OJMO

Open Journal of Mathematical Optimization

Thu-Le Tran, Clément Elvira, Hong-Phuong Dang & Cédric Herzet

One to beat them all: "RYU" – a unifying framework for the construction of safe balls

Volume 6 (2025), article no. 7 (16 pages)

https://doi.org/10.5802/ojmo.45

Article submitted on December 1, 2024, revised on June 11, 2025,
accepted on August 29, 2025.

© The author(s), 2025.

This article is licensed under the
Creative Commons Attribution 4.0 International License.
http://creativecommons.org/licenses/by/4.0/

Open Journal of Mathematical Optimization is a member of the
Centre Mersenne for Open Scientific Publishing
http://www.centre-mersenne.org/
e-ISSN: 2777-5860

One to beat them all: "RYU" – a unifying framework for the construction of safe balls

Thu-Le Tran
Univ. Rennes, IRMAR - UMR 6625, F-35000 Rennes, France
Mathematic department, School of Education, Can Tho Univ, Vietnam
ttle@ctu.edu.vn

Clément Elvira
IETR UMR CNRS 6164, CentraleSupelec Rennes Campus, 35576 Cesson Sévigné, France
clement.elvira@centralesupelec.fr

Hong-Phuong Dang
IETR UMR CNRS 6164, CentraleSupelec Rennes Campus, 35576 Cesson Sévigné, France
hong-phuong.dang@centralesupelec.fr

Cédric Herzet
Univ. Rennes, Ensai, CNRS, CREST - UMR 9194, F-35000 Rennes, France
cedric.herzet@ensai.fr

Abstract

In this paper, we present a new framework, called "RYU", for constructing "safe" regions – specifically, sets that are guaranteed to contain the dual solution of a target optimization problem. Our framework applies to the standard case where the objective function is composed of two components: a closed, proper, convex function with Lipschitz-smooth gradient and another closed, proper, convex function. We show that the RYU framework not only encompasses but also improves upon the state-of-the-art methods proposed over the past decade for this class of optimization problems.

Digital Object Identifier 10.5802/ojmo.45

Keywords Convex optimization, screening rules, safe regions.

Acknowledgments This research is funded by Ministry of Education and Training in Viet Nam under grant number B2025-TCT-02.

1 Introduction

1.1 Context and state of the art

In this paper, we consider the following family of optimization problems:

$$\text{find } \mathbf{x}^\star \in \operatorname*{arg\,min}_{\mathbf{x}\in\mathbf{R}^n} \; P(\mathbf{x}) \triangleq f(\mathbf{A}\mathbf{x}) + g(\mathbf{x}) \tag{P}$$

where $\mathbf{A} \in \mathbf{R}^{m\times n}$ is some known matrix, $f\colon \mathbf{R}^m \to \mathbf{R}$, $g\colon \mathbf{R}^n \to \mathbf{R}\cup\{+\infty\}$ are proper, closed, convex functions and f is "α^{-1}-Lipschitz smooth over $\mathbf{R}^m$", that is f is differentiable everywhere on $\mathbf{R}^m$ and its gradient obeys the following regularity condition for some positive scalar $\alpha > 0$:

$$\forall\, \mathbf{z}, \mathbf{z}' \in \mathbf{R}^m : \; \|\nabla f(\mathbf{z}) - \nabla f(\mathbf{z}')\|_2 \le \alpha^{-1}\|\mathbf{z}-\mathbf{z}'\|_2. \tag{1}$$

We assume moreover that (P) is well-posed in the sense that at least one minimizer $\mathbf{x}^\star$ exists. Instances of problems satisfying these hypotheses are common in the literature of machine learning, statistics or signal processing, and include (among many others) least-squares sparse regression [4], logistic sparse regression [16] or the "Elastic Net" problem [27].

The focus of this paper is on the construction of *"safe regions"*, i.e., subsets of $\mathbf{R}^m$ provably containing the unique solution of the dual problem of (P). More specifically, our goal is to identify some subset $\mathcal{S} \subseteq \mathbf{R}^m$ such that $\mathbf{u}^\star \in \mathcal{S}$ where

$$\mathbf{u}^\star = \operatorname*{arg\,max}_{\mathbf{u}\in\mathbf{R}^m} \; D(\mathbf{u}) \triangleq -f^*(-\mathbf{u}) - g^*(\mathbf{A}^{\mathrm{T}}\mathbf{u}) \tag{D}$$

© Thu-Le Tran & Clément Elvira & Hong-Phuong Dang & Cédric Herzet;
licensed under Creative Commons License Attribution 4.0 International

and f^*, g^* denote the convex conjugates of f, g, respectively.

The construction of safe regions has become an active field of research during the last decade (see e.g., [5, 10, 14, 15, 18, 21, 22, 23, 24, 25, 26]) and has been triggered by the so-called "safe feature elimination" technique (also referred to as "safe screening"), a procedure to accelerate the resolution of (P), first proposed in [6] and further extended in many contributions, see e.g., [7, 8, 12, 19]. One central element in the effectiveness of these acceleration methods is the identification (preferably at low computational cost) of *small* safe regions with some specific geometry (e.g., ball, ellipsoid, dome, etc). In this paper, we focus on safe regions having a "ball" geometry, that is

$$\mathcal{S} = \mathcal{B}(\mathbf{c}, r) \triangleq \{\mathbf{u} \in \mathbf{R}^m \mid \|\mathbf{u} - \mathbf{c}\|_2 \leq r\} \tag{2}$$

for some $\mathbf{c} \in \mathbf{R}^m$ and $r > 0$. In this respect, the state-of-the-art safe ball for the general family of optimization problems considered in this paper is indubitably the so-called "GAP ball" proposed in [10, 18]. It is defined for any couple $(\mathbf{x}, \mathbf{u}) \in \mathrm{dom}(P) \times \mathrm{dom}(-D)$ as

$$\mathcal{B}_{\texttt{GAP}}(\mathbf{x}, \mathbf{u}) \triangleq \mathcal{B}(\mathbf{c}_{\texttt{GAP}}, r_{\texttt{GAP}}) \tag{3}$$

where

$$\mathbf{c}_{\texttt{GAP}} \triangleq \mathbf{u} \tag{4a}$$

$$r_{\texttt{GAP}} \triangleq \sqrt{\frac{2\texttt{GAP}(\mathbf{x}, \mathbf{u})}{\alpha}} \tag{4b}$$

and $\texttt{GAP}(\mathbf{x}, \mathbf{u}) \triangleq P(\mathbf{x}) - D(\mathbf{u})$ is the so-called duality gap. The popularity of the GAP ball is due to the two following assets:

1. The construction of the ball is valid for any problem satisfying our blanket hypotheses, that is f, g are proper, closed, convex and f is α^{-1}-Lipschitz smooth over $\mathbf{R}^m$.
2. A GAP ball can be constructed from any primal-dual feasible couple $(\mathbf{x}, \mathbf{u})$. In particular, under strong duality assumption and continuity of the duality gap over its domain, the radius of the ball can be made arbitrarily small by choosing $(\mathbf{x}, \mathbf{u})$ sufficiently close to some primal-dual solution $(\mathbf{x}^\star, \mathbf{u}^\star)$.

These features have led the GAP ball to be widely applied and to allow for substantial acceleration performance in many setups, see e.g., [7, 9, 11, 12, 13].

Other constructions of safe balls, requiring either additional hypotheses on f and g or the knowledge of some specific primal-dual couple $(\mathbf{x}, \mathbf{u})$, have also been proposed in the literature. All the safe ball constructions (to the best of our knowledge) falling into the optimization framework considered in this paper are gathered in Table 1 and will be reviewed in greater details in Section 4. We note that, although requiring additional assumptions, some of these works put to the forth that the construction of safe balls smaller than the GAP ball is possible. In particular, some contributions have highlighted two avenues for improvement. In [15], the authors proposed a safe ball (referred to as "FNE ball") for the LASSO problem and proved that it is a (potentially strict) *subset* of the GAP ball. More recently, the authors of [26] introduced the so-called "dynamic EDPP ball" and emphasized that the latter has *a smaller radius* than the GAP ball constructed with the same primal-dual pair [26, Thm. 10].

In this paper, we provide a new mathematical framework gathering and extending these results to the general family of optimization problems (P).

1.2 Contributions

The contribution of this paper is two-fold. We first introduce a new safe[1] ball referred to as "RYU ball". This ball is defined $\forall\ (\mathbf{x}, \mathbf{u}) \in \mathrm{dom}(P) \times \mathrm{dom}(-D)$ as follows:

$$\mathcal{B}_{\texttt{RYU}}(\mathbf{x}, \mathbf{u}) \triangleq \mathcal{B}(\mathbf{c}_{\texttt{RYU}}, r_{\texttt{RYU}}) \tag{5}$$

where

$$\mathbf{c}_{\texttt{RYU}} \triangleq \frac{1}{2}(\mathbf{u} - \nabla f(\mathbf{A}\mathbf{x})) \tag{6a}$$

$$r_{\texttt{RYU}} \triangleq \sqrt{\frac{\texttt{GAP}(\mathbf{x}, \mathbf{u})}{\alpha} - \frac{\|\mathbf{u} + \nabla f(\mathbf{A}\mathbf{x})\|_2^2}{4}}, \tag{6b}$$

[1] As defined in Section 1.1, a safe region is any subset of $\mathbf{R}^m$ that contains the unique dual solution $\mathbf{u}^\star$ of (D).

Table 1 Summary of the main safe-ball constructions proposed in the literature during the last decade. The first column provides the name and the references associated to the safe ball, the second describes its connection with the proposed RYU ball, the third indicates the constraints on the primal-dual couple $(\mathbf{x}, \mathbf{u})$ used in the construction. The last column specifies the setup considered by the authors in their work.

