Enhanced computation of the proximity operator for perspective functions

Luis M. Briceño-Arias[1], Cristóbal Vivar-Vargas[2]

[1]Universidad Técnica Federico Santa María, Departamento de Matemática, Santiago, Chile
luis.briceno@usm.cl

[2]Universidad Técnica Federico Santa María, Departamento de Matemática, Santiago, Chile
cristobal.vivar@usm.cl

Abstract. In this paper we provide an explicit expression for the proximity operator of a perspective of any proper lower semicontinuous convex function defined on a Hilbert space. Our computation enhances and generalizes known formulae for the case when the Fenchel conjugate of the convex function has open domain or when it is radial. We provide several examples of non-radial functions for which the domain of its conjugate is not open and we compute the proximity operators of their perspectives.

Keywords Convex analysis · Fenchel conjugate · Perspective function · Proximity operator · Recession function

MSC (2020): 46N10 · 47J20 · 49J53 · 49N15 · 90C25.

1 Introduction

In this paper we enhance the computation of the proximity operator of the perspective of a proper lower semicontinuous convex function defined in a real Hilbert space $\mathcal{H}$. This construction is introduced in [25] and appears naturally in optimal mass transportation theory [3, 26], dynamical formulation of the 2-Wasserstein distance [3, 26], information theory [18], physics [5], operator theory [17], statistics [23], matrix analysis [16], signal processing and inverse problems [21, 20, 22], JKO [19] schemes for gradient flows in the space of probability measures [4, 11], and transportation and mean field games problems with state-dependent potentials [8, 9], among other disciplines.

The proximity operator of the perspective of a proper lower semicontinuous convex function $f\colon \mathcal{H} \to]-\infty, +\infty]$ is first obtained in [13, Theorem 3.1] in the case when the domain of f^* is open, where f^* denotes the Fenchel-Legendre conjugate of f. Some examples in the case when the domain f^* is closed are also explored in [13, Section 3.2]. These results need the solution of an inclusion in $\mathcal{H}$ in order to compute the proximity operator of the perspective of f. In the case when f is radial, the proximity operator of the perspective of f is studied without any assumptions on the domain of its conjugate in [14, Proposition 2.3]. This calculus needs a projection onto a particular convex subset of $\mathbb{R}^2$, which is not always easy to compute. Previous results do not allow the computation of the proximity operator of the perspective of non-radial functions f such that the domain of f^* is not open. This is the case, for instance, of entropy-based penalizations including the log-sum, which arises, e.g., in optimal transport problems and mathematical programming [1, 10].

The goal of this paper is to provide a simple computation of the proximity operator of the perspective of any proper lower semicontinuous convex function without any further assumption. Our computation generalizes the results in [13] and [14] and the solution of a scalar equation is needed, which can be obtained via standard root-finding methods [24]. We provide several examples of non-radial functions f such that the domain of f^* is not open, for which the proximity operator is not available in the literature.

The paper is organized as follows. In Section 2 we present our notation and preliminary results. The proximity operator for the perspective of f is presented in Section 3. Examples are studied in Section 4.

2 Notation and preliminaries

2.1 Notation

Throughout this paper, $\mathcal{H}$ is a real Hilbert space endowed with the inner product $\langle \cdot \mid \cdot \rangle$ and associated norm $\|\cdot\|$. $\mathcal{H} \oplus \mathbb{R}$ denotes the Hilbert direct sum between $\mathcal{H}$ and $\mathbb{R}$.

Let $f : \mathcal{H} \to]-\infty, +\infty]$. The domain of f is $\operatorname{dom} f = \{x \in \mathcal{H} \mid f(x) < +\infty\}$ and f is proper if $\operatorname{dom} f \neq \varnothing$. Denote by $\Gamma_0(\mathcal{H})$ the class of proper lower semicontinuous convex functions from $\mathcal{H}$ to $]-\infty, +\infty]$. Suppose that $f \in \Gamma_0(\mathcal{H})$. The recession function of f is

$$(\forall x_0 \in \operatorname{dom} f)(\forall x \in \mathcal{H}) \quad \operatorname{rec} f(x) = \lim_{t\to+\infty} \frac{f(x_0 + tx) - f(x_0)}{t} \tag{2.1}$$

and it satisfies (see [2, Proposition 7.13 & Proposition 13.49])

$$\operatorname{rec} f = \sigma_{\operatorname{dom} f^*} = \sigma_{\overline{\operatorname{dom}} f^*}. \tag{2.2}$$

The Fenchel conjugate of f is

$$f^* : \mathcal{H} \to]-\infty, +\infty] : u \mapsto \sup_{x\in\mathcal{H}} \left(\langle x \mid u\rangle - f(x)\right). \tag{2.3}$$

We have $f^* \in \Gamma_0(\mathcal{H})$, $f^{**} = f$, and we have the Fenchel-Young inequality [2, Proposition 13.15]

$$(\forall x \in \mathcal{H})(\forall u \in \mathcal{H}) \quad f(x) + f^*(y) \geq \langle x \mid u\rangle. \tag{2.4}$$

The subdifferential of f is the set-valued operator

$$\partial f : \mathcal{H} \to 2^{\mathcal{H}} : x \mapsto \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \quad \langle y - x \mid u \rangle + f(x) \leq f(y)\} \tag{2.5}$$

and $\operatorname{dom} \partial f = \{x \in \mathcal{H} \mid \partial f(x) \neq \varnothing\}$. We have the Fenchel-Young identity [2, Proposition 16.10]

$$(\forall x \in \mathcal{H})(\forall u \in \mathcal{H}) \quad u \in \partial f(x) \quad \Leftrightarrow \quad f(x) + f^*(u) = \langle x \mid u \rangle. \tag{2.6}$$

The *proximity operator* of f is

$$\operatorname{prox}_f : \mathcal{H} \to \mathcal{H} : x \mapsto \underset{y \in \mathcal{H}}{\arg\min} \left(f(y) + \frac{1}{2} \|x - y\|^2 \right), \tag{2.7}$$

which is characterized by

$$(\forall x \in \mathcal{H})(\forall p \in \mathcal{H}) \quad p = \operatorname{prox}_f x \quad \Leftrightarrow \quad x - p \in \partial f(p) \tag{2.8}$$

and satisfies

$$(\forall \gamma \in \,]0, +\infty[) \quad \operatorname{prox}_{\gamma f} = \operatorname{Id} - \gamma \operatorname{prox}_{f^*/\gamma} \circ (\operatorname{Id}/\gamma), \tag{2.9}$$

where $\operatorname{Id}\colon \mathcal{H} \to \mathcal{H}$ denotes the identity operator.

Let $C \subset \mathcal{H}$ be a nonempty closed convex set. The indicator function of C is

$$\iota_C : \mathcal{H} \to \,]-\infty, +\infty] : x \mapsto \begin{cases} 0, & \text{if } x \in C; \\ +\infty, & \text{if } x \notin C, \end{cases} \tag{2.10}$$

its support function is

$$\sigma_C : \mathcal{H} \to \,]-\infty, +\infty] : u \mapsto \sup_{x \in C} \langle x \mid u \rangle, \tag{2.11}$$

and we have $\sigma_C = (\iota_C)^*$. The projection operator onto C is $P_C = \operatorname{prox}_{\iota_C}$, which is characterized by

$$(\forall x \in \mathcal{H})(\forall y \in C) \quad \langle y - P_C x \mid x - P_C x \rangle \leq 0. \tag{2.12}$$

For further background on convex analysis, the reader is referred to [2].

2.2 Perspective functions and properties

Now, we review essential properties of perspective functions. We refer the reader to [12] for further background.

Definition 2.1. Let $f \in \Gamma_0(\mathcal{H})$. The *perspective* of f is:

$$\widetilde{f}\colon \mathcal{H} \times \mathbb{R} \to \,]-\infty, +\infty]\colon (x, \eta) \mapsto \begin{cases} \eta f\left(\dfrac{x}{\eta}\right), & \text{if } \eta > 0; \\ (\operatorname{rec} f)(x), & \text{if } \eta = 0; \\ +\infty, & \text{if } \eta < 0. \end{cases} \tag{2.13}$$

Lemma 2.1. *Let $f \in \Gamma_0(\mathcal{H})$. Then the following hold:*

(i) $\widetilde{f} \in \Gamma_0(\mathcal{H} \oplus \mathbb{R})$.

(ii) $\widetilde{f}$ *is not radial, i.e., there exist (x, η) and (y, ν) in $\mathcal{H} \times \mathbb{R}$ such that $\|(x, \eta)\| = \|(y, \nu)\|$ and $\widetilde{f}(x, \eta) \neq \widetilde{f}(y, \nu)$.*

(iii) *Let $C = \{(x, \eta) \in \mathcal{H} \times \mathbb{R} \mid \eta + f^*(x) \leq 0\}$. Then $\left(\widetilde{f}\right)^* = \iota_C$.*

Proof. (i): [12, Proposition 2.3(ii)].

