mathrm{prox}_{\gamma_C^*}(x + y) \Leftrightarrow y = x + y - \mathrm{prox}_{\delta_{C^0}}(x + y) \\
\Leftrightarrow\ & x = \mathrm{P}_{C^0}(x + y).
\end{aligned}
$$

From this follows that

$$
\partial\gamma_C(y) = \{x \in \mathcal{H} : x = \mathrm{P}_{C^0}(x + y)\}.
$$

In addition, if C is a closed convex cone, then it follows from [1, Theorem 6.29] that

$$
\partial\gamma_C(y) = \{x \in \mathcal{H} : x = x + y - \mathrm{P}_C(x + y)\} = \{x \in \mathcal{H} : y = \mathrm{P}_C(x + y)\}.
$$

3 Numerical Experiments

Our numerical tests are implemented in MATLAB on a PC with an Intel Core i5-8400 CPU with 2.8GHz and 16 GB RAM. While the numerical tests in [14] were based on the partial inverse algorithm introduced by Spingarn in [19], we use here the *parallel splitting algorithm* from [1, Proposition 27.8].

Theorem 3.1 (parallel splitting algorithm)**.** *Let n be an integer such that $n \geq 2$ and $f_i : \mathbb{R}^s \to \overline{\mathbb{R}}$ be a proper, lower semicontinuous and convex function for $i = 1, ..., n$. Suppose that the problem*

$$(P^{DR}) \quad \min_{x \in \mathbb{R}^s} \left\{ \sum_{i=1}^{n} f_i(x) \right\}$$

has at least one solution and that $\operatorname{dom} f_1 \cap \bigcap_{i=2}^{n} \operatorname{int} \operatorname{dom} f_i \neq \emptyset$*. Let $(\mu_k)_{k\in\mathbb{N}}$ be a sequence in $[0, 2]$ such that $\sum_{k\in\mathbb{N}} \mu_k(2 - \mu_k) = +\infty$, let $\nu > 0$, and let $(x_{i,0})_{i=1}^{n} \in \mathbb{R}^s \times \cdots \times \mathbb{R}^s$. Set*

$$(\forall k \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} r_k = \frac{1}{n} \sum\limits_{i=1}^{n} x_{i,k}, \\ y_{i,k} = \operatorname{prox}_{\nu f_i} x_{i,k}, \ i = 1, ..., n, \\ q_k = \frac{1}{n} \sum\limits_{i=1}^{n} y_{i,k}, \\ x_{i,k+1} = x_{i,k} + \mu_k(2q_k - r_k - y_{i,k}), \ i = 1, ..., n. \end{array} \right.$$

Then $(r_k)_{k\in\mathbb{N}}$ converges to a solution of problem (P^{DR}).

In order to use the parallel splitting algorithm given in the previous theorem, we need to rewrite the *extended multifacility location problem* $(EP_N^{M,\beta})$ in (1.5) into an optimization problem with an objective function, which is a sum of proper, convex and lower semicontinuous functions.

The first way to reformulate this location problem is based on the introduction of an additional variable as presented in (1.6):

$$(EP_N^{M,\beta}) \min_{(x_1,...,x_m,t)\in\mathbb{R}^d\times\cdots\times\mathbb{R}^d\times\mathbb{R}} \left\{ t + \sum_{i=1}^{n} \delta_{\operatorname{epi}\left(\sum\limits_{j=1}^{m} w_{ij}\|\cdot - p_i\|^{\beta_i}\right)}(x_1, ..., x_m, t) \right\}. \tag{3.1}$$

We define the functions

$$f_1 : \mathbb{R}^d \times \cdots \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}, \ f_1(x_1, ..., x_m, t) = t \text{ and}$$
$$f_i : \mathbb{R}^d \times \cdots \times \mathbb{R}^d \times \mathbb{R} \to \overline{\mathbb{R}}, \ f_i(x_1, ..., x_m, t) = \delta_{\operatorname{epi}\left(\sum\limits_{j=1}^{m} w_{ij}\|\cdot - p_i\|^{\beta_i}\right)}(x_1, ..., x_m, t),$$

