Projecting onto rectangular hyperbolic paraboloids in Hilbert space

Heinz H. Bauschke*, Manish Krishan Lal†, and Xianfu Wang‡

Dedicated to Henry Wolkowicz on the occasion of his 75th birthday.

Abstract

In $\mathbb{R}^3$, a hyperbolic paraboloid is a classical saddle-shaped quadric surface. Recently, Elser has modeled problems arising in Deep Learning using rectangular hyperbolic paraboloids in $\mathbb{R}^n$. Motivated by his work, we provide a rigorous analysis of the associated projection. In some cases, finding this projection amounts to finding a certain root of a quintic or cubic polynomial. We also observe when the projection is not a singleton and point out connections to graphical and set convergence.

2020 Mathematics Subject Classification: Primary 41A50, 90C26; Secondary 15A63, 46C05.

Keywords: cross, graphical convergence, rectangular hyperbolic paraboloid, projection onto a nonconvex set.

1 Introduction

Throughout this paper, we assume that

$$X \text{ is a real Hilbert space with an inner product } \langle\cdot,\cdot\rangle : X \times X \to \mathbb{R},$$

*Department of Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: heinz.bauschke@ubc.ca.

†Department of Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: manish.krishanlal@ubc.ca.

‡Department of Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: shawn.wang@ubc.ca.

and induced norm $\|\cdot\|$, and that $\alpha \in \mathbb{R} \smallsetminus \{0\}$ and $\beta > 0$. Define the β-weighted norm on the product space $X \times X \times \mathbb{R}$ by

$$(\forall (x,y,\gamma) \in X \times X \times \mathbb{R}) \ \|(x,y,\gamma)\| := \sqrt{\|x\|^2 + \|y\|^2 + \beta^2|\gamma|^2}.$$

Now define the set

$$C_\alpha := \{(x,y,\gamma) \in X \times X \times \mathbb{R} \mid \langle x,y \rangle = \alpha\gamma\}. \tag{1}$$

The set C_α is a special bilinear constraint set in optimization, and it corresponds to a rectangular (a.k.a. orthogonal) hyperbolic paraboloid in geometry [8]. Motivated by Deep Learning, Elser recently presented in [6] a formula for the projection $P_{C_\alpha}(x_0,y_0,\gamma_0)$ when $x_0 \neq \pm y_0$. However, complete mathematical justifications were not presented, and the case when $x_0 = \pm y_0$ was not considered. The goal of this paper is to provide a complete analysis of P_{C_α} that is applicable to all possible cases.

The paper is organized as follows. We collect auxiliary results in Section 2. Our main result is proved in Section 3 which also contains a numerical illustration. The formula for the projection onto the set C_α is presented in Section 4.

As usual, the distance function and projection mapping associated to C_α are denoted by $d_{C_\alpha}(x_0,y_0,\gamma_0) := \inf_{(x,y,\gamma)\in C_\alpha} \|(x,y,\gamma) - (x_0,y_0,\gamma_0)\|$ and $P_{C_\alpha}(x_0,y_0,\gamma_0) := \operatorname{argmin}_{(x,y,\gamma)\in C_\alpha} \|(x,y,\gamma) - (x_0,y_0,\gamma_0)\|$, respectively. We say that $x, x_0 \in X$ are *conically dependent* if there exists $s \geq 0$ such that $x = sx_0$ or $x_0 = sx$.

2 Auxiliary results

We start with some elementary properties of C_α, and justify the existence of projections onto these sets.

Proposition 2.1. The following hold:

(i) The set C_α is closed. If X is infinite-dimensional, then C_α is not weakly closed; in fact, $\overline{C_\alpha}^{\text{weak}} = X \times X \times \mathbb{R}$.

(ii) C_α is prox-regular in $X \times X \times \mathbb{R}$. Hence, for every point in $(x_0,y_0,\gamma_0) \in C_\alpha$, there exists a neighborhood such that the projection mapping P_{C_α} is single-valued.

Proof. (i): Clearly, C_α is closed. Thus assume that X is infinite-dimensional. By [3, Proposition 2.1], for every $\gamma \in \mathbb{R}$, $\overline{\{(x,y) \in X \times X |\ \langle x,y \rangle = \alpha\gamma\}}^{\text{weak}} = X \times X$. Thus,

$$X \times X \times \mathbb{R} = \bigcup_{\gamma\in\mathbb{R}} \left(\overline{\{(x,y) \in X \times X |\ \langle x,y \rangle = \alpha\gamma\}}^{\text{weak}} \times \{\gamma\} \right)$$

$$\subseteq \overline{\{(x,y,\gamma)\in X\times X\times\mathbb{R} \mid \langle x,y\rangle=\alpha\gamma\}}^{\text{weak}} \subseteq X\times X\times\mathbb{R}.$$