Safe region	Relation	Cstr. on $(\mathbf{x}, \mathbf{u})$	Hyp. on f and g
GAP [10, 18, 19]	$\supseteq \mathcal{B}_{\text{RYU}}(\mathbf{x}, \mathbf{u})$	feasible	(H1)–(H2)
$\mathbf{x}$-GAP [14]	$\supseteq \mathcal{B}_{\text{RYU}}(\mathbf{x}, \mathbf{u})$	feasible	(H1)–(H2), $g = \lambda\|\cdot\|_1$
Dyn. EDPP [26]	$= \mathcal{B}_{\text{RYU}}(t^\star \mathbf{x}, \mathbf{u})$	feasible	$f = \frac{1}{2}\|\mathbf{y} - \cdot\|_2^2$, $g =$ gauge
FNE [15]	$= \mathcal{B}_{\text{RYU}}(\mathbf{x}, \mathbf{u})$	$\mathbf{A}^{\mathrm{T}}\mathbf{u} \in \partial g(\mathbf{x})$	$f = \frac{1}{2}\|\mathbf{y} - \cdot\|_2^2$, $g = \lambda\|\cdot\|_1$
SASVI [17]	$= \mathcal{B}_{\text{RYU}}(\mathbf{0}_n, \mathbf{u})$	$\mathbf{u} \in \text{dom}(-D)$	$f = \frac{1}{2}\|\mathbf{y} - \cdot\|_2^2$, $g = \lambda\|\cdot\|_1$
EDPP [24]	$= \mathcal{B}_{\text{RYU}}(\mathbf{x}, \mathbf{u})$	(18)–(19)	$f = \frac{1}{2}\|\mathbf{y} - \cdot\|_2^2$, $g = \lambda\|\cdot\|_1$
DPP [24]	$\supseteq \mathcal{B}_{\text{RYU}}(\mathbf{x}, \mathbf{u})$	(18)–(19)	$f = \frac{1}{2}\|\mathbf{y} - \cdot\|_2^2$, $g = \lambda\|\cdot\|_1$
SAFE [6]	$\supseteq \mathcal{B}_{\text{RYU}}(\mathbf{0}_n, \mathbf{u})$	$\mathbf{u} \in \text{dom}(-D)$	$f = \frac{1}{2}\|\mathbf{y} - \cdot\|_2^2$, $g = \lambda\|\cdot\|_1$
SLORES [25]	$\supseteq \mathcal{B}_{\text{RYU}}(\mathbf{x}, \mathbf{u})$	(18)–(19)	$f =$ logistic, $g = \lambda\|\cdot\|_1$
SFER [21]	$= \mathcal{B}_{\text{RYU}}(\mathbf{x}, \mathbf{u})$	(18)–(19)	$f =$ logistic, $g = \lambda\|\cdot\|_1$

see Theorem 2.[2] The name "RYU" stems from "**R**efined Fenchel-**Y**oung ineq**u**ality" and refers to the fact that the safeness of the ball is a consequence of the (double) application of a refined version of the well-known Fenchel-Young inequality (see Appendix B). Our ball construction is valid under the same generic assumptions as the GAP ball (see Section 3 for a detailed discussion about our working hypotheses). In particular: *i)* it can be applied to any problem (P) involving a proper, closed, convex function f which is α^{-1}-Lipschitz smooth over $\mathbf{R}^m$, and a proper, closed, convex function g; *ii)* primal-dual feasibility is the only assumption required for the pair $(\mathbf{x}, \mathbf{u})$.

Second, we show that our safe ball construction generalizes or improves over all the existing results of the literature. More specifically, we prove that all the existing safe balls correspond to particular cases or supersets of the proposed ball. These results are summarized in the second column of Table 1 and correspond to Propositions 5 to 10 of the paper. As a byproduct, our analysis also provides a unified review of safe balls for problem (D) under hypotheses (H1)–(H2) by connecting existing results in a common framework.

Interestingly, we note that the GAP ball is always a superset of the RYU ball, that is:

$$\mathcal{B}_{\text{RYU}}(\mathbf{x}, \mathbf{u}) \subseteq \mathcal{B}_{\text{GAP}}(\mathbf{x}, \mathbf{u}), \tag{7}$$

where the inclusion is shown to be strict as long as $(\mathbf{x}, \mathbf{u})$ is not a primal-dual optimal couple, see Theorem 5. Moreover, a rapid inspection of (4b) and (6b) shows that the squared radius of the RYU ball is never greater than half the squared radius of the GAP ball.

Since our construction is valid for any feasible[3] couple $(\mathbf{x}, \mathbf{u})$ and holds under general assumptions on f, g, the results in Table 1 therefore emphasize that the proposed framework unifies and generalizes all the methodologies previously proposed in the literature.

1.3 Paper organization

The rest of the paper is organized as follows. In Section 2, we detail the notational conventions used in the paper. In Section 3, we describe the working hypotheses considered in our derivations and discuss some of their implications. Section 4 is dedicated to the presentation of our new safe ball and its connection with the previous results of the literature. Most of the technical details are deferred to Appendices A to C.

2 Notations

Unless mentioned explicitly, we will use the following notational conventions throughout the paper. Vectors are denoted by lowercase bold letters (e.g., $\mathbf{x}, \mathbf{z}$) and matrices by uppercase bold letters (e.g., $\mathbf{A}$). We use the

[2] Note that the quantity under the square root in (6b) is necessarily nonnegative (see Theorem 14).

[3] Strictly speaking, the proposed construction applies to any couple $(\mathbf{x}, \mathbf{u}) \in \mathbf{R}^n \times \mathbf{R}^m$ but (similarly to the GAP ball) it leads to a ball with infinite radius when $(\mathbf{x}, \mathbf{u}) \notin \text{dom}(P) \times \text{dom}(-D)$. This is the reason why we restrict our construction to $\text{dom}(P) \times \text{dom}(-D)$ in the paper.

symbol "$^{\mathrm{T}}$" to denote the transpose of a vector or a matrix. The "all-zero" vector of dimension n is written $\mathbf{0}_n$. $\langle \mathbf{z} \mid \mathbf{z}' \rangle$ denotes the standard inner product between $\mathbf{z}$ and $\mathbf{z}'$. We use the notation x_j to refer to the jth entry of a vector $\mathbf{x}$. For matrices, we use $\mathbf{a}_j$ to denote the jth column of $\mathbf{A}$. We let $\overline{\mathbf{R}} = \mathbf{R} \cup \{-\infty, +\infty\}$ where $\mathbf{R}$ refers to the set of real numbers. Given an extended real-valued function $h \colon \mathbf{R}^d \to \overline{\mathbf{R}}$, we let

$$\mathrm{dom}(h) \triangleq \left\{\mathbf{z} \in \mathbf{R}^d \,\middle|\, h(\mathbf{z}) < +\infty\right\}.$$

The subdifferential of $h \colon \mathbf{R}^d \to \overline{\mathbf{R}}$ at $\mathbf{z} \in \mathbf{R}^d$ is denoted $\partial h(\mathbf{z})$. It is defined for any $\mathbf{z} \in \mathrm{dom}(h)$ as

$$\partial h(\mathbf{z}) \triangleq \left\{\mathbf{g} \in \mathbf{R}^d \,\middle|\, \forall\, \mathbf{z}' \in \mathbf{R}^d :\; h(\mathbf{z}') \geq h(\mathbf{z}) + \langle \mathbf{g} \mid \mathbf{z}' - \mathbf{z} \rangle\right\}.$$

We refer to the elements of $\partial h(\mathbf{z})$ as "subgradients", see [2, Def. 3.2]. Finally,

$$\begin{aligned} h^* \colon \mathbf{R}^d &\longrightarrow \overline{\mathbf{R}} \\ \mathbf{z}^* &\longmapsto \sup_{\mathbf{z} \in \mathbf{R}^d} \langle \mathbf{z}^*, \mathbf{z} \rangle - h(\mathbf{z}), \end{aligned}$$

denotes the convex conjugate of h, see [2, Def. 4.1].

3 Optimization framework

In this paper, we consider problem (P) with the following minimal assumptions:

(H1) f and g are proper, closed and convex functions.
(H2) f is α^{-1}-Lipschitz smooth over $\mathbf{R}^m$.

We note that (H1) and (H2) correspond to the general hypotheses involved in the construction of the GAP ball, see [19]. In the rest of this section we elaborate on some properties of problems (P)–(D) induced by these hypotheses.

First, since f (resp. g) is proper, closed and convex from (H1), its convex conjugate f^* (resp. g^*) is proper, closed and convex, see [2, Thms. 4.5 and 4.13]. Moreover, the convexity and α^{-1}-Lipschitz smoothness over $\mathbf{R}^m$ of f in (H2) imply that f^* is α-strongly convex, see [2, Thm. 5.26], that is:

$$\forall\, \mathbf{z}, \mathbf{z}' \in \mathrm{dom}(f^*) \text{ and } \mathbf{g} \in \partial f^*(\mathbf{z}) :\; f^*(\mathbf{z}') \geq f^*(\mathbf{z}) + \langle \mathbf{g} \mid \mathbf{z}' - \mathbf{z} \rangle + \tfrac{\alpha}{2} \|\mathbf{z}' - \mathbf{z}\|_2^2. \tag{8}$$

Under the properness assumption in (H1), the duality gap, defined as

$$\begin{aligned} \mathtt{GAP} \colon \mathbf{R}^n \times \mathbf{R}^m &\longrightarrow \overline{\mathbf{R}} \\ (\mathbf{x}, \mathbf{u}) &\longmapsto P(\mathbf{x}) - D(\mathbf{u}) \end{aligned}$$

is always a nonnegative quantity, see [1, Prop. 15.21.(i)]. Moreover, $\mathtt{GAP}(\mathbf{x}, \mathbf{u}) < +\infty$ if and only if $(\mathbf{x}, \mathbf{u}) \in \mathrm{dom}(P) \times \mathrm{dom}(-D)$. Hypotheses (H1)–(H2) also imply that strong duality holds for some primal-dual couple as emphasized by the following lemma:

▶ **Lemma 1.** *Let $\mathbf{x}^\star$ be a minimizer of* (P)*. If* (H1)–(H2) *hold, then there exists $\mathbf{u}^\star \in \mathbf{R}^m$ such that*

$$\mathtt{GAP}(\mathbf{x}^\star, \mathbf{u}^\star) = 0. \tag{9}$$

Proof. If (H1) is verified, we have from [1, Thms. 15.23 and 15.24.(viii)] that strong duality holds and a maximizer $\mathbf{u}^\star$ to (D) exists provided that

$$\mathrm{relint}(\mathrm{dom}(f)) \cap \mathrm{relint}(\mathbf{A}\,\mathrm{dom}(g)) \neq \emptyset, \tag{10}$$

where $\mathrm{relint}(\,\cdot\,)$ denotes the relative interior of a set. Now, under our α^{-1}-Lipschitz smoothness assumption (H2), we have that $\mathrm{dom}(f) = \mathbf{R}^m$ so that condition (10) reduces to $\mathrm{relint}(\mathbf{A}\,\mathrm{dom}(g)) \neq \emptyset$. Since g is a proper convex function, its domain $\mathrm{dom}(g)$ is non-empty (by definition of properness) and convex [2, §2.3.1]. The set $\mathbf{A}\,\mathrm{dom}(g)$ is thus also non-empty convex as the image of a non-empty convex set under a linear operator $\mathbf{A}$ [1, Prop. 3.5]. Therefore, the relative interior of $\mathbf{A}\,\mathrm{dom}(g)$ is non-empty as a consequence of [2, Thm. 3.17]. ◀