$i = 2, ..., n + 1$, then $\operatorname{dom} f_1 = \mathbb{R}^d \times \cdots \times \mathbb{R}^d \times \mathbb{R}$ and

$$\left(0_{\mathbb{R}^d}, ..., 0_{\mathbb{R}^d}, \max_{1\leq i\leq n} \left\{ \sum_{j=1}^{m} w_{ij}\|p_i\|^{\beta_i} \right\} + 1 \right) \in \operatorname{int} \operatorname{dom} f_i = \operatorname{int} \operatorname{epi} \left(\sum_{j=1}^{m} w_{ij}\| \cdot - p_i\|^{\beta_i} \right)$$

for all $i = 2, ..., n + 1$, i.e., it holds that $\operatorname{dom} f_1 \cap \bigcap_{i=2}^{n+1} \operatorname{int} \operatorname{dom} f_i \neq \emptyset$. Therefore, the sequences generated by the algorithm from Theorem 3.1 converges to a solution of the location problem $(EP_N^{M,\beta})$ and the following formulae for the proximal points associated to the functions $f_1, ..., f_{n+1}$ can be formulated by using (1.9) and Lemma 2.9

$$\begin{aligned} & (\overline{y}_1, ..., \overline{y}_m, \overline{\theta}) = \operatorname{prox}_{\nu f_1}(x_1, ..., x_m, t) \\ \Leftrightarrow \quad & (x_1, ..., x_m, t) - (\overline{y}_1, ..., \overline{y}_m, \overline{\theta}) \in \partial(\nu f_1)(\overline{y}_1, ..., \overline{y}_m, \overline{\theta}) = (0_{\mathbb{R}^d}, ..., 0_{\mathbb{R}^d}, \nu) \\ \Leftrightarrow \quad & x_i = \overline{y}_i, \ i = 1, ..., m, \text{ and } \overline{\theta} = t - \nu \Leftrightarrow (\overline{y}_1, ..., \overline{y}_m, \overline{\theta}) = (x_1, ..., x_n, t - \nu) \end{aligned}$$

and

$$
\begin{aligned}
(\overline{y}_1,...,\overline{y}_m,\overline{\theta}) &= \operatorname{prox}_{\nu f_i}(x_1,...,x_m,t) = \operatorname{prox}_{\nu\delta_{\operatorname{epi}\left(\sum\limits_{j=1}^{m} w_{ij}\|\cdot - p_i\|^{\beta_i}\right)}}(x_1,...,x_m,t) \\
&= \mathrm{P}_{\operatorname{epi}\left(\sum\limits_{j=1}^{m} w_{ij}\|\cdot - p_i\|^{\beta_i}\right)}(x_1,...,x_m,t) \\
&= \mathrm{P}_{\operatorname{epi}\left(\sum\limits_{j=1}^{m} w_{ij}\|\cdot\|^{\beta_i}\right)}(x_1-p_i,...,x_m-p_i,t) + (p_i,...,p_i,0). \qquad (3.2)
\end{aligned}
$$

The second way to rewrite the extended multifacility location problem $(EP_N^{M,\beta})$ into an optimization problem of the form of (P^{DR}) makes use of the ideas of Cornejo and Michelot given in [14] and splits the sums of weighted norms by $n \cdot m$ additional variables (see also (1.7)):

$$
(EP_N^{M,\beta}) \quad \min_{\substack{t,\ t_{ij}\in\mathbb{R},\ x_j\in\mathbb{R}^d,\\ j=1,...,m,i=1,...,n}} \left\{ t + \sum_{j=1}^{m}\sum_{i=1}^{n} \delta_{\operatorname{epi}\left(w_{ij}\|\cdot - p_i\|^{\beta_i}\right)}(x_j,t_{ij}) + \sum_{i=1}^{n} \delta_{\operatorname{epi}\tau_i}(t_{i1},...,t_{im},t) \right\}, \qquad (3.3)
$$

where $\tau_i(t_{i1},...,t_{im}) := \sum_{j=1}^{m} t_{ij},\ i=1,...,n$. Now, let

$$
\widetilde{x} := (x_1,...,x_m) \in \mathbb{R}^d \times \cdots \times \mathbb{R}^d,\ \widetilde{t} := (t_{ij})_{i=1,...,n,j=1,...,m},
$$