(ii): Set $F\colon X\times X\times\mathbb{R}\to\mathbb{R}\colon (x,y,\gamma)\mapsto\langle x,y\rangle-\alpha\gamma$. Then $C_\alpha=F^{-1}(0)$ and $\nabla F(x,y,\gamma)=(y,x,-\alpha)\neq(0,0,0)$ because $\alpha\neq 0$. The prox-regularity of C_α now follows from [9, Example 6.8] when $X=\mathbb{R}^n$ or from [4, Proposition 2.4] in the general case. Finally, the single-valuedness of the projection locally around every point in C_α follows from [4, Proposition 4.4]. ■

To study the projection onto C_α, it is convenient to introduce

$$\widetilde{C}_\alpha := \big\{(u,v,\gamma)\in X\times X\times\mathbb{R} \;\big|\; \|u\|^2-\|v\|^2=2\alpha\gamma\big\}, \tag{2}$$

which is the standard form of a rectangular hyperbolic paraboloid. Define a linear operator $A: X\times X\times\mathbb{R}\to X\times X\times\mathbb{R}$ by sending (u,v,γ) to (x,y,γ), where

$$x=\frac{u-v}{\sqrt{2}} \quad\text{and}\quad y=\frac{u+v}{\sqrt{2}}.$$

In terms of block matrix notation, we have

$$\begin{bmatrix} x\\ y\\ \gamma\end{bmatrix}=\begin{bmatrix}\frac{1}{\sqrt{2}}\mathrm{Id} & -\frac{1}{\sqrt{2}}\mathrm{Id} & 0\\ \frac{1}{\sqrt{2}}\mathrm{Id} & \frac{1}{\sqrt{2}}\mathrm{Id} & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix} u\\ v\\ \gamma\end{bmatrix} \Leftrightarrow \begin{bmatrix} u\\ v\\ \gamma\end{bmatrix}=\begin{bmatrix}\frac{1}{\sqrt{2}}\mathrm{Id} & \frac{1}{\sqrt{2}}\mathrm{Id} & 0\\ -\frac{1}{\sqrt{2}}\mathrm{Id} & \frac{1}{\sqrt{2}}\mathrm{Id} & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix} x\\ y\\ \gamma\end{bmatrix}.$$

Thus, we may and do identify A with its block matrix representation

$$A=\begin{bmatrix}\frac{1}{\sqrt{2}}\mathrm{Id} & -\frac{1}{\sqrt{2}}\mathrm{Id} & 0\\ \frac{1}{\sqrt{2}}\mathrm{Id} & \frac{1}{\sqrt{2}}\mathrm{Id} & 0\\ 0 & 0 & 1\end{bmatrix},$$

and we denote the adjoint of A by $A^\intercal$. Note that A corresponds to a rotation by $\pi/4$ about the γ-axis. The relationship between C_α and $\widetilde{C}_\alpha$ is summarized as follows.

Proposition 2.2. The following hold:

(i) A is a surjective isometry (i.e., a unitary operator): $AA^\intercal=A^\intercal A=\mathrm{Id}$.
(ii) $A\widetilde{C}_\alpha=C_\alpha$ and $\widetilde{C}_\alpha=A^\intercal C_\alpha$.
(iii) $P_{C_\alpha}=A\,P_{\widetilde{C}_\alpha}A^\intercal$.

Proof. It is straightforward to verify (i) and (ii). To show (iii), let $(x_0,y_0,\gamma_0)\in X\times X\times\mathbb{R}$. In view of (i) and (ii), we have $(x,y,\gamma)\in P_{C_\alpha}(x_0,y_0,\gamma_0)$ if and only if $(x,y,\gamma)\in C_\alpha$ and

$$\|(x,y,\gamma)-(x_0,y_0,\gamma_0)\|=d_{C_\alpha}(x_0,y_0,\gamma_0)=d_{A\widetilde{C}_\alpha}(x_0,y_0,\gamma_0)=d_{\widetilde{C}_\alpha}(A^\intercal[x_0,y_0,\gamma_0]^\intercal),$$

and this is equivalent to

$$\|A^\intercal[x,y,\gamma]^\intercal - A^\intercal[x_0,y_0,\gamma_0]^\intercal\| = d_{\widetilde{C}_\alpha}(A^\intercal[x_0,y_0,\gamma_0]^\intercal).$$

Since $A^\intercal[x,y,\gamma]^\intercal \in \widetilde{C}_\alpha$, this gives $A^\intercal[x,y,\gamma]^\intercal \in P_{\widetilde{C}_\alpha}(A^\intercal[x_0,y_0,\gamma_0]^\intercal)$, i.e., $[x,y,\gamma]^\intercal \in A\,P_{\widetilde{C}_\alpha}(A^\intercal[x_0,y_0,\gamma_0]^\intercal)$. The converse inclusion is proved similarly. ■

Exploiting the structure of $\widetilde{C}_\alpha$ is crucial for showing the existence of $P_{\widetilde{C}_\alpha}(u_0,v_0,\gamma_0)$ for every $(u_0,v_0,\gamma_0) \in X \times X \times \mathbb{R}$.