We note that any $\mathbf{u}^\star$ verifying (9) must obviously be a solution of (D). Hence, we have from Theorem 1 that a maximizer of (D) exists under (H1)–(H2). Moreover, the α-strong convexity of f^* implies that this maximizer is unique, see [2, Thm. 5.25]. Finally, since strong duality holds, the following conditions must be satisfied by any primal-dual optimal couple $(\mathbf{x}^\star, \mathbf{u}^\star)$, see [1, Thm. 19.1]:

$$\mathbf{u}^\star = -\nabla f(\mathbf{A}\mathbf{x}^\star) \tag{11}$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{u}^\star \in \partial g(\mathbf{x}^\star). \tag{12}$$

4 The RYU framework and its connection to the state of the art

The main theoretical result of this paper is the following new safe ball:

▶ **Theorem 2** (RYU ball). *Assume* (H1)–(H2) *hold true. Then, we have for any* $(\mathbf{x}, \mathbf{u}) \in \mathrm{dom}(P) \times \mathrm{dom}(-D)$*:*

$$\mathbf{u}^\star \in \mathcal{B}_{\texttt{RYU}}(\mathbf{x}, \mathbf{u}) \triangleq \mathcal{B}(\mathbf{c}_{\texttt{RYU}}, r_{\texttt{RYU}})$$

where

$$\mathbf{c}_{\texttt{RYU}} \triangleq \frac{1}{2}(\mathbf{u} - \nabla f(\mathbf{A}\mathbf{x}))$$
$$r_{\texttt{RYU}} \triangleq \sqrt{\frac{\texttt{GAP}(\mathbf{x}, \mathbf{u})}{\alpha} - \frac{\|\mathbf{u} + \nabla f(\mathbf{A}\mathbf{x})\|_2^2}{4}}.$$

The name "RYU" stands for "**R**efined Fenchel-**Y**oung ineq**u**ality", a central element appearing in the proof of the safeness of the proposed region, detailed in Appendix B.1.

A close inspection of the hypotheses of Theorem 2 reveals that the construction of the RYU ball can be carried out under exactly the same assumptions as the GAP ball, that is: *i)* it holds for any problem satisfying hypotheses (H1)–(H2) on f and g; *ii)* it is valid for any primal-dual feasible couple $(\mathbf{x}, \mathbf{u})$. Despite of its generality, a careful examination of the definition of the radius of the GAP and RYU balls in (4b) and (6b) indicates that – given a feasible primal-dual pair $(\mathbf{x}, \mathbf{u})$ – the squared radius of the RYU ball is *always* at least twice as small as the squared radius of the GAP ball. In fact, as emphasized in Section 4.1 below, the GAP ball is a strict superset of the proposed RYU ball for any feasible primal-dual $(\mathbf{x}, \mathbf{u})$ different from $(\mathbf{x}^\star, \mathbf{u}^\star)$.

In the rest of this section, we explain how the safe balls previously proposed in the literature relate to the RYU ball. More specifically, we emphasize that the previous results of the state of the art can be seen as either particular cases or supersets of the proposed ball. Our results are contained in Propositions 5 to 10 and are summarized in the second column of Table 1. The third and fourth columns of the table specify assumptions necessary for constructing the corresponding safe ball: the third column outlines potential constraints on the primal-dual pair $(\mathbf{x}, \mathbf{u})$ used in the construction of the ball, while the fourth column details the nature of the functions f, g defining (P).

Before proceedings to the connection between the RYU ball and the existing results of the literature, let us make two important remarks regarding the choice of the couple $(\mathbf{x}, \mathbf{u})$ involved in the construction of the safe ball.

First, an approach which has been considered (often implicitly) in many contributions of the literature consists in choosing a feasible pair $(\mathbf{x}, \mathbf{u})$ such that

$$\mathbf{A}^{\mathrm{T}}\mathbf{u} \in \partial g(\mathbf{x}). \tag{14}$$

Interestingly, when (14) is satisfied the function GAP can be related to two other well-known quantities, namely the Fenchel divergence of f (see [3, Def. 2]) and the Bregman divergence of f^*. In particular, the following lemma holds:

▶ **Lemma 3.** *Assume* (H1)–(H2) *hold and let*[4]

$$\texttt{Fen}(\mathbf{x}, \mathbf{u}) \triangleq f(\mathbf{A}\mathbf{x}) + f^*(-\mathbf{u}) + \langle \mathbf{u} \mid \mathbf{A}\mathbf{x} \rangle \tag{15}$$
$$\texttt{Breg}(\mathbf{x}, \mathbf{u}) \triangleq f^*(-\mathbf{u}) - f^*(\nabla f(\mathbf{A}\mathbf{x})) + \langle \mathbf{A}\mathbf{x} \mid \mathbf{u} + \nabla f(\mathbf{A}\mathbf{x}) \rangle. \tag{16}$$

If (14) *is verified, then* $\forall\, (\mathbf{x}, \mathbf{u}) \in \mathrm{dom}(P) \times \mathrm{dom}(-D)$*:*

$$\texttt{GAP}(\mathbf{x}, \mathbf{u}) = \texttt{Fen}(\mathbf{x}, \mathbf{u}) = \texttt{Breg}(\mathbf{x}, \mathbf{u}). \tag{17}$$

---

[4] $\texttt{Fen}(\mathbf{x}, \mathbf{u})$ corresponds to the Fenchel divergence of f evaluated at $(\mathbf{A}\mathbf{x}, -\mathbf{u})$; $\texttt{Breg}(\mathbf{x}, \mathbf{u})$ is reminiscent from the Bregman divergence of f^* evaluated at $(-\mathbf{u}, \nabla f(\mathbf{A}\mathbf{x}))$. We acknowledge that, in contrast to the standard definition of the Bregman divergence, (16) does not impose differentiability of f^*. However, under (H2), the differentiability of f implies that $\partial f(\mathbf{A}\mathbf{x}) = \{\nabla f(\mathbf{A}\mathbf{x})\}$, and, since f is also convex, proper, and closed under (H1) (see Theorem 11), we have $\mathbf{A}\mathbf{x} \in \partial f^*(\nabla f(\mathbf{A}\mathbf{x}))$. Furthermore, when f^* is differentiable, we recover $\mathbf{A}\mathbf{x} = \nabla f^*(\nabla f(\mathbf{A}\mathbf{x}))$, and (16) coincides with the standard definition of the Bregman divergence.

We refer the reader to Appendix C.1 for a proof of this result. The connection between the duality gap and the Fenchel/Bregman divergences is of interest in two respects. On the one hand, these divergences are sometimes more straightforward to express than the duality gap and thus give an alternative formulation to the proposed RYU ball under the particular assumption (14). On the other hand, some contributions of the literature (see Section 4.5) have directly expressed their safe ball as a function of $\mathtt{Breg}(\mathbf{x}, \mathbf{u})$. The connection established in Theorem 3 will thus allow us to make a direct link between these works and the RYU framework proposed in this paper.

Second, we mention that the following definition of primal-dual couple $(\mathbf{x}, \mathbf{u})$ has been considered in many contributions of the literature (see e.g., [17, 21, 24, 25]):

$$(\mathbf{x}, \mathbf{u}) \triangleq \left(\mathbf{x}_\gamma^\star, -\gamma^{-1}\nabla f(\mathbf{A}\mathbf{x}_\gamma^\star)\right) \tag{18}$$

where $\gamma > 0$ and

$$\mathbf{x}_\gamma^\star \in \operatorname*{arg\,min}_{\mathbf{x}\in\mathbf{R}^n} f(\mathbf{A}\mathbf{x}) + \gamma g(\mathbf{x}). \tag{19}$$

This type of construction appears for example in “sequential” settings where one wants to solve (P) for $g(\,\cdot\,) = \lambda\|\cdot\|$ where $\|\cdot\|$ denotes some norm on $\mathbf{R}^n$ and the solution of a similar problem with $g(\,\cdot\,) = \lambda_0\|\cdot\|$ has already been computed previously. The solution of the latter problem can then be expressed as in (19) with $\gamma = \frac{\lambda_0}{\lambda}$ and $g(\,\cdot\,) = \lambda\|\cdot\|$.

The next lemma emphasizes that (18)–(19) correspond in fact to a particular strategy to build primal-dual couples obeying (14):

▶ **Lemma 4.** *If* $(\mathbf{x}, \mathbf{u})$ *is defined as in* (18)–(19)*, then it verifies* (14)*.*

Proof. Since γ is assumed positive, (19) can be equivalently rewritten:

$$\mathbf{x}_\gamma^\star \in \operatorname*{arg\,min}_{\mathbf{x}\in\mathbf{R}^n} \gamma^{-1} f(\mathbf{A}\mathbf{x}) + g(\mathbf{x}). \tag{20}$$

The functions $\gamma^{-1}f$ and g are convex, proper and closed from (H1) and $\gamma^{-1}f$ is $(\gamma\alpha)^{-1}$-Lipschitz smooth over $\mathbf{R}^m$. One obtains (14) by expanding the optimality conditions (11)–(12) associated to (20) and its dual. ◀

Since many safe balls proposed in the literature rely on the particular construction (18)–(19), the result in Theorem 4 emphasizes that these constructions in fact consider a primal-dual feasible pair $(\mathbf{x}, \mathbf{u})$ verifying (14) and that (from Theorem 3) the connection (17) between the duality gap and the Fenchel/Bregman divergences is thus in force.