$$
\begin{aligned}
&f_1 : \underbrace{\mathbb{R}^d \times \cdots \times \mathbb{R}^d}_{m-\text{times}} \times \mathbb{R}^{mn} \times \mathbb{R} \to \mathbb{R},\ f_1(\widetilde{x},\widetilde{t},t) := t, \\
&f_{ij} : \mathbb{R}^d \times \cdots \times \mathbb{R}^d \times \mathbb{R}^{mn} \times \mathbb{R} \to \overline{\mathbb{R}},\ f_{ij}(\widetilde{x},\widetilde{t},t) := \delta_{\operatorname{epi}\left(w_{ij}\|\cdot - p_i\|^{\beta_i}\right)}(x_j,t_{ij}),
\end{aligned}
$$

$j=1,...,m,\ i=1,...,n$, and

$$
\widetilde{f}_i : \mathbb{R}^d \times \cdots \times \mathbb{R}^d \times \mathbb{R}^{mn} \times \mathbb{R} \to \overline{\mathbb{R}},\ \widetilde{f}_i(\widetilde{x},\widetilde{t},t) := \delta_{\operatorname{epi}\tau_i}(t_{i1},...,t_{im},t),\ i=1,...,n.
$$

As

$$
\begin{aligned}
&\operatorname{dom} f_1 = \mathbb{R}^d \times \cdots \times \mathbb{R}^d \times \mathbb{R}^{mn} \times \mathbb{R}, \\
&\operatorname{dom} f_{ij} = \left\{ (\widetilde{x},\widetilde{t},t) \in \mathbb{R}^d \times \cdots \times \mathbb{R}^d \times \mathbb{R}^{mn} \times \mathbb{R} : (x_j,t_{ij}) \in \operatorname{epi}(w_{ij}\|\cdot - p_i\|^{\beta_i}) \right\}, \\
&i=1,...,n,\ j=1,...,m, \\
&\operatorname{dom} \widetilde{f}_i = \left\{ (\widetilde{x},\widetilde{t},t) \in \mathbb{R}^d \times \cdots \times \mathbb{R}^d \times \mathbb{R}^{mn} \times \mathbb{R} : (t_{i1},...,t_{im},t) \in \operatorname{epi}\tau_i \right\}, \\
&i=1,...,n
\end{aligned}
$$

and

$$
\left(0_{\mathbb{R}^d},...,0_{\mathbb{R}^d}, \max_{\substack{1\le i\le n,\\ 1\le j\le m}} \{w_{ij}\|p_i\|^{\beta_i}\}+1,..., \max_{\substack{1\le i\le n,\\ 1\le j\le m}} \{w_{ij}\|p_i\|^{\beta_i}\}+1, m \max_{\substack{1\le i\le n,\\ 1\le j\le m}} \{w_{ij}\|p_i\|^{\beta_i}\}+m+1\right)
$$

$$
\in \operatorname{dom} f_1 \cap \left(\bigcap_{\substack{1\le i\le n,\\ 1\le j\le m}} \operatorname{int}\operatorname{dom} f_{ij} \right) \cap \left(\bigcap_{1\le i\le n} \operatorname{int}\operatorname{dom} \widetilde{f}_i \right),
$$

convergence in the sense of Theorem 3.1 can be guaranteed. Now, let $\overline{\overline{y}} := (\overline{y}_1, ..., \overline{y}_m)$ and

$$\overline{\overline{\theta}} := (\overline{\theta}_{ij})_{1\leq i\leq n,\ 1\leq j\leq m},$$

then one has by (1.9) for the corresponding proximal points of the functions $f_1,\ f_{ij},\ j = 1, ..., m,\ i = 1, ..., n$, and $\widetilde{f}_i,\ i = 1, ..., n$,

$$(\overline{\overline{y}}, \overline{\overline{\theta}}, \overline{\theta}) = \operatorname{prox}_{\nu f_1}(\widetilde{x}, \widetilde{t}, t) = (\underbrace{0_{\mathbb{R}^d}, ..., 0_{\mathbb{R}^d}}_{m-\text{times}}, \underbrace{0, ..., 0}_{mn-\text{times}}, t - \nu)$$