(ii): Let $(x, \eta) \in \mathcal{H} \times]0, +\infty[$ be such that $x/\eta \in \operatorname{dom} f$. Then $\|(x, \eta)\| = \|(x, -\eta)\|$ and $\widetilde{f}(x, \eta) = \eta f(x/\eta) < +\infty = \widetilde{f}(x, -\eta)$. Hence f is not radial.

(iii): [12, Proposition 2.3(iv)]. □ □

2.3 Preliminaries on proximity operators

We now provide some preliminary results needed in the following sections.

Lemma 2.2. [7, Lemma 3.1(iii)-(iv)] *Let $f \in \Gamma_0(\mathcal{H})$, let $\gamma \in \left]0,+\infty\right[$, and let $(x,p) \in \mathcal{H} \times \mathcal{H}$. Then the following propositions are equivalent:*

(i) $p = \operatorname{prox}_{\gamma f} x$.

(ii) $f(p) + f^*((x-p)/\gamma) = \langle p \mid (x-p)/\gamma \rangle$.

(iii) $(\forall y \in \mathcal{H}) \quad \langle x-p \mid y-p \rangle + \gamma f(p) \leq \gamma f(y)$.

Proposition 2.1. *Let $f \in \Gamma_0(\mathcal{H})$, let $\gamma \in \left]0,+\infty\right[$, and let $x \in \mathcal{H}$. Then the following hold:*

(i) *We have*

$$-\left\langle x - P_{\overline{\operatorname{dom}} f}\, x \;\middle|\; \operatorname{prox}_{\gamma f} x - P_{\overline{\operatorname{dom}} f}\, x \right\rangle + \left\| P_{\overline{\operatorname{dom}} f}\, x - \operatorname{prox}_{\gamma f} x \right\|^2 \leq \gamma \left(f\left(P_{\overline{\operatorname{dom}} f}\, x\right) - f\left(\operatorname{prox}_{\gamma f} x\right)\right) \quad (2.14)$$

(ii) $f\left(\operatorname{prox}_{\gamma f} x\right) \leq f\left(P_{\overline{\operatorname{dom}} f}\, x\right)$.

Proof. (i): Let x and p in $\mathcal{H}$ be such that $p = \operatorname{prox}_{\gamma f} x$. Then, by setting $y = P_{\overline{\operatorname{dom}} f}\, x$, it follows from Lemma 2.2 that

$$-\left\langle x - P_{\overline{\operatorname{dom}} f}\, x \;\middle|\; p - P_{\overline{\operatorname{dom}}}\, x \right\rangle + \left\| P_{\overline{\operatorname{dom}} f}\, x - p \right\|^2 \leq \gamma \left(f\left(P_{\overline{\operatorname{dom}} f}\, x\right) - f(p)\right), \quad (2.15)$$

and (2.14) follows.

(ii): Since $\overline{\operatorname{dom}} f$ is nonempty, closed, convex, and $\operatorname{prox}_{\gamma f} x \in \operatorname{dom} \partial f \subset \overline{\operatorname{dom}} f$, it follows from (2.12) that

$$\left\langle x - P_{\overline{\operatorname{dom}} f}\, x \;\middle|\; \operatorname{prox}_{\gamma f} x - P_{\overline{\operatorname{dom}} f}\, x \right\rangle \leq 0. \quad (2.16)$$

Hence, the result follows from (i). ☐

3 Main results

The proximity operator of a perspective function when $\operatorname{dom} f^*$ is open is computed in [13, Theorem 3.1]. In the case of radial functions, this hypothesis is removed in [14, Proposition 2.3]. In this section we compute the proximity operator of $\widetilde{f}$ for any $f \in \Gamma_0(\mathcal{H})$.

Theorem 3.1. *Let $f \in \Gamma_0(\mathcal{H})$, let $\gamma \in \left]0,+\infty\right[$, and let $(x,\eta) \in \mathcal{H} \times \mathbb{R}$. Then the following hold:*

(i) *Suppose that $\eta + \gamma f^*\big(P_{\overline{\operatorname{dom}} f^*}(x/\gamma)\big) \leq 0$. Then*

$$\operatorname{prox}_{\gamma \widetilde{f}}(x,\eta) = \left(x - \gamma P_{\overline{\operatorname{dom}} f^*}\left(\frac{x}{\gamma}\right), 0 \right). \quad (3.1)$$

(ii) *Suppose that $\eta + \gamma f^*\big(P_{\overline{\operatorname{dom}} f^*}(x/\gamma)\big) > 0$. Then there exists a unique $\mu \in \left]0, \eta + \gamma f^*\big(P_{\overline{\operatorname{dom}} f^*}(x/\gamma)\big)\right]$ such that*

$$\mu = \eta + \gamma f^*\left(\operatorname{prox}_{\frac{\mu}{\gamma} f^*}\left(\frac{x}{\gamma}\right)\right). \quad (3.2)$$

Furthermore

$$\operatorname{prox}_{\gamma \widetilde{f}}(x,\eta) = \left(x - \gamma \operatorname{prox}_{\frac{\mu}{\gamma} f^*}\left(\frac{x}{\gamma}\right), \mu \right). \quad (3.3)$$

Proof. First note that Lemma 2.1(i) asserts that $\widetilde{f} \in \Gamma_0(\mathcal{H}\oplus\mathbb{R})$. Let $(p,\mu) \in \mathcal{H}\times\mathbb{R}$ be such that $(p,\mu) = \operatorname{prox}_{\gamma\widetilde{f}}(x,\eta)$. It follows from Lemma 2.1 and Lemma 2.2 that

$$\begin{aligned}
(p,\mu) = \operatorname{prox}_{\gamma\widetilde{f}}(x,\eta) \;&\Leftrightarrow\; \widetilde{f}(p,\mu) + \left(\widetilde{f}\right)^*\left(\frac{x-p}{\gamma}, \frac{\eta-\mu}{\gamma}\right) = \left\langle (p,\mu) \;\middle|\; \left(\frac{x-p}{\gamma}, \frac{\eta-\mu}{\gamma}\right)\right\rangle \\
&\Leftrightarrow\; \widetilde{f}(p,\mu) + \iota_C\left(\frac{x-p}{\gamma}, \frac{\eta-\mu}{\gamma}\right) = \left\langle p \;\middle|\; \frac{x-p}{\gamma}\right\rangle + \mu\left(\frac{\eta-\mu}{\gamma}\right) \\
&\Leftrightarrow\; \widetilde{f}(p,\mu) = \left\langle p \;\middle|\; \frac{x-p}{\gamma}\right\rangle + \mu\left(\frac{\eta-\mu}{\gamma}\right) \text{ and } \frac{\eta-\mu}{\gamma} + f^*\left(\frac{x-p}{\gamma}\right) \leq 0. \qquad (3.4)
\end{aligned}$$

Moreover, since $(p,\mu) \in \operatorname{dom}\partial\widetilde{f} \subset \operatorname{dom}\widetilde{f}$, we have $\mu \in [0,+\infty[$. Then, let us consider two cases.

(i) Suppose that $\mu = 0$. Then (3.4), (2.2), Lemma 2.2, and (2.9) imply

$$\begin{aligned}
(p,0) = \operatorname{prox}_{\gamma\widetilde{f}}(x,\eta) \;&\Leftrightarrow\; (\operatorname{rec} f)(p) = \left\langle p \;\middle|\; \frac{x-p}{\gamma}\right\rangle \quad\text{and}\quad \frac{\eta}{\gamma} + f^*\left(\frac{x-p}{\gamma}\right) \leq 0 \\
&\Leftrightarrow\; \sigma_{\overline{\operatorname{dom}} f^*}(p) + \iota_{\overline{\operatorname{dom}} f^*}\left(\frac{x-p}{\gamma}\right) = \left\langle p \;\middle|\; \frac{x-p}{\gamma}\right\rangle \quad\text{and}\quad \frac{\eta}{\gamma} + f^*\left(\frac{x-p}{\gamma}\right) \leq 0 \\
&\Leftrightarrow\; p = \operatorname{prox}_{\gamma\sigma_{\overline{\operatorname{dom}} f^*}} x \quad\text{and}\quad \frac{\eta}{\gamma} + f^*\left(\frac{x-p}{\gamma}\right) \leq 0 \\
&\Leftrightarrow\; p = x - \gamma P_{\overline{\operatorname{dom}} f^*}\left(\frac{x}{\gamma}\right) \quad\text{and}\quad \eta + \gamma f^*\left(P_{\overline{\operatorname{dom}} f^*}\left(\frac{x}{\gamma}\right)\right) \leq 0. \qquad (3.5)
\end{aligned}$$