4.1 GAP balls

The GAP ball first proposed in [10] and later generalized in [18, 19] is defined in (4a)–(4b). Its construction is valid under assumptions (H1)–(H2) and can take any primal-dual $(\mathbf{x}, \mathbf{u})$ as input. The next result shows that the RYU ball is always a subset of the GAP ball:

▶ **Proposition 5** (The GAP ball contains the RYU ball)**.** *Assume* (H1)–(H2) *hold. Then, for any primal-dual pair* $(\mathbf{x}, \mathbf{u}) \in \mathrm{dom}(P) \times \mathrm{dom}(-D)$*:*

$$\mathcal{B}_{\mathtt{RYU}}(\mathbf{x}, \mathbf{u}) \subseteq \mathcal{B}(\mathbf{c}_{\mathtt{GAP}}, r_{\mathtt{GAP}}). \tag{21}$$

Moreover, the inclusion is strict as soon as the primal-dual pair $(\mathbf{x}, \mathbf{u})$ *is not optimal.*

In [14, §IV.B] a variant of the GAP ball for the specific case where $g = \lambda\|\cdot\|_1$ was proposed. Although the construction of the ball presented in [14] holds in a slightly more general setup,[5] we focus hereafter on the case where f is proper, closed, convex and satisfies (H2). The center and radius of this ball (referred to as $\mathbf{x}$-GAP since its center depends on $\mathbf{x}$ instead of $\mathbf{u}$ as in the standard GAP ball) reads as follows:

$$\mathbf{c}_{\mathbf{x}\text{-}\mathtt{GAP}} \triangleq -\nabla f(\mathbf{A}\mathbf{x}) \tag{22a}$$

$$r_{\mathbf{x}\text{-}\mathtt{GAP}} \triangleq \sqrt{\frac{2\mathtt{GAP}(\mathbf{x}, \mathbf{u})}{\alpha}} \tag{22b}$$

where $(\mathbf{x}, \mathbf{u})$ can be any primal-dual feasible couple. Similarly to Theorem 5, the next result shows that the $\mathbf{x}$-GAP ball is also a superset of the proposed RYU region:

[5] In particular, a slightly weaker version of (H2) is considered.

▶ **Proposition 6** (The x-GAP ball contains the RYU ball)**.** *Assume* (H1)–(H2) *hold. Then, for any primal-dual pair* $(\mathbf{x}, \mathbf{u}) \in \mathrm{dom}(P) \times \mathrm{dom}(-D)$*:*

$$\mathcal{B}_{\texttt{RYU}}(\mathbf{x}, \mathbf{u}) \subseteq \mathcal{B}(\mathbf{c}_{\mathbf{x}\texttt{-GAP}}, r_{\mathbf{x}\texttt{-GAP}}). \tag{23}$$

Moreover, the inclusion is strict as soon as the primal-dual pair $(\mathbf{x}, \mathbf{u})$ *is not optimal.*

The proof of Theorems 5 and 6 is given in Appendix C.2.

4.2 Dynamic EDPP ball

In [26], the authors focused on the particular family of problems where

$$f = \tfrac{1}{2}\|\mathbf{y} - \cdot\,\|_2^2 \tag{24}$$

$$g = \lambda\|\cdot\| \tag{25}$$

for some vector $\mathbf{y} \in \mathbf{R}^m$, scalar $\lambda > 0$ and norm $\|\cdot\|$.[6] They introduced a new safe ball (see [26, Thm. 9]), dubbed "dynamic EDPP ball" and valid for any primal-dual feasible couple $(\mathbf{x}, \mathbf{u})$. The center and radius of the dynamic EDPP ball reads as follows:

$$\mathbf{c}_{\texttt{dyn.EDPP}} = \frac{1}{2}(\mathbf{y} + \mathbf{u} - t^\star \mathbf{A}\mathbf{x}) \tag{26a}$$

$$r_{\texttt{dyn.EDPP}} = \frac{1}{2}\sqrt{\|\mathbf{y} - \mathbf{u}\|_2^2 - \|t^\star \mathbf{A}\mathbf{x}\|_2^2} \tag{26b}$$

where $t^\star$ is defined as (with the conventions $0/0 = 0$ and $1/0 = +\infty$):

$$t^\star = \max\left(0, \frac{\langle \mathbf{A}\mathbf{x} \mid \mathbf{y} + \mathbf{u}\rangle - 2\lambda\|\mathbf{x}\|}{\|\mathbf{A}\mathbf{x}\|_2^2}\right). \tag{27}$$

The connection between the RYU and dynamic EDPP balls is established in the following proposition:

▶ **Proposition 7** (Dynamic EDPP ball is a special case of the RYU ball)**.** *Assume* (24)–(25) *holds. Then, for any* $(\mathbf{x}, \mathbf{u}) \in \mathrm{dom}(P) \times \mathrm{dom}(-D)$*:*

i. *The quantity* $t^\star$ *defined in* (27) *verifies*

$$t^\star \in \operatorname*{arg\,min}_{t \geq 0} \sqrt{\texttt{GAP}(t\mathbf{x}, \mathbf{u}) - \frac{1}{4}\|\mathbf{u} + \nabla f(t\mathbf{A}\mathbf{x})\|_2^2}.$$

ii. *We have*

$$\mathcal{B}(\mathbf{c}_{\texttt{dyn.EDPP}}, r_{\texttt{dyn.EDPP}}) = \mathcal{B}_{\texttt{RYU}}(t^\star\mathbf{x}, \mathbf{u}).$$

A proof of this result is available in Appendix C.3. In particular, the dynamic EDPP ball was shown to exhibit a smaller radius as compared to the GAP ball constructed with the (feasible) primal-dual pair $(\mathbf{x}, \mathbf{u})$, see [26, Thm. 10]. Item ii. of Theorem 7 (combined with Theorem 5) elucidates this connection by showing that the dynamic EDPP ball is in fact a subset of the GAP ball $\mathcal{B}_{\texttt{GAP}}(t^\star\mathbf{x}, \mathbf{u})$. Finally, item i. of Theorem 7 provides a novel interpretation of the definition of $t^\star$, that is $t^\star$ corresponds to a nonnegative rescaling of the primal vector $\mathbf{x}$ minimizing the radius of the RYU ball.

4.3 FNE, EDPP, DPP and SASVI balls

FNE [15], (E)DPP [24] and SASVI [17] balls are safe regions designed for the same problem where

$$f = \tfrac{1}{2}\|\mathbf{y} - \cdot\,\|_2^2 \tag{28}$$

$$g = \lambda\|\cdot\|_1. \tag{29}$$

[6] More precisely, the authors of [26] considered gauge functions instead of norms for the definition of g. Although the results presented in this section still hold in this more general setup, we stick to norms to simplify the exposition.

They all assume (explicitly or implicitly) that the primal-dual couple $(\mathbf{x}, \mathbf{u})$ used in the construction verifies (14). We start with the description of the FNE ball which corresponds to the most general construction. We address the EDPP, DPP and SASVI balls at the end of the section as particular cases or relaxation of the FNE region.

The FNE ball is defined by the following center and radius:

$$\mathbf{c}_{\mathtt{FNE}} = \mathbf{u} + \frac{1}{2}(\mathbf{y} - \mathbf{A}\mathbf{x} - \mathbf{u}) \tag{30a}$$

$$r_{\mathtt{FNE}} = \frac{1}{2}\left\|\mathbf{y} - \mathbf{A}\mathbf{x} - \mathbf{u}\right\|_2. \tag{30b}$$

In [15, Thm. 1], the authors showed that the FNE ball is safe for any primal-dual feasible pair $(\mathbf{x}, \mathbf{u})$ satisfying

$$\langle \mathbf{u} \mid \mathbf{A}\mathbf{x}\rangle = \lambda\|\mathbf{x}\|_1. \tag{31}$$

The following result shows that the FNE ball in fact corresponds to a particular case of the RYU ball when (28)–(29) and (31) hold:

▶ **Proposition 8** (FNE ball is a special case of the RYU ball)**.** *Assume* (28)–(29) *holds. Then, for any* $(\mathbf{x}, \mathbf{u}) \in \mathrm{dom}(P) \times \mathrm{dom}(-D)$ *satisfying* (31)*:*

$$\mathcal{B}(\mathbf{c}_{\mathtt{FNE}}, r_{\mathtt{FNE}}) = \mathcal{B}_{\mathtt{RYU}}(\mathbf{x}, \mathbf{u}).$$

A proof of this result is available in Appendix C.4. We note that the authors also showed in [15, Lem. 1] that the FNE ball is a strict subset of the GAP ball as long as $(\mathbf{x}, \mathbf{u}) \neq (\mathbf{x}^\star, \mathbf{u}^\star)$. Interestingly, in view of Theorem 8, this result turns out to be a particular case of Theorem 5 in the more general framework of the RYU ball.

The EDPP and SASVI balls represent specific instances of the FNE ball, resulting from particular choices of the pair $(\mathbf{x}, \mathbf{u})$. On the one hand, the SASVI ball (see [17, §2.2]) corresponds to the case where $(\mathbf{x}, \mathbf{u}) = (\mathbf{0}_n, \mathbf{u})$ for some dual feasible point $\mathbf{u}$, i.e.,

$$\mathbf{c}_{\mathtt{SASVI}} = \frac{1}{2}(\mathbf{y} + \mathbf{u}) \tag{32a}$$

$$r_{\mathtt{SASVI}} = \frac{1}{2}\left\|\mathbf{y} - \mathbf{u}\right\|_2. \tag{32b}$$

It is easy to see this couple trivially verifies (31). Using Theorem 8 with $(\mathbf{x}, \mathbf{u}) = (\mathbf{0}_n, \mathbf{u})$, $\mathbf{u} \in \mathrm{dom}(-D)$, then directly leads to

$$\mathcal{B}(\mathbf{c}_{\mathtt{SASVI}}, r_{\mathtt{SASVI}}) = \mathcal{B}_{\mathtt{RYU}}(\mathbf{0}_n, \mathbf{u}).$$

On the other hand, the center and radius of the EDPP ball (see [24, Thm. 13]) obeys the same definition (30a)–(30b) as those of the FNE ball but for $(\mathbf{x}, \mathbf{u})$ defined as in (18)–(19) for some $\gamma > 0$. Taking into account that

$$\partial\|\mathbf{x}\|_1 = \{\mathbf{z} \in \mathbf{R}^n \mid \langle \mathbf{z} \mid \mathbf{x}\rangle = \|\mathbf{x}\|_1, \|\mathbf{z}\|_\infty \leq 1\}, \tag{33}$$

it is easy to see that (31) together with feasibility of $(\mathbf{x}, \mathbf{u})$ is in fact an equivalent rewriting of (14) for $g = \lambda\|\cdot\|_1$. Hence, in view of Theorem 4, the couple $(\mathbf{x}, \mathbf{u})$ considered in the EDPP construction verifies (31) and this ball is nothing but a particular instance of FNE ball. Theorem 8 thus applies to the EDPP ball as well.

Finally, it was shown in [24, Thm. 13] that the DPP ball is always a superset of the EDPP ball. This directly leads to the inclusion reported in Table 1.