and by (1.9) and Lemma 2.9

$$\begin{aligned}
&(\overline{\overline{y}}, \overline{\overline{\theta}}, \overline{\theta}) = \operatorname{prox}_{\nu f_{ij}}(\widetilde{x}, \widetilde{t}, t) \Leftrightarrow (\widetilde{x}, \widetilde{t}, t) - (\overline{\overline{y}}, \overline{\overline{\theta}}, \overline{\theta}) \in \partial(\nu f_{ij})(\overline{\overline{y}}, \overline{\overline{\theta}}, \overline{\theta})\\
\Leftrightarrow\ &(x_j, t_{ij}) - (\overline{y}_j, \overline{\theta}_{ij}) \in \partial(\nu\delta_{\operatorname{epi}(w_{ij}\|\cdot - p_i\|^{\beta_i})})(\overline{y}_j, \overline{\theta}_{ij}) \text{ and}\\
&\overline{y}_l = x_l,\ \overline{\theta}_{sl} = t_{sl},\ \overline{\theta} = t,\ s = 1, ..., n,\ l = 1, ..., m,\ sl \neq ij,\\
\Leftrightarrow\ &(\overline{y}_j, \overline{\theta}_{ij}) = \operatorname{prox}_{\nu\delta_{\operatorname{epi}(w_{ij}\|\cdot - p_i\|^{\beta_i})}}(x_j, t_{ij}) = \mathrm{P}_{\operatorname{epi}(w_{ij}\|\cdot - p_i\|^{\beta_i})}(x_j, t_{ij})\\
&= \mathrm{P}_{\operatorname{epi}(w_{ij}\|\cdot\|^{\beta_i})}(x_j - p_i, t_{ij}) + (p_i, 0) \text{ and}\\
&\overline{y}_l = x_l,\ \overline{\theta}_{sl} = t_{sl},\ \overline{\theta} = t,\ s = 1, ..., n,\ l = 1, ..., m,\ sl \neq ij,
\end{aligned} \tag{3.4}$$

$j = 1, ..., m,\ i = 1, ..., n$. Moreover, by (1.9) and [1, Example 28.17] follows

$$\begin{aligned}
&(\overline{\overline{y}}, \overline{\overline{\theta}}, \overline{\theta}) = \operatorname{prox}_{\nu \widetilde{f}_i}(\widetilde{x}, \widetilde{t}, t) \Leftrightarrow (\widetilde{x}, \widetilde{t}, t) - (\overline{\overline{y}}, \overline{\overline{\theta}}, \overline{\theta}) \in \partial(\nu \widetilde{f}_i)(\overline{\overline{y}}, \overline{\overline{\theta}}, \overline{\theta})\\
\Leftrightarrow\ &(t_{i1}, ..., t_{im}, t) - (\overline{\theta}_{i1}, ..., \overline{\theta}_{im}, \overline{\theta}) \in \partial\left(\nu\delta_{\operatorname{epi}\tau_i}\right)(\overline{\theta}_{i1}, ..., \overline{\theta}_{im}, \overline{\theta}) \text{ and}\\
&(t_{l1}, ..., t_{lm}, t) = (\overline{\theta}_{l1}, ..., \overline{\theta}_{lm}, \overline{\theta}),\ l = 1, ..., n,\ l \neq i,\ (x_1, ..., x_m) = (\overline{y}_1, ..., \overline{y}_m)\\
\Leftrightarrow\ &(\overline{\theta}_{i1}, ..., \overline{\theta}_{im}, \overline{\theta}) = \operatorname{prox}_{\nu\delta_{\operatorname{epi}\tau_i}}(t_{i1}, ..., t_{im}, t) = \mathrm{P}_{\operatorname{epi}\tau_i}(t_{i1}, ..., t_{im}, t)\\
&= \begin{cases}
(\overline{\theta}_{i1}, ..., \overline{\theta}_{im}, \overline{\theta})^T, & \text{if } \sum\limits_{j=1}^{m} t_{ij} - t \leq 0,\\
(\overline{\theta}_{i1}, ..., \overline{\theta}_{im}, \overline{\theta})^T - \frac{\sum_{j=1}^{m} t_{ij} - t}{m+1}(1, ..., 1, -1)^T, & \text{if } \sum\limits_{j=1}^{m} t_{ij} - t > 0,
\end{cases}\\
&\text{and } (t_{l1}, ..., t_{lm}, t) = (\overline{\theta}_{l1}, ..., \overline{\theta}_{lm}, \overline{\theta}),\ l = 1, ..., n,\ l \neq i,\ (x_1, ..., x_m) = (\overline{y}_1, ..., \overline{y}_m),\\
&i = 1, ..., n.
\end{aligned}$$