(ii) Suppose that $\mu > 0$. Then it follows from Lemma 2.2, (2.4), and (2.9) that

$$\begin{aligned}
(p,\mu) = \operatorname{prox}_{\gamma\widetilde{f}}(x,\eta) \;&\Leftrightarrow\; \mu f\left(\frac{p}{\mu}\right) = \left\langle p \;\middle|\; \frac{x-p}{\gamma}\right\rangle + \mu\left(\frac{\eta-\mu}{\gamma}\right) \\
&\qquad \text{and } \frac{\eta-\mu}{\gamma} + f^*\left(\frac{x-p}{\gamma}\right) \leq 0 \\
&\Leftrightarrow\; f\left(\frac{p}{\mu}\right) = \left\langle \frac{p}{\mu} \;\middle|\; \frac{x-p}{\gamma}\right\rangle + \frac{\eta-\mu}{\gamma} \\
&\qquad \text{and } \frac{\eta-\mu}{\gamma} + f^*\left(\frac{x-p}{\gamma}\right) \leq 0 \\
&\Leftrightarrow\; f\left(\frac{p}{\mu}\right) + f^*\left(\frac{x-p}{\gamma}\right) = \left\langle \frac{p}{\mu} \;\middle|\; \frac{x-p}{\gamma}\right\rangle + \frac{\eta-\mu}{\gamma} + f^*\left(\frac{x-p}{\gamma}\right) \\
&\qquad \text{and } \frac{\eta-\mu}{\gamma} + f^*\left(\frac{x-p}{\gamma}\right) \leq 0 \\
&\Leftrightarrow\; f\left(\frac{p}{\mu}\right) + f^*\left(\frac{x-p}{\gamma}\right) = \left\langle \frac{p}{\mu} \;\middle|\; \frac{x-p}{\gamma}\right\rangle \\
&\qquad \text{and } \frac{\eta-\mu}{\gamma} = -f^*\left(\frac{x-p}{\gamma}\right) \\
&\Leftrightarrow\; \frac{p}{\mu} = \operatorname{prox}_{\frac{\gamma}{\mu}f}\left(\frac{x}{\mu}\right) \text{ and } \mu = \eta + \gamma f^*\left(\frac{x-p}{\gamma}\right) \\
&\Leftrightarrow\; p = x - \gamma\operatorname{prox}_{\frac{\mu}{\gamma}f^*}\left(\frac{x}{\gamma}\right) \text{ and } \mu = \eta + \gamma f^*\left(\operatorname{prox}_{\frac{\mu}{\gamma}f^*}\left(\frac{x}{\gamma}\right)\right). \qquad (3.6)
\end{aligned}$$

Hence, Proposition 2.1((ii)) implies that

$$0 < \mu = \eta + \gamma f^*\left(\operatorname{prox}_{\frac{\mu}{\gamma}f^*}\left(\frac{x}{\gamma}\right)\right) \leq \eta + \gamma f^*\left(P_{\overline{\operatorname{dom}} f^*}\left(\frac{x}{\gamma}\right)\right). \tag{3.7}$$

Altogether, if $\eta + \gamma f^*(P_{\overline{\operatorname{dom}} f^*}(x/\gamma)) \leq 0$ and supposing that $\mu > 0$, we arrive at a contradiction with (3.7) and therefore (3.1) follows from (3.5). Conversely, if $\eta + \gamma f^*(P_{\overline{\operatorname{dom}} f^*}(x/\gamma)) > 0$ and supposing that $\mu = 0$, we arrive at a contradiction with (3.5) and therefore (3.3) and (3.2) follow from (3.6) and [7, Lemma 3.2(ii) & Lemma 3.2(iii)]. ☐

In the case when f^* has open domain, Theorem 3.1 recovers [13, Theorem 3.1], as the following example illustrates.

Example 3.1. Let $f \in \Gamma_0(\mathcal{H})$, let $\gamma \in \,]0, +\infty[$, let $\eta \in \mathbb{R}$, and let $x \in \mathcal{H}$. Then the following hold:

(i) Suppose that $\eta + \gamma f^*(x/\gamma) \leq 0$. Then $\operatorname{prox}_{\gamma \widetilde{f}}(x, \eta) = (0, 0)$.

(ii) Suppose that $\operatorname{dom} f^*$ is open and that $\eta + \gamma f^*(x/\gamma) > 0$. Then

$$\operatorname{prox}_{\gamma \widetilde{f}}(x, \eta) = (x - \gamma p, \eta + \gamma f^*(p)), \tag{3.8}$$

where p is the unique solution to the inclusion

$$x \in \gamma p + (\eta + \gamma f^*(p))\partial f^*(p). \tag{3.9}$$

If f^* is differentiable at p, then p is characterized by $y = \gamma p + (\eta + \gamma f^*(p))\nabla f^*(p)$.

Proof. (i): Note that $\eta + \gamma f^*(x/\gamma) \leq 0$ implies that $x/\gamma \in \operatorname{dom} f^* \subset \overline{\operatorname{dom}} f^*$ and then $P_{\overline{\operatorname{dom}} f^*}(x/\gamma) = x/\gamma$. Therefore, it follows from Theorem 3.1(i) that

$$\operatorname{prox}_{\gamma \widetilde{f}}(x, \eta) = \left(x - \gamma P_{\overline{\operatorname{dom}} f^*}\left(\frac{x}{\gamma}\right), 0\right) = \left(x - \gamma \frac{x}{\gamma}, 0\right) = (0, 0). \tag{3.10}$$

(ii): Suppose that $\eta + \gamma f^*(P_{\overline{\operatorname{dom}} f^*}(x/\gamma)) \leq 0$. Then, since $\operatorname{dom} f^*$ is open, $P_{\overline{\operatorname{dom}} f^*}(x/\gamma) \in \operatorname{dom} f^* = \operatorname{int}(\operatorname{dom} f^*)$ and it follows from Theorem 3.1(i) that $\operatorname{prox}_{\gamma \widetilde{f}}(x, \eta) = (x - \gamma P_{\overline{\operatorname{dom}} f^*}(x/\gamma), 0)$. Hence, (3.4), (2.2), the fact that $\iota_{\overline{\operatorname{dom}} f^*}(P_{\overline{\operatorname{dom}} f^*}(x/\gamma)) = 0$, and (2.6) yield

$$\begin{aligned}
\widetilde{f}\left(x - \gamma P_{\overline{\operatorname{dom}} f^*}\left(\frac{x}{\gamma}\right), 0\right) &= \operatorname{rec} f\left(x - \gamma P_{\overline{\operatorname{dom}} f^*}\left(\frac{x}{\gamma}\right)\right) \\
&= \sigma_{\overline{\operatorname{dom}} f^*}\left(x - \gamma P_{\overline{\operatorname{dom}} f^*}\left(\frac{x}{\gamma}\right)\right) \\
&= \left\langle x - \gamma P_{\overline{\operatorname{dom}} f^*}\left(\frac{x}{\gamma}\right) \,\middle|\, P_{\overline{\operatorname{dom}} f^*}\left(\frac{x}{\gamma}\right)\right\rangle.
\end{aligned} \tag{3.11}$$

Therefore, by [2, Corollary 7.6(i)], $P_{\overline{\operatorname{dom}} f^*}(x/\gamma) \in \overline{\operatorname{dom}} f^* \setminus \operatorname{int}(\operatorname{dom} f^*)$ which is a contradiction. Therefore $\eta + \gamma f^*(P_{\overline{\operatorname{dom}} f^*}(x/\gamma)) > 0$ and it follows from Theorem 3.1(ii) that

$$\operatorname{prox}_{\gamma \widetilde{f}}(x, \eta) = \left(x - \gamma \operatorname{prox}_{\frac{\mu}{\gamma}f^*}\left(\frac{x}{\gamma}\right), \mu\right), \tag{3.12}$$

where μ is the solution of the equation

$$\mu = \eta + \gamma f^*\left(\operatorname{prox}_{\frac{\mu}{\gamma}f^*}\left(\frac{x}{\gamma}\right)\right). \tag{3.13}$$

Now, given $p \in \mathcal{H}$, it follows from (2.8) that

$$p = \operatorname{prox}_{\frac{\mu}{\gamma}f^*}\left(\frac{x}{\gamma}\right) \Leftrightarrow x \in \gamma p + \mu \partial f^*(p) = \gamma p + (\eta + \gamma f^*(p))\partial f^*(p). \tag{3.14}$$

Hence, (3.8) and (3.9) follow from (3.12) and (3.14). Lastly, the claim when f^* is differentiable follows from $\partial f^*(p) = \{\nabla f^*(p)\}$. ☐ ☐

Remark 3.1. Note that [13, Theorem 3.1] needs the solution to the inclusion (3.9) for computing $\operatorname{prox}_{\gamma\widetilde{f}}$. By contrast, Theorem 3.1 only needs to solve a scalar equation, which can be obtained via standard root-finding methods [24].

The next result illustrates Theorem 3.1 in the particular case of radial functions.