4.4 SAFE ball

The SAFE ball is the first safe region proposed in the seminal paper [6]. It applies in the case where

$$f = \tfrac{1}{2}\|\mathbf{y} - \cdot\|_2^2 \tag{34}$$

$$g = \lambda\|\cdot\|_1. \tag{35}$$

Its center and radius are defined $\forall\, \mathbf{u} \in \mathrm{dom}(-D)$ as

$$\mathbf{c}_{\mathtt{SAFE}} = \mathbf{y}$$

$$r_{\mathtt{SAFE}} = \|\mathbf{y} - \mathbf{u}\|_2.$$

In Appendix C.5, we show that the SAFE ball is a relaxation of the proposed RYU ball for $(\mathbf{x}, \mathbf{u}) = (\mathbf{0}_n, \mathbf{u})$ with $\mathbf{u} \in \mathrm{dom}(-D)$. More specifically, we prove that the following result holds:

▶ **Proposition 9.** *Assume* (34)–(35) *holds. Then, for any* $\mathbf{u} \in \mathrm{dom}(-D)$*:*

$$\mathcal{B}(\mathbf{c}_{\texttt{SAFE}}, r_{\texttt{SAFE}}) \supseteq \mathcal{B}_{\texttt{RYU}}(\mathbf{0}_n, \mathbf{u}). \tag{36}$$

4.5 SLORE and SFER balls

We end up this section by considering the SLORES and SFER balls respectively proposed in [25, Thm. 2] and [21, Cor. 1]. The focus of these papers is on problem (P) with the following definitions for f and g:

$$f(\mathbf{z}) = \sum_{i=1}^{m} \log(1 + \mathrm{e}^{-z_i}), \tag{37}$$

$$g(\mathbf{z}) = \lambda \|\mathbf{z}\|_1. \tag{38}$$

The construction of these balls is moreover based on the knowledge of a primal-dual couple $(\mathbf{x}, \mathbf{u})$ verifying (18)–(19) for some $\gamma > 0$. The expression of the center and radius of the SLORE and SFER balls respectively read as

$$\mathbf{c}_{\texttt{SLORES}} = \gamma \mathbf{u} \tag{39a}$$

$$r_{\texttt{SLORES}} = \sqrt{\frac{1}{2} \texttt{Breg}(\mathbf{x}, \mathbf{u})}, \tag{39b}$$

and

$$\mathbf{c}_{\texttt{SFER}} = \frac{1}{2}(1+\gamma)\mathbf{u} \tag{40a}$$

$$r_{\texttt{SFER}} = \sqrt{\frac{1}{4} \texttt{Breg}(\mathbf{x}, \mathbf{u}) - \frac{1}{4} \|(1-\gamma)\mathbf{u}\|_2^2}. \tag{40b}$$

In [21, Thm. 3], it was shown that

$$\mathcal{B}(\mathbf{c}_{\texttt{SFER}}, r_{\texttt{SFER}}) \subseteq \mathcal{B}(\mathbf{c}_{\texttt{SLORES}}, r_{\texttt{SLORES}}). \tag{41}$$

The next result shows that the SFER ball is a particular instance of RYU ball, thereby proving the results in Table 1:

▶ **Proposition 10.** *Assume* (37)–(38) *hold and* $(\mathbf{x}, \mathbf{u})$ *is defined as* (18)–(19) *for some* $\gamma > 0$*. Then, we have*

$$\mathcal{B}(\mathbf{c}_{\texttt{SFER}}, r_{\texttt{SFER}}) = \mathcal{B}_{\texttt{RYU}}(\mathbf{x}, \mathbf{u}). \tag{42}$$

A proof of this result is available in Appendix C.6.

5 Conclusion

In this paper, we introduced a new framework for constructing safe balls for a broad class of optimization problems. Specifically, our approach addresses cases where the cost function consists of a closed, proper, convex, Lipschitz-smooth term combined with another closed, proper, convex term, and only relies on the knowledge of a primal-dual feasible pair. The proposed construction not only unifies existing methods but also extends and improves upon all prior approaches from the last decade, providing a comprehensive framework for generating safe balls within this family of optimization problems and connecting existing results of the literature.

A Convex analysis

This appendix reminds two standard results from convex analysis. The first result relates the subdifferential of a function to the subdifferential of its convex conjugate:

▶ **Lemma 11** (Subdifferential inversion)**.** *Let* $h\colon \mathbf{R}^d \to \overline{\mathbf{R}}$ *be a proper, closed and convex function. Then, for all* $\mathbf{z}, \mathbf{z}^* \in \mathbf{R}^d$*:*

$$\mathbf{z}^* \in \partial h(\mathbf{z}) \iff \mathbf{z} \in \partial h^*(\mathbf{z}^*). \tag{43}$$

A proof of this lemma can be found in [2, Thm. 4.20].

The second result recalls two Fenchel-Young inequalities:

▶ **Lemma 12** (Fenchel-Young inequalities). *Let $h\colon \mathbf{R}^d \to \overline{\mathbf{R}}$ be a proper and convex function. Then, for all $\mathbf{z}, \mathbf{z}^* \in \mathbf{R}^d$:*

$$h(\mathbf{z}) + h^*(\mathbf{z}^*) \geq \langle \mathbf{z}^* \mid \mathbf{z} \rangle \tag{44}$$

with equality if and only if $\mathbf{z}^* \in \partial h(\mathbf{z})$.

If h is moreover closed and α-strongly convex, then for all $\mathbf{z}, \mathbf{z}^* \in \mathbf{R}^d$:

$$h(\mathbf{z}) + h^*(\mathbf{z}^*) \geq \langle \mathbf{z}^* \mid \mathbf{z} \rangle + \frac{\alpha}{2}\|\mathbf{z} - \nabla h^*(\mathbf{z}^*)\|_2^2. \tag{45}$$

We note that (44) corresponds to the standard formulation of the well-known Fenchel-Young inequality. A proof of this result follows from [2, Thm. 4.6] and [2, Thm. 4.20]. (45) is a refined version of the Fenchel-Young inequality which applies to closed and strongly convex functions. Since the latter is less common in the literature, a proof is provided hereafter.[7]

Proof of (45). Assume that h is α-strongly convex and let $\mathbf{z}, \mathbf{z}^* \in \mathbf{R}^d$. First note that since h is proper, closed and α-strongly convex, we have from [2, Thm. 5.26.(b)] that h^* is α^{-1}-Lipschitz smooth over $\mathbf{R}^m$. In particular, $\mathrm{dom}(h^*) = \mathbf{R}^d$ and h^* is differentiable at any $\mathbf{z}^* \in \mathbf{R}^d$, that is

$$\partial h^*(\mathbf{z}^*) = \{\nabla h^*(\mathbf{z}^*)\}. \tag{46}$$

Second, if $\mathbf{z} \notin \mathrm{dom}(h)$, then the left-hand side of (45) is infinite and the inequality trivially holds true since the right-hand side is finite. We conclude the proof by showing that (45) is also valid for $\mathbf{z} \in \mathrm{dom}(h)$. Using the fact that h is proper, closed and convex, we obtain from Theorem 11 with the pair $(\nabla h^*(\mathbf{z}^*), \mathbf{z}^*)$ that

$$\mathbf{z}^* \in \partial h(\nabla h^*(\mathbf{z}^*)),$$

and from [1, Prop. 16.4.(i)],

$$\nabla h^*(\mathbf{z}^*) \in \mathrm{dom}(h).$$

Invoking the first-order characterization of α-strong convexity of h at $\nabla h^*(\mathbf{z}^*) \in \mathrm{dom}(h)$ (see [2, Thm. 5.24]) then leads to

$$h(\mathbf{z}) \geq h(\nabla h^*(\mathbf{z}^*)) + \langle \mathbf{z}^* \mid \mathbf{z} - \nabla h^*(\mathbf{z}^*)\rangle + \tfrac{\alpha}{2}\|\mathbf{z} - \nabla h^*(\mathbf{z}^*)\|^2. \tag{47}$$

Finally, considering the standard Fenchel-Young inequality (44) with $\mathbf{z} = \nabla h^*(\mathbf{z}^*)$ and using (46), we have that the following equality holds:

$$h(\nabla h^*(\mathbf{z}^*)) + h^*(\mathbf{z}^*) = \langle \mathbf{z}^* \mid \nabla h^*(\mathbf{z}^*)\rangle. \tag{48}$$

We obtain the desired result (45) by re-injecting (48) into (47). ◀

B Proofs related to construction of the RYU framework

B.1 Proof of Theorem 2

We first notice that our result in Theorem 2 is an equivalent rewriting[8] of the following proposition:

▶ **Proposition 13.** *If hypotheses* (H1)–(H2) *hold true then the following inequality is satisfied for any $(\mathbf{x}, \mathbf{u}) \in \mathbf{R}^n \times \mathbf{R}^m$:*

$$\|\mathbf{u}^\star - \mathbf{u}\|_2^2 + \|\mathbf{u}^\star + \nabla f(\mathbf{A}\mathbf{x})\|_2^2 \leq \frac{2\mathtt{GAP}(\mathbf{x}, \mathbf{u})}{\alpha}. \tag{49}$$

In the rest of this section, we thus concentrate on the proof of Theorem 13. Our arguments leverage the following lemma whose proof is postponed to Appendix B.2:

[7] Another version of the proof can be found in [20, Lem. 5].
[8] See Appendix B.3.

▶ **Lemma 14.** *If hypotheses* (H1)–(H2) *hold true, then the following inequality is satisfied for any* $(\mathbf{x}, \mathbf{u}) \in \mathbf{R}^n \times \mathbf{R}^m$*:*

$$\|\mathbf{u} + \nabla f(\mathbf{A}\mathbf{x})\|_2^2 \leq \frac{2\mathtt{GAP}(\mathbf{x}, \mathbf{u})}{\alpha}. \tag{50}$$

Theorem 13 can be proved by applying Theorem 14 for two different choices of couple $(\mathbf{x}, \mathbf{u})$. We remind the reader that: *i)* we assume that (P) admits (at least) one minimizer $\mathbf{x}^\star$; *ii)* there exists a unique maximizer $\mathbf{u}^\star$ to (D) under (H1)–(H2), see Theorem 1. A first application of Theorem 14 with $(\mathbf{x}^\star, \mathbf{u})$ then leads to

$$\|\mathbf{u} + \nabla f(\mathbf{A}\mathbf{x}^\star)\|_2^2 \leq \frac{2(P(\mathbf{x}^\star) - D(\mathbf{u}))}{\alpha}.$$

Since Theorem 1 also ensures that strong duality holds, we have $P(\mathbf{x}^\star) = D(\mathbf{u}^\star)$ and $\mathbf{u}^\star = -\nabla f(\mathbf{A}\mathbf{x}^\star)$. Therefore, the previous inequality can also be rewritten as

$$\|\mathbf{u} - \mathbf{u}^\star\|_2^2 \leq \frac{2(D(\mathbf{u}^\star) - D(\mathbf{u}))}{\alpha}. \tag{51}$$

A second application of Theorem 14 with the pair $(\mathbf{x}, \mathbf{u}^\star)$ yields

$$\|\mathbf{u}^\star + \nabla f(\mathbf{A}\mathbf{x})\|_2^2 \leq \frac{2(P(\mathbf{x}) - D(\mathbf{u}^\star))}{\alpha}. \tag{52}$$

Summing up (51) and (52) leads to the desired result (49).