The tables below illustrate the performance of our method using the formulae from Corollary 2.3 and 2.5 for the projection onto the epigraph of the sum of powers of weighted norms (EpiSumNorms) compared with the concept proposed by Cornejo and Michelot in [14], where only the projection onto the epigraph of a weighted norm (EpiNorm) is needed (see Corollary 2.8). We solved the problem $(EP_N^{M,\beta})$ in $\mathbb{R}^2$ and $\mathbb{R}^3$ for different choices of given and new facilities. The performance results are visualized by the associated figures, where we use the following notations:

NumGivFac:	Number of given facilities
NumNewFac:	Number of new facilities
NumIt:	Number of iterations of the algorithm
MaxNumIt:	Maximal number of iterations
CPUtime:	CPU time in seconds.

We used the following parameters for initialization: $\mu_n = 1$ for all $n \in \mathbb{N}$. Moreover, let us point out that we tested the algorithm of Theorem 3.1 for different values of the parameter ν, where some results are printed in the tables below and selected ones are visualized in the corresponding figures (in the tables the best results of the methods EpiSumNorms and EpiNorm concerning the CPU time and number of iterations are marked in bold, respectively). Notice also that in the context of the problem (3.1) the iterate r_k of Theorem 3.1 is of the form $r_k = (x_1, ..., x_m, t)$ and in the framework of (3.3) of the form $r_k = (x_1, ..., x_m, \widetilde{t}, t)$ with $\widetilde{t} = (t_{ij})_{i=1,...,n,\ j=1,...,m}$, where $x_1, ..., x_m$ converge to the optimal locations and t to the optimal objective value.
To be more precise, for our numerical experiments we proceed as follows. The points $p_1, ..., p_n$ were generated by the MATLAB command RANDN and the corresponding weights $a_1, ..., a_n$ by RAND. Further, we used in all numerical tests as starting point the origin (i.e. $x_1 = \ldots = x_m = 0_{\mathbb{R}^d}$) and ran the algorithm of the method EpiSumNorms for all examples five hundred thousand iterations. Then, we saved the determined solutions as the optimal solutions $\overline{x} = (\overline{x}_1, ..., \overline{x}_m)$ to the associated optimization problems and set the maximal number of iterations to one hundred thousand (i.e. MaxNumIt=100000). Finally, we ran for all examples the algorithm of the method EpiSumNorms a second time as well as the algorithm of the method EpiNorm and noticed the number of iterations as well as the time needed to generate a solution which is within the maximum bound from the optimal location(s) $\overline{x} = (\overline{x}_1, ..., \overline{x}_m)$ of 0.001, i.e. $\|\overline{x} - x\| \leq 0.001$, respectively.

First, we consider the situation where $\beta_i = 1$ for all $i = 1, ..., n$.

Table 1: Performance evaluation for NumGiFac 25 and NumNewFac 5 in $\mathbb{R}^2$

MaxNumIt	EpiSumNorms		EpiNorm	
= 100000	NumIt	CPUtime	NumIt	CPUtime
$\nu = 0.1$	50879	83.8313	>100000	-
$\nu = 1$	5076	7.7140	>100000	-
$\nu = 5$	989	1.4959	21182	184.7405
$\nu = 30$	**185**	**0.3054**	**2180**	**16.3661**
$\nu = 100$	688	1.1056	2216	16.0718
$\nu = 500$	3507	5.7172	15121	122.9190
$\nu = 1000$	7012	11.4579	30390	281.2456