Proposition 3.1. *Let $\phi \in \Gamma_0(\mathbb{R})$ be even, set $f = \phi \circ \|\cdot\|$, let $\gamma \in \,]0, +\infty[$, and let $(x, \eta) \in \mathcal{H} \times \mathbb{R}$. Then the following hold:*

(i) *Suppose that $\eta + \gamma\phi^* \left(P_{\overline{\operatorname{dom}\,\phi^*}} (\|x\| / \gamma)\right) \leq 0$. Then*

$$\operatorname{prox}_{\gamma\widetilde{f}}(x, \eta) = \begin{cases} \left(\left(1 - \gamma \dfrac{P_{\overline{\operatorname{dom}\,\phi^*}} (\|x\|/\gamma)}{\|x\|}\right) x, 0\right), & \text{if } x \neq 0; \\ (0, 0), & \text{if } x = 0. \end{cases} \tag{3.15}$$

(ii) *Suppose that $\eta + \gamma\phi^* \left(P_{\overline{\operatorname{dom}\,\phi^*}} (\|x\| / \gamma)\right) > 0$. Then there exists a unique $\mu \in \,]0, +\infty[$ such that*

$$\mu = \eta + \gamma\phi^* \left(\operatorname{prox}_{\frac{\mu}{\gamma}\phi^*} \left(\frac{\|x\|}{\gamma}\right)\right). \tag{3.16}$$

Furthermore

$$\operatorname{prox}_{\gamma\widetilde{f}}(x, \eta) = \begin{cases} \left(\left(1 - \gamma \dfrac{\operatorname{prox}_{\frac{\mu}{\gamma}\phi^*} (\|x\|/\gamma)}{\|x\|}\right) x, \mu\right), & \text{if } x \neq 0; \\ (0, \eta + \gamma\phi^*(0)), & \text{if } x = 0. \end{cases} \tag{3.17}$$

Proof. First, since ϕ is even, [2, Proposition 13.21] implies that ϕ^* is even and [2, Example 13.8] yields

$$f^* = \phi^* \circ \| \cdot \|. \tag{3.18}$$

Hence,

$$f(0) = \phi(0) = \inf_{x \in \mathbb{R}} \phi(x) = -\phi^*(0) = -f^*(0), \tag{3.19}$$

and it follows from $\phi^* \in \Gamma_0(\mathbb{R})$ and [7, Lemma 3.1(i) & Lemma 3.2(ii)] that

$$(\forall x \in \mathcal{H}) \quad P_{\overline{\operatorname{dom}\,f^*}}(x) = \begin{cases} P_{\overline{\operatorname{dom}\,\phi^*}}(\|x\|) \dfrac{x}{\|x\|}, & \text{if } x \neq 0; \\ 0, & \text{if } x = 0. \end{cases} \tag{3.20}$$

Therefore (3.18), (3.19), and (3.20) yield

$$(\forall x \in \mathcal{H}) \quad f^* \left(P_{\overline{\operatorname{dom}\,f^*}} x\right) = \phi^* \left(P_{\overline{\operatorname{dom}\,\phi^*}}(\|x\|)\right). \tag{3.21}$$

(i): In this case (3.21) yields

$$\eta + \gamma f^* \left(P_{\overline{\operatorname{dom}\,f^*}} \left(\frac{x}{\gamma}\right)\right) \leq 0. \tag{3.22}$$

Therefore, the result follows from Theorem 3.1(i) and (3.20).

(ii): In this case (3.21) yields

$$\eta + \gamma f^* \left(P_{\overline{\operatorname{dom}\,f^*}} \left(\frac{x}{\gamma}\right)\right) > 0. \tag{3.23}$$

Hence, it follows from Theorem 3.1(ii) and [7, Lemma 3.2(iv)] that there exists a unique $\mu \in \,]0,+\infty[$ such that

$$\mu = \eta + \gamma f^*\left(\operatorname{prox}_{\frac{\mu}{\gamma} f^*}\left(\frac{x}{\gamma}\right)\right) = \eta + \gamma\phi^*\left(\operatorname{prox}_{\frac{\mu}{\gamma}\phi^*}\left(\frac{\|x\|}{\gamma}\right)\right), \tag{3.24}$$

and

$$\operatorname{prox}_{\gamma\widetilde{f}}(x,\eta) = \begin{cases} \left(\left(1-\gamma\dfrac{\operatorname{prox}_{\frac{\mu}{\gamma}\phi^*}(\|x\|/\gamma)}{\|x\|}\right)x, \mu\right), & \text{if } x \neq 0;\\ (0,\mu), & \text{if } x = 0. \end{cases} \tag{3.25}$$

Furthermore, for $x = 0$, (3.24) and (3.25) yield $\mu = \eta + \gamma\phi^*(\operatorname{prox}_{\mu\phi^*/\gamma}(0))$ and $\operatorname{prox}_{\gamma\widetilde{f}}(x,\eta) = (-\gamma \operatorname{prox}_{\mu f^*/\gamma}(0), \mu)$. Next, since $\phi(0) = -\phi^*(0)$ and $\mu \geq 0$, it follows from (2.6) and (2.8) that

$$0 = \phi(0) + \phi^*(0) \;\Leftrightarrow\; 0 \in \frac{\mu}{\gamma}\partial\phi^*(0) \;\Leftrightarrow\; 0 = \operatorname{prox}_{\frac{\mu}{\gamma}\phi^*}(0). \tag{3.26}$$

Therefore, the result follows from (3.24), (3.25), and (3.26). □ □

Remark 3.2. Note that [14, Proposition 2.3] can be obtained from Proposition 3.1. Indeed, let us define

$$\begin{aligned} \mathcal{R} &= \left\{(\nu,\chi) \in \mathbb{R}^2 \;\middle|\; \chi + \phi^*(\nu) \leq 0\right\}\\ \text{and} \quad \mathcal{S} &= \left\{(\nu,\chi) \in \mathbb{R}^2 \;\middle|\; \chi + \phi^*\left(P_{\overline{\operatorname{dom}}\,\phi^*}\nu\right) \leq 0\right\} \end{aligned} \tag{3.27}$$

and note that $\mathcal{R}$ is nonempty, closed, and convex, since it is the level set of the proper lower semicontinuous convex function $(\nu,\chi) \mapsto \chi + \phi^*(\nu)$. Now let (x,η) and $(\overline{\nu},\overline{\chi})$ in $\mathcal{H}\times\mathbb{R}$. It follows from (2.12) that $(\overline{\nu},\overline{\chi}) = P_{\mathcal{R}}(\|x\|/\gamma, \eta/\gamma)$ if and only if

$$(\forall(\nu,\chi) \in \mathcal{R}) \quad (\nu - \overline{\nu})\left(\frac{\|x\|}{\gamma} - \overline{\nu}\right) + (\chi - \overline{\chi})\left(\frac{\eta}{\gamma} - \overline{\chi}\right) \leq 0. \tag{3.28}$$

Moreover, since $\mathcal{R} \subset \mathcal{S}$, we have three cases:

(i) $(\|x\|/\gamma, \eta/\gamma) \in \mathcal{R}$: In this case we have $\|x\|/\gamma \in \operatorname{dom}\phi^*$, and, then, $P_{\overline{\operatorname{dom}}\,\phi^*}(\|x\|/\gamma) = \|x\|/\gamma$. Hence, it follows from Proposition 3.1(i) that

$$\operatorname{prox}_{\gamma f}(x,\eta) = (0,0). \tag{3.29}$$

(ii) $(\|x\|/\gamma, \eta/\gamma) \in \mathcal{S}\setminus\mathcal{R}$: In this case, note that $(\nu,\chi) \in \mathcal{R}$ implies $\nu \in \operatorname{dom}\phi^* \subset \overline{\operatorname{dom}}\,\phi^*$ and, hence, (2.12) yields

$$\begin{aligned} (\forall(\nu,\chi) \in \mathcal{R}) \quad 0 &\geq \left(\nu - P_{\overline{\operatorname{dom}}\,\phi^*}\left(\frac{\|x\|}{\gamma}\right)\right)\left(\frac{\|x\|}{\gamma} - P_{\overline{\operatorname{dom}}\,\phi^*}\left(\frac{\|x\|}{\gamma}\right)\right)\\ &= \left(\nu - P_{\overline{\operatorname{dom}}\,\phi^*}\left(\frac{\|x\|}{\gamma}\right)\right)\left(\frac{\|x\|}{\gamma} - P_{\overline{\operatorname{dom}}\,\phi^*}\left(\frac{\|x\|}{\gamma}\right)\right)\\ &\quad + \left(\chi - \frac{\eta}{\gamma}\right)\left(\frac{\eta}{\gamma} - \frac{\eta}{\gamma}\right). \end{aligned} \tag{3.30}$$

Therefore, (3.28) implies $(\overline{\nu},\overline{\chi}) = (P_{\overline{\operatorname{dom}}\,\phi^*}(\|x\|/\gamma), \eta/\gamma) = P_{\mathcal{R}}(\|x\|/\gamma, \eta/\gamma)$ and Proposition 3.1(i) yields

$$\begin{aligned} \operatorname{prox}_{\gamma\widetilde{f}}(x,\eta) &= \begin{cases} \left(\left(1 - \frac{\gamma\overline{\nu}}{\|x\|}\right)x, \eta - \gamma\overline{\chi}\right), & \text{if } x \neq 0;\\ (0,0), & \text{if } x = 0. \end{cases}\\ &= \begin{cases} \left(\left(1 - \gamma\dfrac{P_{\overline{\operatorname{dom}}\,\phi^*}(\|x\|/\gamma)}{\|x\|}\right)x, 0\right), & \text{if } x \neq 0;\\ (0,0), & \text{if } x = 0. \end{cases} \end{aligned} \tag{3.31}$$