B.2 Proof of Theorem 14

If $(\mathbf{x}, \mathbf{u}) \notin \mathrm{dom}(P) \times \mathrm{dom}(-D)$, then (50) is trivially satisfied since the left-hand side is finite whereas the right-hand side is equal to $+\infty$. In the rest of the proof, we thus assume that $(\mathbf{x}, \mathbf{u}) \in \mathrm{dom}(P) \times \mathrm{dom}(-D)$.

As discussed in Section 3, hypotheses (H1)–(H2) imply that f^* is proper, closed and α-strongly convex. Applying the refined Fenchel-Young inequality (45) with $h = f^*$, $\mathbf{z} = -\mathbf{u}$ and $\mathbf{z}^* = \mathbf{A}\mathbf{x}$ then leads to

$$f^*(-\mathbf{u}) + f^{**}(\mathbf{A}\mathbf{x}) \geq \langle -\mathbf{u} \mid \mathbf{A}\mathbf{x}\rangle + \frac{\alpha}{2}\|\mathbf{u} + \nabla f^{**}(\mathbf{A}\mathbf{x})\|_2^2. \tag{53}$$

Using [2, Thm. 4.8], we have that $f^{**} = f$ since f is proper, closed and convex by (H1). Therefore, (53) can be equivalently rewritten as:

$$f(\mathbf{A}\mathbf{x}) + f^*(-\mathbf{u}) \geq \langle -\mathbf{u} \mid \mathbf{A}\mathbf{x}\rangle + \frac{\alpha}{2}\|\mathbf{u} + \nabla f(\mathbf{A}\mathbf{x})\|_2^2. \tag{54}$$

In order to conclude the proof, we need to add $g(\mathbf{x}) + g^*(\mathbf{A}^{\mathrm{T}}\mathbf{u})$ to both sides of this inequality. Prior to this operation, we have nevertheless to ensure that $g(\mathbf{x}) + g^*(\mathbf{A}^{\mathrm{T}}\mathbf{u}) < +\infty$. To that end, we notice that $f(\mathbf{x}) > -\infty$ and $f^*(-\mathbf{u}) > -\infty$ since f and f^* are proper. Hence,

$$\begin{aligned} P(\mathbf{x}) < +\infty &\implies g(\mathbf{x}) < +\infty \\ -D(\mathbf{u}) < +\infty &\implies g^*(\mathbf{A}^{\mathrm{T}}\mathbf{u}) < +\infty. \end{aligned}$$

Since we assume that $(\mathbf{x}, \mathbf{u}) \in \mathrm{dom}(P) \times \mathrm{dom}(-D)$, the left-hand sides of these implications are satisfied, so that $g(\mathbf{x}) + g^*(\mathbf{A}^{\mathrm{T}}\mathbf{u}) < +\infty$.

Adding $g(\mathbf{x}) + g^*(\mathbf{A}^{\mathrm{T}}\mathbf{u})$ to both sides of (54) then leads to

$$P(\mathbf{x}) - D(\mathbf{u}) \geq g(\mathbf{x}) + g^*(\mathbf{A}^{\mathrm{T}}\mathbf{u}) - \langle \mathbf{u} \mid \mathbf{A}\mathbf{x}\rangle + \frac{\alpha}{2}\|\mathbf{u} + \nabla f(\mathbf{A}\mathbf{x})\|_2^2.$$

Finally, since g is proper, closed and convex from (H1), we can apply the Fenchel-Young inequality (44) with $h = g$, $\mathbf{z} = \mathbf{x}$ and $\mathbf{z}^* = \mathbf{A}^{\mathrm{T}}\mathbf{u}$ to obtain (50).

B.3 Equivalent rewriting of the RYU ball

In this section, we show the following set equality:

$$\{\mathbf{u}^\star \in \mathbf{R}^m \mid \|\mathbf{u}^\star - \mathbf{c}_{\mathtt{RYU}}\|_2 \leq r_{\mathtt{RYU}}\} = \left\{\mathbf{u}^\star \in \mathbf{R}^m \;\middle|\; \|\mathbf{u}^\star - \mathbf{u}\|_2^2 + \|\mathbf{u}^\star + \nabla f(\mathbf{A}\mathbf{x})\|_2^2 \leq \frac{2\mathtt{GAP}(\mathbf{x}, \mathbf{u})}{\alpha}\right\}.$$

This result is a consequence of the parallelogram identity which reads (see, e.g., [1, Lem. 2.12.(ii)])

$$\forall\, \mathbf{z}, \mathbf{z}' \in \mathbf{R}^m : \ \|\mathbf{z}\|_2^2 + \|\mathbf{z}'\|_2^2 = \frac{1}{2}\|\mathbf{z} - \mathbf{z}'\|_2^2 + \frac{1}{2}\|\mathbf{z} + \mathbf{z}'\|_2^2 . \tag{55}$$

More precisely, applying (55) with

$$\begin{aligned} \mathbf{z} &= \mathbf{u}^\star - \mathbf{u} \\ \mathbf{z}' &= \mathbf{u}^\star + \nabla f(\mathbf{A}\mathbf{x}), \end{aligned}$$

we obtain

$$\begin{aligned} \|\mathbf{u}^\star - \mathbf{u}\|_2^2 + \|\mathbf{u}^\star + \nabla f(\mathbf{A}\mathbf{x})\|_2^2 &= \frac{1}{2}\|\mathbf{u} + \nabla f(\mathbf{A}\mathbf{x})\|_2^2 + 2\left\|\mathbf{u}^\star - \frac{\mathbf{u} - \nabla f(\mathbf{A}\mathbf{x})}{2}\right\|_2^2 \\ &= \frac{1}{2}\|\mathbf{u} + \nabla f(\mathbf{A}\mathbf{x})\|_2^2 + 2\|\mathbf{u}^\star - \mathbf{c}_{\mathtt{RYU}}\|_2^2 . \end{aligned}$$

Using this equality, we have

$$\begin{aligned} \|\mathbf{u}^\star - \mathbf{u}\|_2^2 + \|\mathbf{u}^\star + \nabla f(\mathbf{A}\mathbf{x})\|_2^2 \leq \frac{2\mathtt{GAP}(\mathbf{x}, \mathbf{u})}{\alpha} &\iff \frac{1}{2}\|\mathbf{u} + \nabla f(\mathbf{A}\mathbf{x})\|_2^2 + 2\|\mathbf{u}^\star - \mathbf{c}_{\mathtt{RYU}}\|_2^2 \leq \frac{2\mathtt{GAP}(\mathbf{x}, \mathbf{u})}{\alpha} \\ &\iff \|\mathbf{u}^\star - \mathbf{c}_{\mathtt{RYU}}\|_2^2 \leq \frac{\mathtt{GAP}(\mathbf{x}, \mathbf{u})}{\alpha} - \frac{1}{4}\|\mathbf{u} + \nabla f(\mathbf{A}\mathbf{x})\|_2^2 . \end{aligned} \tag{57}$$

One finally obtains the desired result by noting that the right-hand side of (57) is equal to $r_{\mathtt{RYU}}^2$ as defined in (6b).

C Proofs of the connections with existing results

This appendix gathers all the proofs related to the comparison of the proposed RYU ball with state-of-the-art safe regions presented in Section 4.1 to Section 4.5.

C.1 Proof of Theorem 3

Let $(\mathbf{x}, \mathbf{u}) \in \mathrm{dom}(P) \times \mathrm{dom}(-D)$. Using the definition of the primal and dual cost functions, we obtain:

$$\begin{aligned} \mathtt{GAP}(\mathbf{x}, \mathbf{u}) &= f(\mathbf{A}\mathbf{x}) + g(\mathbf{x}) + f^*(-\mathbf{u}) + g^*(\mathbf{A}^{\mathrm{T}}\mathbf{u}) \\ &= f(\mathbf{A}\mathbf{x}) + f^*(-\mathbf{u}) + \langle \mathbf{A}\mathbf{x} \mid \mathbf{u} \rangle + g(\mathbf{x}) + g^*(\mathbf{A}^{\mathrm{T}}\mathbf{u}) - \langle \mathbf{A}\mathbf{x} \mid \mathbf{u} \rangle \\ &= \mathtt{Fen}(\mathbf{x}, \mathbf{u}) + g(\mathbf{x}) + g^*(\mathbf{A}^{\mathrm{T}}\mathbf{u}) - \langle \mathbf{A}\mathbf{x} \mid \mathbf{u} \rangle . \end{aligned}$$

If (14) holds, we then have from Theorem 12 that

$$g(\mathbf{x}) + g^*(\mathbf{A}^{\mathrm{T}}\mathbf{u}) - \langle \mathbf{A}\mathbf{x} \mid \mathbf{u} \rangle = 0.$$

This shows the first equality in (17).