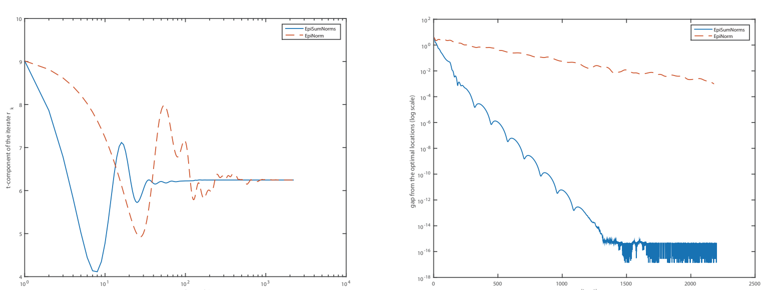

Figure 1: Comparison of the methods EpiSumNorms (blue solid line) and EpiNorm (red dashed line) in $\mathbb{R}^2$ for $\nu = 30$

Table 2: Performance evaluation for NumGiFac 30 and NumNewFac 10 in $\mathbb{R}^2$

MaxNumIt = 100000	EpiSumNorms		EpiNorm	
	NumIt	CPUtime	NumIt	CPUtime
$\nu = 0.1$	43058	106.5284	>100000	-
$\nu = 1$	4306	9.7233	>100000	-
$\nu = 10$	411	0.9267	26596	639.1180
$\nu = 18$	**269**	**0.6146**	14416	303.8991
$\nu = 50$	538	1.2406	**3478**	**62.2674**
$\nu = 100$	1122	2.6104	4762	88.3561
$\nu = 1000$	11324	26.8627	56615	1633.1276

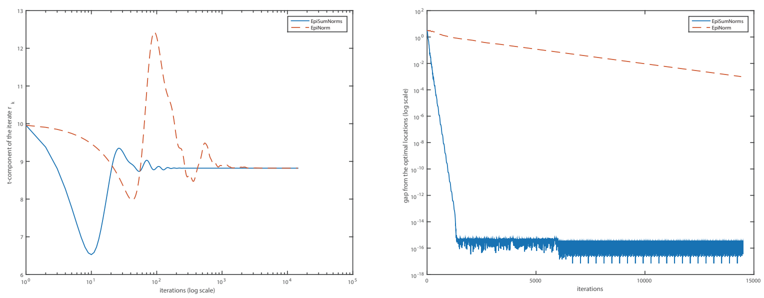

Figure 2: Comparison of the methods EpiSumNorms (blue solid line) and EpiNorm (red dashed line) in $\mathbb{R}^2$ for $\nu = 18$

Table 3: Performance evaluation for NumGiFac 60 and NumNewFac 20 in $\mathbb{R}^3$

MaxNumIt = 100000	EpiSumNorms		EpiNorm	
	NumIt	CPUtime	NumIt	CPUtime
$\nu = 1$	15415	89.1836	>100000	-
$\nu = 10$	1541	8.5561	>100000	-
$\nu = 98$	**592**	**3.3485**	28920	5332.9115
$\nu = 205$	1129	6.3206	**15697**	**2784.3976**
$\nu = 500$	2687	15.7556	16369	2831.1355
$\nu = 1000$	5346	29.7313	31429	5859.4418
$\nu = 5000$	26715	163.1581	>100000	-

In Table 1 it is shown that the parallel splitting algorithm converges very slow when employed in connection with the method proposed in [14], while our method performs much better. The corresponding figure shows that our method EpiSumNorms generates after 185 iterations a solution which is within the maximum bound from the optimal solution, while the method EpiNorm needs 2180 iterations. Take also note that in this example the location problem has in the form of EpiNorm 125 additional variables, while the examples in

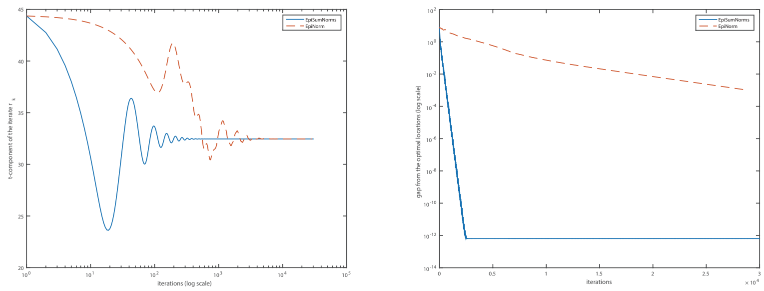

Figure 3: Comparison of the methods EpiSumNorms (blue solid line) and EpiNorm (red dashed line) in $\mathbb{R}^3$ for $\nu = 98$

the Table 2 and 3 have 300 and 1200 additional variables, respectively. For this reason our method by far outperforms the concept EpiNorm on such optimization problems regarding the accuracy as well as the CPU speed and number of iterations.