(iii) $(\|x\|/\gamma, \eta/\gamma) \in \mathbb{R}^2 \setminus \mathcal{S}$: In this case, recalling that $\phi^*(0) = -\phi(0)$, we obtain from Proposition 3.1(ii) that $\operatorname{prox}_{\gamma\widetilde{f}}(0,\eta) = (0, \eta - \gamma\phi(0))$. On the other hand, set

$$(\overline{\nu}, \overline{\chi}) = \left(\operatorname{prox}_{\frac{\mu}{\gamma}\phi^*}\left(\frac{\|x\|}{\gamma}\right), -\phi^*\left(\operatorname{prox}_{\frac{\mu}{\gamma}\phi^*}\left(\frac{\|x\|}{\gamma}\right)\right)\right), \tag{3.32}$$

where $\mu \in]0,+\infty[$ is the unique solution to (3.16) guaranteed by Proposition 3.1(ii). It follows from Lemma 2.2 that

$$(\forall \nu \in \operatorname{dom}\phi^*) \quad (\nu - \overline{\nu})\left(\frac{\|x\|}{\gamma} - \overline{\nu}\right) \leq \frac{\mu}{\gamma}(\phi^*(\nu) - \phi^*(\overline{\nu})). \tag{3.33}$$

Now, let $(\nu, \chi) \in \mathcal{R}$ and recall that $\nu \in \operatorname{dom}\phi^*$. Hence, (3.33), (3.16), and (3.32) yield

$$\begin{aligned}
(\nu - \overline{\nu})\left(\frac{\|x\|}{\gamma} - \overline{\nu}\right) + (\chi - \overline{\chi})\left(\frac{\eta}{\gamma} - \overline{\chi}\right) &\leq \frac{\mu}{\gamma}(\phi^*(\nu) - \phi^*(\overline{\nu})) + (\chi - \overline{\chi})\left(\frac{\eta}{\gamma} - \overline{\chi}\right) \\
&= \left(\frac{\eta}{\gamma} + \phi^*(\overline{\nu})\right)(\phi^*(\nu) - \phi^*(\overline{\nu})) + (\chi + \phi^*(\overline{\nu}))\left(\frac{\eta}{\gamma} + \phi^*(\overline{\nu})\right) \\
&= \left(\frac{\eta}{\gamma} + \phi^*(\overline{\nu})\right)(\phi^*(\nu) + \chi) \\
&= \frac{\mu}{\gamma}(\phi^*(\nu) + \chi) \\
&\leq 0.
\end{aligned} \tag{3.34}$$

Therefore, $(\overline{\nu}, \overline{\chi}) = P_{\mathcal{R}}(\|x\|/\gamma, \eta/\gamma)$ and [14, Proposition 2.3(iii)] is obtained from Proposition 3.1(ii).

Altogether, we deduce that [14, Proposition 2.3] is deduced from Proposition 3.1. Note that, the formulae in [14, Proposition 2.3] need the computation on $\mathcal{R} \subset \mathbb{R}^2$, which can be complicated in some instances, as the following example illustrates.

Example 3.2. In the context of Proposition 3.1, let $\phi\colon x \mapsto x^2/2$. Then, $\phi^* = \phi$, $\operatorname{dom}\phi^* = \mathbb{R}$, and, for every $\tau \in]0,+\infty[$, $\operatorname{prox}_{\tau\phi^*} = \operatorname{Id}/(1+\tau)$. Therefore, given $(x,\eta) \in \mathcal{H} \times \mathbb{R}$, Proposition 3.1 yields

$$\operatorname{prox}_{\gamma\widetilde{f}}(x,\eta) = \begin{cases} (0,0), & \text{if } \eta + \|x\|^2/(2\gamma) \leq 0; \\ (0,\eta), & \text{if } \eta + \|x\|^2/(2\gamma) > 0 \text{ and } x = 0; \\ \left(\frac{\mu}{\gamma+\mu}x, \mu\right), & \text{if } \eta + \|x\|^2/(2\gamma) > 0 \text{ and } x \neq 0, \end{cases} \tag{3.35}$$

where $\mu \in]0,+\infty[$ is the unique solution to

$$\mu = \eta + \frac{\gamma}{2(\gamma+\mu)^2}\|x\|^2. \tag{3.36}$$

On the other hand, the proximity operator of the perspective proposed in [14, Proposition 2.3] needs the projection onto

$$\mathcal{R} = \left\{(\nu,\chi) \in \mathbb{R}^2 \;\middle|\; \chi + \nu^2/2 \leq 0\right\},$$

which involves additional computations.

4 Examples

In this section, we provide several instances in which f is non-radial, $\operatorname{dom} f^*$ is not open, and Theorem 3.1 allows us to compute the proximity operator of f. For this class of functions, the computation of the proximity operator of their perspectives is not available in the literature.

Let us start with the computation of the proximity operator of the perspective of the perspective of a lower semicontinuous convex function.

Example 4.1. Let $\mathcal{G}$ be a real Hilbert space, let $g \in \Gamma_0(\mathcal{G})$, and set $f = \widetilde{g}$. Then, Lemma 2.1(i) yields $f \in \Gamma_0(\mathcal{G}\times\mathbb{R})$ and, since f is positively homogeneous [12, Proposition 2.3(i)], we have

$$\widetilde{f}\colon ((x,\eta),\delta) \mapsto \begin{cases} \eta g\left(\dfrac{x}{\eta}\right), & \text{if } \eta > 0 \text{ and } \delta \geq 0;\\ (\operatorname{rec} g)(x), & \text{if } \eta = 0 \text{ and } \delta \geq 0;\\ +\infty, & \text{if } \eta < 0 \text{ or } \delta < 0. \end{cases} \tag{4.1}$$

Moreover, by defining $C = \{(x,\eta) \in \mathcal{G}\times\mathbb{R} \mid \eta + g^*(x) \leq 0\}$, it follows from Lemma 2.1(iii) that

$$f^* = \iota_C \quad \text{and} \quad (\forall \tau \in \left]0,+\infty\right[) \quad \operatorname{prox}_{\tau f^*} = P_{\overline{\operatorname{dom} f^*}} = P_C. \tag{4.2}$$

Hence, since $g^* \in \Gamma_0(\mathcal{G})$, $\operatorname{dom} f^* = C$ is closed. Now, in order to compute the proximity operator of $\widetilde{f}$, fix $(x,\eta) \in \mathcal{G}\times\mathbb{R}$, $\delta \in \mathbb{R}$, $\gamma \in \left]0,+\infty\right[$, and note that

$$\delta + \gamma f^*\left(P_{\overline{\operatorname{dom} f^*}}\left(\frac{x}{\gamma},\frac{\eta}{\gamma}\right)\right) = \delta + \gamma\iota_C\left(P_C\left(\frac{x}{\gamma},\frac{\eta}{\gamma}\right)\right) = \delta. \tag{4.3}$$

Therefore, by considering $\mathcal{H} = \mathcal{G}\oplus\mathbb{R}$, we deduce from Lemma 2.1(i) and Theorem 3.1(i) that, if $\delta \leq 0$, $\operatorname{prox}_{\gamma\widetilde{f}}$ is computed in (3.1). On the other hand, if $\delta > 0$, Theorem 3.1(ii) asserts that there exists a unique $\mu \in \left]0,\delta\right]$ solution to $\mu = \delta + \gamma f^*(\operatorname{prox}_{\mu f^*/\gamma}(x/\gamma))$ and $\operatorname{prox}_{\gamma\widetilde{f}}$ is obtained in (3.3). Altogether, noting that (4.2) implies that $\mu = \delta$, we derive from (2.9) and again from Theorem 3.1 that

$$\begin{aligned} \operatorname{prox}_{\gamma\widetilde{f}}((x,\eta),\delta) &= \left((x,\eta) - \gamma P_C\left(\frac{x}{\gamma},\frac{\eta}{\gamma}\right), \max\{0,\delta\}\right)\\ &= \left(\operatorname{prox}_{\gamma\widetilde{g}}(x,\eta), \max\{0,\delta\}\right)\\ &= \begin{cases} \left(x - \gamma P_{\overline{\operatorname{dom} g^*}}\left(\frac{x}{\gamma}\right), 0, \max\{0,\delta\}\right), & \text{if } \ell(x,\gamma) \leq 0;\\ \left(x - \gamma \operatorname{prox}_{\frac{\nu}{\gamma}g^*}\left(\frac{x}{\gamma}\right), \nu, \max\{0,\delta\}\right), & \text{if } \ell(x,\gamma) > 0, \end{cases} \end{aligned} \tag{4.4}$$

where $\nu \in \left]0,\ell(x,\gamma)\right]$ is the unique solution to $\nu = \eta + \gamma g^*(\operatorname{prox}_{\nu g^*/\gamma}(x/\gamma))$ and we denote $\ell(x,\eta) = \eta + \gamma g^*(P_{\overline{\operatorname{dom} g^*}}(x/\gamma))$.