The second inequality can be obtained by noticing that (from Theorem 12)

$$f(\mathbf{A}\mathbf{x}) + f^*(\nabla f(\mathbf{A}\mathbf{x})) = \langle \nabla f(\mathbf{A}\mathbf{x}) \mid \mathbf{A}\mathbf{x} \rangle \tag{58}$$

since f is proper, convex and $\partial f(\mathbf{A}\mathbf{x}) = \{\nabla f(\mathbf{A}\mathbf{x})\}$. Hence,

$$\begin{aligned} \mathtt{Fen}(\mathbf{x}, \mathbf{u}) &= f(\mathbf{A}\mathbf{x}) + f^*(-\mathbf{u}) + \langle \mathbf{A}\mathbf{x} \mid \mathbf{u} \rangle \\ &= f^*(-\mathbf{u}) - f^*(\nabla f(\mathbf{A}\mathbf{x})) + \langle \mathbf{A}\mathbf{x} \mid \mathbf{u} + \nabla f(\mathbf{A}\mathbf{x}) \rangle \\ &= \mathtt{Breg}(\mathbf{x}, \mathbf{u}). \end{aligned}$$

C.2 Proof of Theorem 5

Using Theorem 13, we have

$$\mathcal{B}(\mathbf{c}_{\mathtt{RYU}}, r_{\mathtt{RYU}}) = \left\{ \mathbf{u}' \in \mathbf{R}^m \,\middle|\, \|\mathbf{u}' - \mathbf{u}\|_2^2 + \|\mathbf{u}' + \nabla f(\mathbf{A}\mathbf{x})\|_2^2 \leq \frac{2\mathtt{GAP}(\mathbf{x}, \mathbf{u})}{\alpha} \right\}, \tag{59}$$

whereas the definitions of the GAP and $\mathbf{x}-$GAP balls (see (4) and (22), respectively) lead to:

$$\mathcal{B}(\mathbf{c}_{\texttt{GAP}}, r_{\texttt{GAP}}) = \left\{ \mathbf{u}' \in \mathbf{R}^m \,\middle|\, \|\mathbf{u}' - \mathbf{u}\|_2^2 \leq \frac{2\texttt{GAP}(\mathbf{x}, \mathbf{u})}{\alpha} \right\} \tag{60}$$

$$\mathcal{B}(\mathbf{c}_{\mathbf{x}\texttt{-GAP}}, r_{\mathbf{x}\texttt{-GAP}}) = \left\{ \mathbf{u}' \in \mathbf{R}^m \,\middle|\, \|\mathbf{u}' + \nabla f(\mathbf{A}\mathbf{x})\|_2^2 \leq \frac{2\texttt{GAP}(\mathbf{x}, \mathbf{u})}{\alpha} \right\}. \tag{61}$$

Since the membership conditions in (60) and (61) are relaxations of the inequality defining the RYU ball in (59), inclusions (21) and (23) necessarily hold.

Finally, to prove strict inclusion it is then sufficient to note that

$$r_{\texttt{RYU}} \leq \tfrac{\texttt{GAP}(\mathbf{x},\mathbf{u})}{\alpha} < r_{\texttt{GAP}}$$
$$r_{\texttt{RYU}} \leq \tfrac{\texttt{GAP}(\mathbf{x},\mathbf{u})}{\alpha} < r_{\mathbf{x}\texttt{-GAP}}$$

whenever $\texttt{GAP}(\mathbf{x}, \mathbf{u}) \neq 0$.

C.3 Proof of Theorem 7

We first note that the RYU ball in Theorem 2 is well-defined. Indeed, functions f and g in (28)–(29) are closed, proper and convex, so that (H1) holds. Moreover, (H2) is verified with $\alpha = 1$. For any feasible $(t\mathbf{x}, \mathbf{u})$ with $t \geq 0$, we thus have by definition:

$$\mathbf{c}_{\texttt{RYU}}(t) = \tfrac{1}{2}(\mathbf{u} + \mathbf{y} - t\mathbf{A}\mathbf{x}) \tag{62}$$

$$r_{\texttt{RYU}}^2(t) = \texttt{GAP}(t\mathbf{x}, \mathbf{u}) - \tfrac{1}{4}\|\mathbf{u} - \mathbf{y} + t\mathbf{A}\mathbf{x}\|_2^2. \tag{63}$$

Using the definitions of f and g in (24)–(25), we note that the duality gap can be expressed as

$$\begin{aligned}
\texttt{GAP}(\mathbf{x}, \mathbf{u}) &= \tfrac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\| - \tfrac{1}{2}\|\mathbf{y}\|_2^2 + \tfrac{1}{2}\|\mathbf{y} - \mathbf{u}\|_2^2 \\
&= \lambda\|\mathbf{x}\| + \tfrac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 - \langle \mathbf{y} \mid \mathbf{u} \rangle + \tfrac{1}{2}\|\mathbf{u}\|_2^2 \\
&= \lambda\|\mathbf{x}\| + \tfrac{1}{2}\|\mathbf{u} - \mathbf{y} + \mathbf{A}\mathbf{x}\|_2^2 - \langle \mathbf{A}\mathbf{x} \mid \mathbf{u} \rangle
\end{aligned}$$

so that

$$r_{\texttt{RYU}}^2(t) = \lambda t\|\mathbf{x}\| - t\langle \mathbf{A}\mathbf{x} \mid \mathbf{u} \rangle + \tfrac{1}{4}\|\mathbf{u} - \mathbf{y} + t\mathbf{A}\mathbf{x}\|_2^2. \tag{64}$$

Proving item i. of Theorem 7 is equivalent to showing that the variable $t^\star$ defined in (27) verifies

$$t^\star \in \operatorname*{arg\,min}_{t \geq 0}\; r_{\texttt{RYU}}^2(t). \tag{65}$$

Since $r_{\texttt{RYU}}^2(t)$ is a convex function, it is sufficient to show that $t^\star$ satisfies the problem's first-order optimality condition, i.e.,

$$\forall\, t \geq 0 : \; (r_{\texttt{RYU}}^2(t^\star))'(t - t^\star) \geq 0, \tag{66}$$

where

$$(r_{\texttt{RYU}}^2(t))' = \lambda\|\mathbf{x}\| - \tfrac{1}{2}\langle \mathbf{A}\mathbf{x} \mid \mathbf{u} + \mathbf{y} \rangle + \tfrac{t}{2}\|\mathbf{A}\mathbf{x}\|_2^2.$$

We distinguish between three cases. First, if $\mathbf{A}\mathbf{x} = \mathbf{0}_n$, then $(r_{\texttt{RYU}}^2(t))' = \lambda\|\mathbf{x}\| \geq 0$ so that

$$0 \in \operatorname*{arg\,min}_{t \geq 0} r_{\texttt{RYU}}^2(t).$$

In this case, the definition of $t^\star$ in (27) also leads to $t^\star = 0$ (by using the conventions $0/0 = 0$ and $1/0 = +\infty$). Second, if $\mathbf{A}\mathbf{x} \neq \mathbf{0}_n$ and

$$\widetilde{t} \triangleq \frac{\langle \mathbf{A}\mathbf{x} \mid \mathbf{y} + \mathbf{u} \rangle - 2\lambda\|\mathbf{x}\|}{\|\mathbf{A}\mathbf{x}\|_2^2} \geq 0,$$

we easily have that

$$\widetilde{t} \in \arg\min_{t \geq 0} r^2_{\mathtt{RYU}}(t)$$

since $(r^2_{\mathtt{RYU}}(\widetilde{t}))' = 0$. In this case, one deduces $t^\star = \frac{\langle \mathbf{Ax} | \mathbf{y}+\mathbf{u}\rangle - 2\lambda\|\mathbf{x}\|}{\|\mathbf{Ax}\|_2^2}$. Finally, if $\mathbf{Ax} \neq \mathbf{0}_n$ and $\widetilde{t} < 0$, we then have that $(r^2_{\mathtt{RYU}}(0))' \geq 0$ since $(r^2_{\mathtt{RYU}}(0))' = -\frac{\|\mathbf{Ax}\|_2^2}{2}\widetilde{t}$ and $\widetilde{t} < 0$. In this case, 0 is a minimizer since it verifies (66) and this corresponds again to the definition of $t^\star$ in (27).

Showing item ii. of Theorem 7 is tantamount to showing that

$$\mathbf{c}_{\mathtt{RYU}}(t^\star) = \mathbf{c}_{\mathtt{dyn.EDPP}} \tag{67}$$

$$r_{\mathtt{RYU}}(t^\star) = r_{\mathtt{dyn.EDPP}}. \tag{68}$$

On the one hand, since item i. of Theorem 7 is true, we directly have from the expression of $\mathbf{c}_{\mathtt{RYU}}(t)$ in (62) that (67) holds. On the other hand, (68) can be shown by examining the following two cases.

If $t^\star = 0$, the equality in (68) follows directly from the definition (64). If $t^\star > 0$, we have $(r^2_{\mathtt{RYU}}(t^\star))' = 0$, i.e.,

$$\lambda\|\mathbf{x}\| = \tfrac{1}{2}\langle \mathbf{Ax} \mid \mathbf{u}+\mathbf{y}\rangle - \tfrac{t^\star}{2}\|\mathbf{Ax}\|_2^2.$$

Plugging this equality into (64) then leads to

$$\begin{aligned} r^2_{\mathtt{RYU}}(t^\star) &= \tfrac{1}{2}\langle t^\star\mathbf{Ax} \mid \mathbf{y}-\mathbf{u}\rangle - \tfrac{1}{2}\left\|t^\star\mathbf{Ax}\right\|_2^2 + \tfrac{1}{4}\left\|t^\star\mathbf{Ax} - \mathbf{y} + \mathbf{u}\right\|_2^2 \\ &= \tfrac{1}{4}\|\mathbf{y}-\mathbf{u}\|_2^2 - \tfrac{1}{4}\|t^\star\mathbf{Ax}\|_2^2 \\ &= r^2_{\mathtt{dyn.EDPP}}. \end{aligned}$$

C.4 Proof of Theorem 8

It is straightforward to see from the definition of f and g in (28)–(29) that (H1)–(H2) are satisfied with $\alpha = 1$. The RYU ball in Theorem 2 is therefore well-defined. We next show that the center and radius of the RYU and FNE balls coincide.