Finally, we consider the situation where $w_i = 1$ and $\beta_i = 2$ for all $i = 1, ..., n$.

Table 4: Performance evaluation for NumGiFac 25 and NumNewFac 5 in $\mathbb{R}^2$

MaxNumIt	EpiSumNorms		EpiNorm	
= 100000	NumIt	CPUtime	NumIt	CPUtime
$\nu = 0.1$	4967	4.6351	>100000	-
$\nu = 1$	663	0.6409	53618	583.0896
$\nu = 5$	**306**	**0.3172**	9504	70.9192
$\nu = 39$	2645	2.8821	**2851**	**21.4682**
$\nu = 100$	6776	7.4103	7120	53.9206
$\nu = 500$	33904	39.3127	35740	340.2518
$\nu = 1000$	67806	84.4253	71456	881.7213

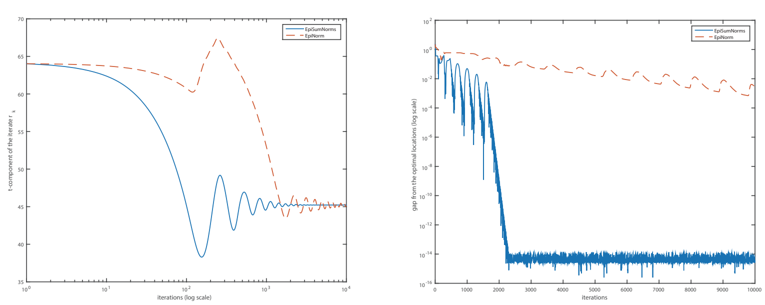

Figure 4: Comparison of the methods EpiSumNorms (blue solid line) and EpiNorm (red dashed line) in $\mathbb{R}^2$ for $\nu = 5$

Table 5: Performance evaluation for NumGiFac 60 and NumNewFac 10 in $\mathbb{R}^3$

MaxNumIt = 100000	EpiSumNorms		EpiNorm	
	NumIt	CPUtime	NumIt	CPUtime
$\nu = 0.1$	3391	8.2135	>100000	-
$\nu = 1$	**1042**	**2.5581**	>100000	-
$\nu = 10$	8270	20.7564	29000	1713.3125
$\nu = 50$	2714	6.8445	32168	2042.8242
$\nu = 110$	1691	4.3914	15167	821.7959
$\nu = 445$	6669	17.5405	**5224**	**273.9350**
$\nu = 1000$	15381	41.1793	14533	788.8773

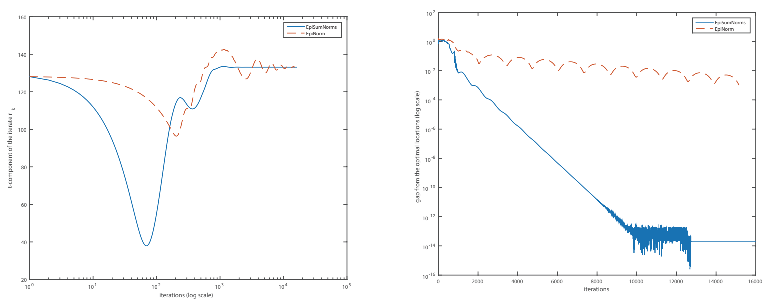

Figure 5: Comparison of the methods EpiSumNorm (blue solid line) and EpiNorm (red dashed line) in $\mathbb{R}^3$ for $\nu = 110$

The examples in the last two tables draw a similar picture as the examples in the previous ones. While the method EpiSumNorms generates a solution within the maximum bound from the optimal solution after few seconds, the method EpiNorm needs several minutes. This also points up the usefulness of our approach made in Section 2.

Acknowledgements

O. Wilfer would like to thank Prof. R. I. Boţ for the opportunity of two research stays at the Faculty of Mathematics of the University of Vienna as well as the many helpful remarks and comments which improved the quality of this article.

References

[1] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces,* CMS Books in Mathematics, Springer, New York, 2011.

[2] R.I. Boţ and E.R. Csetnek, On the convergence rate of a forward-backward type primal-dual splitting algorithm for convex optimization problems, *Optimization* 64 (2015) 5–23.