In the particular case when $\mathcal{G} = \mathbb{R}$ and $g\colon \xi \mapsto \xi^2/2$, we have $g = g^*$, $\operatorname{dom} g = \mathbb{R}$, and, for every $\tau \in \left]0,+\infty\right[$, $\operatorname{prox}_{\tau g^*} = \operatorname{Id}/(1+\tau)$. Therefore, (4.4) reduces to

$$\operatorname{prox}_{\gamma f}((\xi,\eta),\delta) = \begin{cases} (0,0,\max\{0,\delta\}), & \text{if } \eta + \xi^2/(2\gamma) \leq 0;\\ \left(\frac{\nu}{\gamma+\nu}\xi, \nu, \max\{0,\delta\}\right), & \text{if } \eta + \xi^2/(2\gamma) > 0, \end{cases} \tag{4.5}$$

where $\nu \in \left]0,\eta + \xi^2/(2\gamma)\right]$ is the unique solution to the cubic equation $\nu = \eta + \gamma\xi^2/(2(\gamma+\nu)^2)$.

The following three examples are motivated by penalty methods for solving convex-constrained mathematical programming problems investigated in [1]. The last example also appears in the dual of entropy-penalized transport problems [10].

Example 4.2. Set

$$\psi\colon \xi \mapsto \begin{cases} -\ln(\xi), & \text{if } \xi > 0;\\ +\infty, & \text{if } \xi \leq 0, \end{cases} \tag{4.6}$$

set $\varphi = \psi + \iota_{]-\infty,1]}$, and set $f = \varphi^*$. Since [2, Example 13.2(iii)] implies that $\psi \in \Gamma_0(\mathbb{R})$ and $\left]-\infty,1\right]$ is closed and convex, we have $\varphi \in \Gamma_0(\mathbb{R})$. Moreover, since $\operatorname{dom}\psi \cap \left]-\infty,1\right[\neq \varnothing$, it follows from [2, Theorem 15.3 & Example 13.2(iii)] and simple computations that $f \in \Gamma_0(\mathbb{R})$ and

$$f\colon \mathbb{R} \to \left]-\infty,+\infty\right] \,:\, \xi \mapsto \begin{cases} -1-\ln(-\xi), & \text{if } \xi < -1;\\ \xi, & \text{if } \xi \geq -1, \end{cases} \tag{4.7}$$

from which we obtain

$$\widetilde{f}\colon \mathbb{R}\times\mathbb{R} \to \,]-\infty,+\infty] : (\xi,\eta) \mapsto \begin{cases} \eta - \eta\ln\left(-\frac{\xi}{\eta}\right), & \text{if } \eta>0 \text{ and } \xi<-\eta;\\ \xi, & \text{if } \eta>0 \text{ and } \xi\geq -\eta;\\ \max\{0,\xi\}, & \text{if } \eta=0;\\ +\infty, & \text{otherwise.}\end{cases} \tag{4.8}$$

Moreover, note that

$$f^* = \varphi^{**} = \psi + \iota_{]-\infty,1]}\colon \xi \mapsto \begin{cases} -\ln(\xi), & \text{if } 0<\xi\leq 1;\\ +\infty, & \text{otherwise}\end{cases} \tag{4.9}$$

and, thus, $\operatorname{dom} f^* = \,]0,1]$ which is neither open nor closed and $P_{\overline{\operatorname{dom}} f^*}\colon \xi\mapsto \operatorname{mid}\{0,\xi,1\}$. Now, in order to compute the proximity operator of $\widetilde{f}$, fix $\xi\in\mathbb{R}$, $\eta\in\mathbb{R}$, $\gamma\in\,]0,+\infty[$, and note that

$$\eta + \gamma f^*\left(P_{\overline{\operatorname{dom}} f^*}\left(\frac{\xi}{\gamma}\right)\right) = \eta - \gamma\ln\left(\operatorname{mid}\left\{0,\frac{\xi}{\gamma},1\right\}\right). \tag{4.10}$$

Therefore, by considering $\mathcal{H}=\mathbb{R}$, Theorem 3.1 yields

$$\operatorname{prox}_{\gamma\widetilde{f}}(\xi,\eta) = \begin{cases} \left(\xi - \gamma P_{\overline{\operatorname{dom}} f^*}\left(\frac{\xi}{\gamma}\right),0\right), & \text{if } \eta-\gamma\ln\left(\operatorname{mid}\left\{0,\frac{\xi}{\gamma},1\right\}\right)\leq 0;\\ \left(\xi - \gamma \operatorname{prox}_{\frac{\mu}{\gamma}f^*}\left(\frac{\xi}{\gamma}\right),\mu\right), & \text{if } \eta-\gamma\ln\left(\operatorname{mid}\left\{0,\frac{\xi}{\gamma},1\right\}\right)> 0.\end{cases} \tag{4.11}$$

where $\mu\in\,]0,\eta-\gamma\ln(\operatorname{mid}\{0,\xi/\gamma,1\})\,[$ is the unique solution to (3.2).

Note that, it follows from (4.9), $]-\infty,1]\cap\operatorname{dom}\psi =]0,1]\neq\varnothing$, and [2, Proposition 24.47] that $\operatorname{prox}_{\mu f^*/\gamma} = P_{]-\infty,1]}\circ\operatorname{prox}_{\mu\psi/\gamma}$. Moreover, [2, Example 24.40] implies that $\operatorname{prox}_{\mu\psi/\gamma}(\xi/\gamma) = (\xi+\sqrt{\xi^2+4\mu\gamma})/(2\gamma)$.

Observing that

$$\frac{\xi+\sqrt{\xi^2+4\mu\gamma}}{2\gamma}\geq 1 \quad\Leftrightarrow\quad \xi\geq\gamma-\mu, \tag{4.12}$$

we obtain

$$\operatorname{prox}_{\frac{\mu}{\gamma}f^*}\left(\frac{\xi}{\gamma}\right) = \begin{cases} \dfrac{\xi+\sqrt{\xi^2+4\mu\gamma}}{2\gamma}, & \text{if } \xi<\gamma-\mu;\\ 1, & \text{if } \xi\geq\gamma-\mu\end{cases} \tag{4.13}$$

and (3.2) reduces to

$$\mu = \begin{cases} \eta, & \text{if } \xi\geq\gamma-\eta;\\ \eta-\gamma\ln\left(\frac{\xi+\sqrt{\xi^2+4\mu\gamma}}{2\gamma}\right), & \text{if } \xi<\gamma-\eta.\end{cases} \tag{4.14}$$

Hence, since for all $(\xi,\eta)\in\mathcal{H}\times\mathbb{R}$, we have

$$\eta-\gamma\ln\left(\operatorname{mid}\left\{0,\frac{\xi}{\gamma},1\right\}\right)\leq 0 \quad\Leftrightarrow\quad \eta\leq 0 \quad\text{and}\quad \xi\geq\gamma e^{\eta/\gamma} \tag{4.15}$$

and

$$\begin{aligned}\eta-\gamma\ln\left(\operatorname{mid}\left\{0,\frac{\xi}{\gamma},1\right\}\right)> 0 \quad\Leftrightarrow\quad & (\eta>0 \text{ and } \xi\geq\gamma-\eta)\\ & \text{or } (\xi<\min\{\gamma e^{\eta/\gamma},\gamma-\eta\}),\end{aligned} \tag{4.16}$$

we deduce that (4.11) can be explicitly written as

$$\operatorname{prox}_{\gamma\widetilde{f}}(\xi,\eta)=\begin{cases}(\max\{0,\xi-\gamma\},0)\,, & \text{if } \eta\leq 0 \text{ and } \xi\geq \gamma e^{\frac{\eta}{\gamma}};\\ (\xi-\gamma,\eta), & \text{if } \eta>0 \text{ and } \xi\geq\gamma-\eta;\\ \left(\dfrac{\xi-\sqrt{\xi^2+4\mu\gamma}}{2},\mu\right), & \text{if } \xi<\min\left\{\gamma e^{\frac{\eta}{\gamma}},\gamma-\eta\right\},\end{cases}\tag{4.17}$$

where $\mu\in\,]0,\eta-\gamma\ln\left(\max\{0,\xi/\gamma\}\right)[$ is the unique solution to

$$\mu=\eta-\gamma\ln\left(\frac{\xi+\sqrt{\xi^2+4\mu\gamma}}{2\gamma}\right).\tag{4.18}$$

Example 4.3. Let $n\in\mathbb{N}$ and consider

$$f\colon\mathbb{R}^n\to\,]0,+\infty[\,\colon x\mapsto\sum_{i=1}^n e^{x_i-1}.\tag{4.19}$$

Then, $f\in\Gamma_0(\mathbb{R}^n)$ and

$$\widetilde{f}\colon\mathbb{R}^n\times\mathbb{R}\to\,]-\infty,+\infty]\colon(x,\eta)\mapsto\begin{cases}\eta\sum_{i=1}^n e^{\frac{x_i}{\eta}-1}, & \text{if } \eta>0;\\ 0, & \text{if } \eta=0 \text{ and } x\leq 0;\\ +\infty, & \text{otherwise.}\end{cases}\tag{4.20}$$