First, using the definition of f in (28), we have

$$\nabla f(\mathbf{Ax}) = -(\mathbf{y} - \mathbf{Ax}). \tag{69}$$

Hence,

$$\mathbf{c}_{\mathtt{RYU}} = \tfrac{1}{2}(\mathbf{u} - \nabla f(\mathbf{Ax})) = \mathbf{u} + \tfrac{1}{2}(\mathbf{y} - \mathbf{Ax} - \mathbf{u}) = \mathbf{c}_{\mathtt{FNE}}.$$

Second, using the definition of g in (29) and (33), it can be seen that condition "$\langle \mathbf{u} \mid \mathbf{Ax}\rangle = \lambda\|\mathbf{x}\|_1$" in (31) together with feasibility of $(\mathbf{x}, \mathbf{u})$ is equivalent to "$\mathbf{A}^{\mathrm{T}}\mathbf{u} = \partial g(\mathbf{x})$" in (14) so that Theorem 3 applies. In particular, we have:

$$\mathtt{GAP}(\mathbf{x}, \mathbf{u}) = \mathtt{Fen}(\mathbf{x}, \mathbf{u}) = f(\mathbf{Ax}) + f^*(-\mathbf{u}) + \langle \mathbf{u} \mid \mathbf{Ax}\rangle.$$

Since

$$f^*(\mathbf{u}) = \tfrac{1}{2}\|\mathbf{u}\|_2^2 + \langle \mathbf{u} \mid \mathbf{y}\rangle,$$

the duality gap can thus also be written as

$$\begin{aligned} \mathtt{GAP}(\mathbf{x}, \mathbf{u}) &= \tfrac{1}{2}\|\mathbf{y} - \mathbf{Ax}\|_2^2 + \tfrac{1}{2}\|\mathbf{u}\|_2^2 - \langle \mathbf{u} \mid \mathbf{y}\rangle + \langle \mathbf{u} \mid \mathbf{Ax}\rangle \\ &= \tfrac{1}{2}\|\mathbf{y} - \mathbf{Ax} - \mathbf{u}\|_2^2. \end{aligned} \tag{70}$$

Going back to the definition of the radius of the RYU ball (with $\alpha = 1$), we finally obtain:

$$\begin{aligned} r^2_{\mathtt{RYU}} &= \mathtt{GAP}(\mathbf{x}, \mathbf{u}) - \tfrac{1}{4}\left\|\mathbf{u} + \nabla f(\mathbf{Ax})\right\|_2^2 \\ &= \tfrac{1}{2}\|\mathbf{y} - \mathbf{Ax} - \mathbf{u}\|_2^2 - \tfrac{1}{4}\|\mathbf{y} - \mathbf{Ax} - \mathbf{u}\|_2^2 \\ &= r^2_{\mathtt{FNE}}, \end{aligned}$$

where we have used (69) and (70) in the second equality.

C.5 Proof of Theorem 9

It is easy to see that the definitions of f and g in (28)–(29) verify (H1)–(H2) with $\alpha = 1$, so that the RYU ball in Theorem 2 is well-defined.

On the one hand, using the definition of f in (28) with $\mathbf{x} = \mathbf{0}_n$, we have

$$\nabla f(\mathbf{A}\mathbf{x}) = -\mathbf{y}.$$

On the other hand, noticing that the couple $(\mathbf{0}_n, \mathbf{u})$ verifies (14) and using the same reasoning as in the proof of Theorem 8, we obtain from (70):

$$\mathtt{GAP}(\mathbf{x}, \mathbf{u}) = \tfrac{1}{2}\|\mathbf{y} - \mathbf{u}\|_2^2.$$

Finally, using Theorem 13, we have

$$\mathcal{B}_{\mathtt{RYU}}(\mathbf{0}_n, \mathbf{u}) = \left\{ \mathbf{u}' \in \mathbf{R}^m \;\middle|\; \|\mathbf{u}' - \mathbf{u}\|_2^2 + \|\mathbf{u}' - \mathbf{y}\|_2^2 \leq \|\mathbf{y} - \mathbf{u}\|_2^2 \right\}, \tag{71}$$

whereas the SAFE ball is defined as

$$\mathcal{B}(\mathbf{c}_{\mathtt{SAFE}}, r_{\mathtt{SAFE}}) = \left\{ \mathbf{u}' \in \mathbf{R}^m \;\middle|\; \|\mathbf{u}' - \mathbf{y}\|_2^2 \leq \|\mathbf{y} - \mathbf{u}\|_2^2 \right\}. \tag{72}$$

Since the membership condition in (72) is a relaxation of the inequality in (71), inclusion (36) holds.

C.6 Proof of Theorem 10

It is easy to see from the definition of f and g in (37)–(38) that (H1)–(H2) hold with $\alpha = 4$. The RYU ball in Theorem 2 is therefore well-defined.

We next show that the center and the radius of the RYU and SFER balls coincide. Since $(\mathbf{x}, \mathbf{u})$ is defined as in (18)–(19), we have from Theorem 4 that this couple satisfies (14). Using Theorem 3 then leads to

$$\mathtt{GAP}(\mathbf{x}, \mathbf{u}) = \mathtt{Breg}(\mathbf{x}, \mathbf{u}). \tag{73}$$

Moreover, we also have from (18)–(19):

$$\nabla f(\mathbf{A}\mathbf{x}) = -\gamma \mathbf{u}. \tag{74}$$

Using (73)–(74), we then easily find that

$$\begin{aligned}
\mathbf{c}_{\mathtt{RYU}} &= \tfrac{1}{2}(\mathbf{u} - \nabla f(\mathbf{A}\mathbf{x})) = \tfrac{1}{2}(1+\gamma)\mathbf{u} = \mathbf{c}_{\mathtt{SFER}} \\
r_{\mathtt{RYU}}^2 &= \tfrac{1}{4}\mathtt{GAP}(\mathbf{x}, \mathbf{u}) - \tfrac{1}{4}\|\mathbf{u} + \nabla f(\mathbf{A}\mathbf{x})\|_2^2 \\
&= \tfrac{1}{4}\,\mathtt{Breg}(\mathbf{x}, \mathbf{u}) - \tfrac{1}{4}\|(1-\gamma)\mathbf{u}\|_2^2 \\
&= r_{\mathtt{SFER}}^2.
\end{aligned}$$

This shows the equality in (42).

References

1 Heinz H. Bauschke and Patrick L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, 2nd edition, 2017.

2 Amir Beck. *First-Order Methods in Optimization*, volume 25 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics; Mathematical Optimization Society, 2017.

3 Mathieu Blondel, André F. T. Martins, and Vlad Niculae. Learning with fenchel-young losses. *J. Mach. Learn. Res.*, 21: article no. 35 (69 pages), 2020.

4 Scott Shaobing Chen, David L. Donoho, and Michael A. Saunders. Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.*, 20(1):33–61, 1999.

5 Liang Dai and Kristiaan Pelckmans. An ellipsoid based, two-stage screening test for BPDN. In *Proceedings of the European Signal Processing Conference (EUSIPCO)*, pages 654–658. IEEE, 2012.

6 Laurent El Ghaoui, Vivian Viallon, and Tarek Rabbani. Safe feature elimination for the Lasso and sparse supervised learning problems. *Pac. J. Optim.*, 8(4):667–698, 2012.

7 Clément Elvira and Cédric Herzet. Safe squeezing for antisparse coding. *IEEE Trans. Signal Process.*, 68:3252–3265, 2020.

8 Clément Elvira and Cédric Herzet. Short and squeezed: accelerating the computation of antisparse representations with safe squeezing. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 5615–5619. IEEE, 2020.

9 Clément Elvira and Cédric Herzet. Safe rules for the identification of zeros in the solutions of the slope problem. *SIAM J. Math. Data Sci.*, 5(1):147–173, 2023.

10 Olivier Fercoq, Alexandre Gramfort, and Joseph Salmon. Mind the duality gap: safer rules for the Lasso. In *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, volume 37, pages 333–342. 2015.

11 Cássio Fraga Dantas and Rémi Gribonval. Stable safe screening and structured dictionaries for faster ℓ_1 regularization. *IEEE Trans. Signal Process.*, 67(14):3756–3769, 2019.

12 Théo Guyard, Cédric Herzet, and Clément Elvira. Screen & relax: accelerating the resolution of Elastic-Net by safe identification of the solution support. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 5443–5447. IEEE, 2022.

13 Cédric Herzet, Clément Dorffer, and Angélique Drémeau. Gather and conquer: Region-based strategies to accelerate safe screening tests. *IEEE Trans. Signal Process.*, 67(12):3300–3315, 2019.

14 Cédric Herzet, Clément Elvira, and Hong-Phuong Dang. Region-free safe screening tests for ℓ_1-penalized convex problems. In *Proceedings of the European Signal Processing Conference (EUSIPCO)*, pages 2061–2065. 2022.

15 Cédric Herzet and Abed Malti. Safe screening tests for Lasso based on firmly non-expansiveness. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 4732–4736. IEEE, 2016.

16 Kwangmoo Koh, Seung-Jean Kim, and Stephen P. Boyd. An interior-point method for large-scale l1-regularized logistic regression. *J. Mach. Learn. Res.*, 8:1519–1555, 2007.

17 Jun Liu, Zheng Zhao, Jie Wang, and Jieping Ye. Safe screening with variational inequalities and its application to lasso. In *Proceedings of the 31st International Conference on Machine Learning (ICML)*, volume 32, pages 289–297. 2014.

18 Eugene Ndiaye, Olivier Fercoq, Alexandre Gramfort, and Joseph Salmon. Gap safe screening rules for sparsity enforcing penalties. *J. Mach. Learn. Res.*, 18(1):4671–4703, 2017.

19 Eugene Ndiaye, Olivier Fercoq, and Joseph Salmon. Screening Rules and its Complexity for Active Set Identification. *J. Convex Anal.*, 28(4):1053–1072, 2021.

20 Eugene Ndiaye, Tam Le, Olivier Fercoq, Joseph Salmon, and Ichiro Takeuchi. Safe grid search with optimal complexity. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, volume 97 of *Proceedings of Machine Learning Research*, pages 4771–4780. 2019.

21 Xianli Pan and Yitian Xu. A safe feature elimination rule for l1-regularized logistic regression. *IEEE Trans. Pattern Anal. Mach. Intell.*, 44(9):4544–4554, 2022.

22 Thu-Le Tran, Clément Elvira, Hong-Phuong Dang, and Cédric Herzet. Beyond gap screening for lasso by exploiting new dual cutting half-spaces. In *Proceedings of the European Signal Processing Conference (EUSIPCO)*, pages 2056–2060. 2022.

23 Hongmei Wang, Kun Jiang, and Yitian Xu. Sequential safe feature elimination rule for l1-regularized regression with kullback–leibler divergence. *Neural Netw.*, 155:523–535, 2022.

24 Jie Wang, Peter Wonka, and Jieping Ye. Lasso screening rules via dual polytope projection. *J. Mach. Learn. Res.*, 16:1063–1101, 2015.

25 Jie Wang, Jiayu Zhou, Jun Liu, Peter Wonka, and Jieping Ye. A safe screening rule for sparse logistic regression. In *NIPS'14: Proceedings of the 28th International Conference on Neural Information Processing Systems*, pages 1053–1061. Curran Associates, Inc., 2014.

26 Hiroaki Yamada and Makoto Yamada. Dynamic Sasvi: Strong safe screening for norm-regularized least squares. In *NIPS'21: Proceedings of the 35th International Conference on Neural Information Processing Systems*, pages 14645–14655. Curran Associates, Inc., 2021.

27 Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *J. R. Stat. Soc., Ser. B, Stat. Methodol.*, 67(2):301–320, 2005.