[3] R.I. Boţ, E.R. Csetnek, C. Hendrich and A. Heinrich, On the convergence rate improvement of a primal-dual splitting algorithm for solving monotone inclusion problems, *Math. Program.* 150 (2015) 251–279.

[4] R.I. Boţ, E.R. Csetnek and C. Hendrich, Recent developments on primal-dual splitting methods with applications to convex minimization, in: *Mathematics Without Boundaries: Surveys in Interdisciplinary Research,* P.M. Pardalos, T.M. Rassias (eds.), Springer-Verlag, New York, 2014, pp. 57–99.

[5] R.I. Boţ, S.-M. Grad and G. Wanka, *Duality in Vector Optimization,* Springer-Verlag, Berlin Heidelberg, 2009.

[6] R.I. Boţ and A. Heinrich, Regression tasks in machine learning via Fenchel duality, *Ann. Oper. Res.* 222 (2014) 197–211.

[7] R.I. Boţ and C. Hendrich, A Douglas-Rachford type primal-dual method for solving inclusions with mixtures of composite and parallel-sum type monotone operators, *SIAM J. Optim.* 23 (2013) 2541–2565.

[8] R. I. Boţ and C. Hendrich, A double smoothing technique for solving unconstrained nondifferentiable convex optimization problems, *Comput. Optim. Appl.* 54 (2013) 239–262.

[9] R.I. Boţ and C. Hendrich, Convergence analysis for a primal-dual monotone + skew splitting algorithm with applications to total variation minimization, *J. Math. Imaging Vision* 49 (2014) 551–568.

[10] R.I. Boţ and N. Lorenz, Optimization problems in statistical learning: duality and optimality conditions, *European J. Oper. Res.* 213 (2011) 395–404.

[11] E.C. Chi and K. Lange, Splitting methods for convex clustering, *J. Comput. Graph. Statist.* 24 (2015) 994–1013.

[12] G. Chierchia, N. Pustelnik, J.-C. Pesquet and B. Pesquet-Popescu, Epigraphical projection and proximal tools for solving constrained convex optimization problems: part I, *CoRR* abs/1210.5844, 2012.

[13] P. L. Combettes and J.-C. Pesquet, A proximal decomposition method for solving convex variational inverse problems, *Inverse Problems* 24 (2008).

[14] O. Cornejo and C. Michelot, A proximal solution for a class of extended minimax location problem. *ICCSA* 3483 (2005) 712–721.

[15] Z. Drezner, The weighted minimax location problem with set-up costs and extensions, *RAIRO Oper. Res.* 25 (1991) 55-64.

[16] H. Idrissi, O. Lefebvre and C. Michelot, Duality constrained multifacility location problems with mixed norms and applications, *Ann. Oper. Res.* 18 (1989) 71–92.

[17] N. Parikh and S. Boyd, Proximal algorithms, *Foundations and Trends in Optimization* 1 (2014) 127–239.

[18] J. Songsiri, Projection onto an $l1$-norm ball with application to identification of sparse autoregressive models, *ASAC* (2011).

[19] J.E. Spingarn, Partial inverse of a monotone operator, *Appl. Math. Optim.* 10 (1983) 247–265.

[20] G. Wanka and O. Wilfer, A Lagrange duality approach for multi-composed optimization problems, *TOP* 25 (2017) 288–313.

[21] G. Wanka and O. Wilfer, Duality results for extended multifacility location problems, *Optimization* 67 (2016) 1095–1119.

[22] G. Wanka and O. Wilfer, Duality results for nonlinear single minimax location problems via multi-composed optimization, *Math. Meth. Oper. Res.* 86 (2017) 401–439.

Manuscript received 8 November 2017
revised 22 November 2018, 7 March 2019
accepted for publication 11 March 2019

Gert Wanka
Faculty of Mathematics, Chemnitz University of Technology
D-09107 Chemnitz, Germany
E-mail address: gert.wanka@mathematik.tu-chemnitz.de

Oleg Wilfer
Faculty of Mathematics, Chemnitz University of Technology
D-09107 Chemnitz, Germany
E-mail address: oleg.wilfer@mathematik.tu-chemnitz.de