In order to compute the proximity operator of $\widetilde{f}$, note that, in view of [2, Example 13.2(v), Proposition 13.23 & Proposition 13.30], we obtain

$$f^*\colon\mathbb{R}^n\to\,]-\infty,+\infty]\colon x\mapsto\sum_{i=1}^n\phi(x_i),\tag{4.21}$$

where

$$\phi\colon\mathbb{R}\to\,]-\infty,+\infty]\colon\xi\mapsto\begin{cases}\xi\ln(\xi), & \text{if } \xi>0;\\ 0, & \text{if } \xi=0;\\ +\infty, & \text{if } \xi<0.\end{cases}\tag{4.22}$$

Hence, $\operatorname{dom} f^*=[0,+\infty[^n$ is closed and $P_{\overline{\operatorname{dom}} f^*}=\max\{0,\cdot\}$, where we denote $\max\{0,\cdot\}\colon y\mapsto(\max\{0,y_i\})_{1\leq i\leq n}$. Now fix $(x,\eta)\in\mathbb{R}^n\times\mathbb{R}$, $\gamma\in\,]0,+\infty[$, and note that

$$\eta+\gamma f^*\left(P_{\overline{\operatorname{dom}} f^*}\left(\frac{x}{\gamma}\right)\right)=\eta+\sum_{i\in I_+(x)}x_i\ln\left(\frac{x_i}{\gamma}\right),\tag{4.23}$$

where $I_+(x)=\{i\in\{1,\ldots,n\}\mid x_i>0\}$ and the sum over the empty set is zero. Therefore, by setting $\mathcal{H}=\mathbb{R}^n$, Theorem 3.1 yields

$$\operatorname{prox}_{\gamma\widetilde{f}}(x,\eta)=\begin{cases}\left(x-\gamma\max\left\{0,\dfrac{x}{\gamma}\right\},0\right), & \text{if } \eta+\sum\limits_{i\in I_+(x)}x_i\ln\left(\dfrac{x_i}{\gamma}\right)\leq 0;\\ \left(x-\gamma\operatorname{prox}_{\frac{\mu}{\gamma}f^*}\left(\dfrac{x}{\gamma}\right),\mu\right), & \text{if } \eta+\sum\limits_{i\in I_+(x)}x_i\ln\left(\dfrac{x_i}{\gamma}\right)>0,\end{cases}\tag{4.24}$$

where $\mu\in\,]0,\eta+\sum_{i\in I_+(x)}x_i\ln(x_i/\gamma)]$ is the unique solution to (3.2).

Furthermore, in view of [2, Proposition 16.9], we have $\mathrm{prox}_{\mu f^*/\gamma}(x/\gamma) \in \mathrm{dom}\,\partial(\mu f^*/\gamma) = \left]0,+\infty\right[^n$. Hence, we obtain from (2.8) that, for every $p \in \left]0,+\infty\right[^n$,

$$\begin{aligned}
p = \mathrm{prox}_{\frac{\mu}{\gamma}f^*}\left(\frac{x}{\gamma}\right) \quad &\Leftrightarrow \quad \frac{x}{\gamma} - p \in \frac{\mu}{\gamma}\partial f^*(p)\\
&\Leftrightarrow \quad (\forall i \in \{1,\ldots,n\}) \quad \frac{x_i}{\gamma} - p_i \in \frac{\mu}{\gamma}\partial\phi(p_i)\\
&\Leftrightarrow \quad (\forall i \in \{1,\ldots,n\}) \quad \frac{x_i}{\mu} - 1 = \ln(p_i) + \frac{\gamma}{\mu}p_i\\
&\Leftrightarrow \quad (\forall i \in \{1,\ldots,n\}) \quad \frac{\gamma}{\mu}e^{\frac{x_i}{\mu}-1} = \frac{\gamma}{\mu}p_i\, e^{\frac{\gamma}{\mu}p_i}\\
&\Leftrightarrow \quad (\forall i \in \{1,\ldots,n\}) \quad p_i = \frac{\mu}{\gamma}W_0\left(\frac{\gamma}{\mu}e^{\frac{x_i}{\mu}-1}\right),
\end{aligned} \tag{4.25}$$

where W_0 is the principal branch of the Lambert W-function. Altogether, by denoting $\min\{0,\cdot\}\colon y \mapsto -\max\{0,-y\}$ and, for every $y \in \mathbb{R}^n$, $e^y = (e^{y_i})_{1\le i\le n}$ and $W_0(y) = (W_0(y_i))_{1\le i\le n}$, (4.24) reduces to

$$\mathrm{prox}_{\gamma\widetilde{f}}(x,\eta) = \begin{cases} (\min\{0,x\},0), & \text{if } \eta + \sum\limits_{i\in I_+(x)} x_i\ln(x_i/\gamma) \le 0;\\ \left(x - \mu W_0\left(\frac{\gamma}{\mu}e^{\frac{x}{\mu}-1}\right),\mu\right), & \text{if } \eta + \sum\limits_{i\in I_+(x)} x_i\ln(x_i/\gamma) > 0, \end{cases} \tag{4.26}$$

where μ is the unique solution to (3.2), which reduces to

$$\mu = \eta + \mu\sum_{i=1}^n W_0\left(\frac{\gamma}{\mu}e^{\frac{x_i}{\mu}-1}\right)\ln\left(\frac{\mu}{\gamma}W_0\left(\frac{\gamma}{\mu}e^{\frac{x_i}{\mu}-1}\right)\right). \tag{4.27}$$

Since W_0 is continuous and strictly increasing in $\left[0,+\infty\right[$ [15], μ can be computed via standard one-dimensional root finding numerical schemes [24, Chapter 9].

Example 4.4. Let $n \in \mathbb{N}$ and let

$$f\colon \mathbb{R}^n \to \mathbb{R} : x \mapsto \ln\left(\sum_{i=1}^n e^{x_i}\right). \tag{4.28}$$

Then $f \in \Gamma_0(\mathbb{R}^n)$ and

$$\widetilde{f}\colon \mathbb{R}^n \times \mathbb{R} \to \left]-\infty,+\infty\right] : (x,\eta) \mapsto \begin{cases} \eta\ln\left(\sum_{i=1}^n e^{\frac{x_i}{\eta}}\right), & \text{if } \eta > 0;\\ \max_{1\le i\le n} x_i, & \text{if } \eta = 0;\\ +\infty, & \text{if } \eta < 0, \end{cases} \tag{4.29}$$

which appears naturally in the dual of entropy-penalized transport problems [10].

In order to compute the proximity operator of $\widetilde{f}$, set ϕ as in (4.22) and let

$$\triangle = \left\{x \in \left[0,+\infty\right[^n \;\middle|\; \sum_{i=1}^n x_i = 1\right\} \tag{4.30}$$

be the probability simplex in $\mathbb{R}^n$. Then [6, Example 3.25] yields

$$f^*\colon \mathbb{R}^n \to \left]-\infty,+\infty\right] : x \mapsto \begin{cases} \sum_{i=1}^n \phi(x_i), & \text{if } \sum_{i=1}^n x_i = 1;\\ +\infty, & \text{otherwise}, \end{cases} \tag{4.31}$$

and hence, $\operatorname{dom} f^* = \triangle$ is closed. Now, fix $x \in \mathbb{R}^n$, $\eta \in \mathbb{R}$, and $\gamma \in \left]0, +\infty\right[$. By setting $\mathcal{H} = \mathbb{R}^n$, Theorem 3.1 implies

$$\operatorname{prox}_{\gamma \widetilde{f}}(x, \eta) = \begin{cases} \left(x - \gamma P_\triangle\left(\frac{x}{\gamma}\right), 0\right), & \text{if } \eta + \gamma f^*\left(P_\triangle\left(\frac{x}{\gamma}\right)\right) \leq 0, \\ \left(x - \gamma \operatorname{prox}_{\frac{\mu}{\gamma} f^*}\left(\frac{x}{\gamma}\right), \mu\right), & \text{if } \eta + \gamma f^*\left(P_\triangle\left(\frac{x}{\gamma}\right)\right) > 0, \end{cases} \tag{4.32}$$

where $\mu \in \left]0, \eta + \gamma f^*\left(P_\triangle(x/\gamma)\right)\right]$ is the unique solution to (3.2). Moreover, since [2, Proposition 16.9] yield $\operatorname{prox}_{\mu f^*/\gamma}(x/\gamma) \in \operatorname{dom} \partial(\mu f^*/\gamma) = \triangle \cap \left]0, +\infty\right[^n$, we obtain from (2.8) that, for every $p \in \triangle \cap \left]0, +\infty\right[^n$,

$$\begin{aligned} p = \operatorname{prox}_{\frac{\mu}{\gamma} f^*}\left(\frac{x}{\gamma}\right) \;&\Leftrightarrow\; \frac{x}{\gamma} - p \in \frac{\mu}{\gamma} \partial f^*(p) \\ &\Leftrightarrow\; (\exists \lambda \in \mathbb{R})(\forall i \in \{1, \ldots, n\}) \quad \frac{x_i - \gamma p_i}{\mu} = \ln(p_i) + 1 + \lambda \\ &\Leftrightarrow\; (\exists \lambda \in \mathbb{R})(\forall i \in \{1, \ldots, n\}) \quad \frac{x_i}{\mu} - 1 - \lambda = \ln(p_i) + \frac{\gamma}{\mu} p_i \\ &\Leftrightarrow\; (\exists \lambda \in \mathbb{R})(\forall i \in \{1, \ldots, n\}) \quad \frac{\gamma}{\mu} e^{\frac{x_i}{\mu} - 1 - \lambda} = \frac{\gamma}{\mu} p_i \, e^{\frac{\gamma}{\mu} p_i} \\ &\Leftrightarrow\; (\exists \lambda \in \mathbb{R})(\forall i \in \{1, \ldots, n\}) \quad p_i = \frac{\mu}{\gamma} W_0\left(\frac{\gamma}{\mu} e^{\frac{x_i}{\mu} - 1 - \lambda}\right), \end{aligned} \tag{4.33}$$

where W_0 is the principal branch of the Lambert W-function. Note that by summing over i in (4.33) we obtain

$$\frac{\mu}{\gamma} \sum_{i=1}^{n} W_0\left(\frac{\gamma}{\mu} e^{\frac{x_i}{\mu} - 1 - \lambda}\right) = 1. \tag{4.34}$$

Altogether, (4.32) reduces to

$$\operatorname{prox}_{\gamma \widetilde{f}}(x, \eta) = \begin{cases} \left(x - \gamma P_\triangle\left(\frac{x}{\gamma}\right), 0\right), & \text{if } \eta + \gamma f^*\left(P_\triangle\left(\frac{x}{\gamma}\right)\right) \leq 0, \\ \left(x - \mu W_0\left(\frac{\gamma}{\mu} e^{\frac{x}{\mu} - \Lambda}\right), \mu\right), & \text{if } \eta + \gamma f^*\left(P_\triangle\left(\frac{x}{\gamma}\right)\right) > 0, \end{cases} \tag{4.35}$$

where we denote $\Lambda = (\lambda + 1)_{1 \leq i \leq n}$, and μ and λ are the solution to the nonlinear system of equations

$$\begin{cases} \mu &= \eta + \mu \sum_{i=1}^{n} W_0\left(\frac{\gamma}{\mu} e^{\frac{x_i}{\mu} - 1 - \lambda}\right) \ln\left(\frac{\mu}{\gamma} W_0\left(\frac{\gamma}{\mu} e^{\frac{x_i}{\mu} - 1 - \lambda}\right)\right); \\ 1 &= \frac{\mu}{\gamma} \sum_{i=1}^{n} W_0\left(\frac{\gamma}{\mu} e^{\frac{x_i}{\mu} - 1 - \lambda}\right), \end{cases} \tag{4.36}$$

which can be solved, for instance, by Newton's type methods.

5 Conclusions

In summary, we provide an explicit formula for the proximity operator of the perspective function of any proper lower semicontinuous convex function defined in real Hilbert spaces. The formula needs to solve a scalar nonlinear equation which can be efficiently solved by several one-dimensional root-finding numerical schemes. Our result generalizes [13] and [14], valid only under additional assumptions. Finally, we derive several new formulae of proximity operators of perspective functions arising in penalization of mathematical programming problems appearing, e.g., in entropy-penalized optimal transport problems.

Acknowledgments. The work of Luis M. Briceño-Arias is supported by Centro de Modelamiento Matemático (CMM), FB210005, BASAL fund for centers of excellence, FONDECYT 1190871, and FONDECYT 1230257 from ANID-Chile.

References

[1] Auslender, A., Cominetti, R., Haddou, M.: Asymptotic analysis for penalty and barrier methods in convex and linear programming. Math. Oper. Res. **22**(1), 43–62 (1997). https://doi.org/10.1287/moor.22.1.43

[2] Bauschke, H.H., Combettes, P.L.: Convex analysis and monotone operator theory in Hilbert spaces. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-48311-5

[3] Benamou, J.D., Brenier, Y.: A computational fluid mechanics solution to the monge-kantorovich mass transfer problem. Numer. Math. **84**(3), 375–393 (2000). https://doi.org/10.1007/s002110050002

[4] Benamou, J.D., Carlier, G., Laborde, M.: An augmented lagrangian approach to wasserstein gradient flows and applications. ESAIM Proc. Surveys **54**, 1–17 (2016). https://doi.org/10.1051/proc/201654001

[5] Bercher, J.F.: Some properties of generalized fisher information in the context of nonextensive thermostatistics. Phys. A **392**(15), 3140–3154 (2013). https://doi.org/10.1016/j.physa.2013.03.062

[6] Boyd, S., Vandenberghe, L.: Convex optimization. Cambridge university press (2004). https://doi.org/10.1017/CBO9780511804441

[7] Briceño-Arias, L.M., Combettes, P.L., Silva, F.J.: Proximity operators of perspective functions with nonlinear scaling. SIAM J. Optim. (2023). To appear, https://arxiv.org/pdf/2303.05337.pdf

[8] Briceño-Arias, L.M., Kalise, D., Silva, F.J.: Proximal methods for stationary mean field games with local couplings. SIAM J. Control Optim. **56**(2), 801–836 (2018). https://doi.org/10.1137/16M1095615

[9] Cardaliaguet, P.: Weak solutions for first order mean field games with local coupling. In: P. Bettiol, P. Cannarsa, G. Colombo, M. Motta, F. Rampazzo (eds.) Analysis and geometry in control theory and its applications, pp. 111–158. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-06917-3_5

[10] Carlier, G., Pegon, P., Tamanini, L.: Convergence rate of general entropic optimal transport costs. Calc. Var. Partial Differential Equations **62**(4), Art. 116, 28 pp. (2023). https://doi.org/10.1007/s00526-023-02455-0

[11] Carrillo, J.A., Craig, K., Wang, L., Wei, C.: Primal dual methods for wasserstein gradient flows. Found. Comput. Math. **22**(2), 389–443 (2022). https://doi.org/10.1007/s10208-021-09503-1

[12] Combettes, P.L.: Perspective functions: Properties, constructions, and examples. Set-Valued Var. Anal. **26**(2), 247–264 (2018). https://doi.org/10.1007/s11228-017-0407-x

[13] Combettes, P.L., Müller, C.L.: Perspective functions: Proximal calculus and applications in high-dimensional statistics. J. Math. Anal. Appl. **457**(2), 1283–1306 (2018). https://doi.org/10.1016/j.jmaa.2016.12.021

[14] Combettes, P.L., Müller, C.L.: Perspective maximum likelihood-type estimation via proximal decomposition. Electron. J. Stat. **14**(1), 207–238 (2020). https://doi.org/10.1214/19-EJS1662

[15] Corless, R.M., H., G.G., Hare, D.E.G., Jeffrey, D.J., Knuth, D.E.: On the lambert W function. Adv. Comput. Math. **5**(4), 329–359 (1996). https://doi.org/10.1007/BF02124750

[16] Dacorogna, B., Maréchal, P.: The role of perspective functions in convexity, polyconvexity, rank-one convexity and separate convexity. J. Convex Anal. **15**(2), 271–284 (2008)

[17] Effros, E.G.: A matrix convexity approach to some celebrated quantum inequalities. Proc. Natl. Acad. Sci. USA **106**(4), 1006–1008 (2009). https://doi.org/10.1073/pnas.080796510

[18] El Gheche, M., Chierchia, G., Pesquet, J.C.: Proximity operators of discrete information divergences. IEEE Trans. Inform. Theory **64**(2), 1092–1104 (2018). https://doi.org/10.1109/TIT.2017.2782789

[19] Jordan, R., Kinderlehrer, D., Otto, F.: The variational formulation of the Fokker–Planck equation. SIAM J. Math. Anal. **29**(1), 1–17 (1998). https://doi.org/10.1137/S0036141096303359

[20] Kuroda, H., Kitahara, D.: Block-sparse recovery with optimal block partition. IEEE Trans. Signal Process. **70**, 1506–1520 (2022). https://doi.org/10.1109/TSP.2022.3156283

[21] Kuroda, H., Kitahara, D., Hirabayashi, A.: A convex penalty for block-sparse signals with unknown structures. In: ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5430–5434. IEEE (2021). https://doi.org/10.1109/ICASSP39728.2021.9414340

[22] Micchelli, C.A., Morales, J.M., Pontil, M.: Regularizers for structured sparsity. Adv. Comput. Math. **38**(3), 455–489 (2013). https://doi.org/10.1007/s10444-011-9245-9

[23] Owen, A.B.: A robust hybrid of lasso and ridge regression. In: J.S. Verducci, X. Shen, J. Lafferty (eds.) Prediction and discovery, vol. 443, pp. 59–71. Amer. Math. Soc., Providence, RI (2007). https://doi.org/10.1090/conm/443/08555

[24] Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical recipes 3rd edition: The art of scientific computing. Cambridge University Press, Cambridge (2007)

[25] Rockafellar, R.T.: Level sets and continuity of conjugate convex functions. Trans. Amer. Math. Soc. **123**, 46–63 (1966). https://doi.org/10.1090/S0002-9947-1966-0192318-X

[26] Villani, C.: Topics in optimal transportation. American Mathematical Society, Providence, RI (2003). http://dx.doi.org/10.1090/gsm/058