Proposition 2.3. (existence of the projection) Let $(u_0,v_0,\gamma_0) \in X \times X \times \mathbb{R}$. Then the minimization problem

$$\text{minimize} \quad f(u,v,\gamma) := \|u-u_0\|^2 + \|v-v_0\|^2 + \beta^2|\gamma-\gamma_0|^2 \tag{3a}$$

$$\text{subject to} \quad h(u,v,\gamma) := \|u\|^2 - \|v\|^2 - 2\alpha\gamma = 0 \tag{3b}$$

always has a solution, i.e., $P_{\widetilde{C}_\alpha}(u_0,v_0,\gamma_0) \neq \varnothing$. If $(u,v,\gamma) \in P_{\widetilde{C}_\alpha}(u_0,v_0,\gamma_0)$, then u,u_0 are conically dependent, and v,v_0 are also conically dependent.

Proof. We only illustrate the case when $u_0 \neq 0, v_0 \neq 0$, since the other cases are similar. We claim that the optimization problem is essentially 3-dimensional. To this end, we expand

$$f(u,v,\gamma) = \underbrace{\|u\|^2 - 2\langle u,u_0\rangle + \|u_0\|^2}+\underbrace{\|v\|^2 - 2\langle v,v_0\rangle + \|v_0\|^2}+\beta^2|\gamma-\gamma_0|^2. \tag{4}$$

The constraint

$$h(u,v,\gamma) = \|u\|^2 - \|v\|^2 - 2\alpha\gamma = 0$$

means that for the variables u,v only the norms $\|u\|$ and $\|v\|$ matter. With $\|u\|$ fixed, the Cauchy-Schwarz inequality in Hilbert space (see, e.g., [7]), shows that $-2\langle u,u_0\rangle$ in the left underbraced part of (4) will be smallest when u,u_0 are conically dependent. Similarly, for fixed $\|v\|$, the second underlined part in f will be smaller when $v = tv_0$ for some $t \geq 0$. It follows that the optimization problem given by (3) is equivalent to

$$\text{minimize} \quad g(s,t,\gamma) := (1-s)^2\|u_0\|^2 + (1-t)^2\|v_0\|^2 + \beta^2|\gamma-\gamma_0|^2 \tag{5a}$$

$$\text{subject to} \quad g_1(s,t,\gamma) := s^2\|u_0\|^2 - t^2\|v_0\|^2 - 2\alpha\gamma = 0, \quad s \geq 0, t \geq 0, \gamma \in \mathbb{R}. \tag{5b}$$

Because g is continuous and coercive, and g_1 is continuous, we conclude that the optimization problem (5) has a solution. ■

Next, we provide a result on set convergence and review graphical convergence, see, e.g., [1,9]. We shall need the *cross*

$$C := \{(x,y) \in X \times X \mid \langle x,y\rangle = 0\}, \tag{6}$$

which was studied in, e.g., [2], as well as

$$\widetilde{C} := \{(u,v) \in X \times X \mid \|u\|^2 - \|v\|^2 = 0\}. \tag{7}$$

Proposition 2.4. The following hold:

(i) $\lim_{\alpha\to 0} \widetilde{C}_\alpha = \widetilde{C} \times \mathbb{R}$.
(ii) $\lim_{\alpha\to 0} C_\alpha = C \times \mathbb{R}$.

Proof. (i): First we show that $\limsup_{\alpha\to 0} \widetilde{C}_\alpha \subseteq \widetilde{C} \times \mathbb{R}$. Let $(u_\alpha, v_\alpha, \gamma_\alpha) \to (u, v, \gamma)$ and $(u_\alpha, v_\alpha, \gamma_\alpha) \in \widetilde{C}_\alpha$ with $\alpha \to 0$. Then $\|u_\alpha\|^2 - \|v_\alpha\|^2 = 2\alpha\gamma_\alpha$ gives $\|u\|^2 - \|v\|^2 = 0$ when $\alpha \to 0$, so $(u, v, \gamma) \in \widetilde{C} \times \mathbb{R}$.

Next we show $\widetilde{C} \times \mathbb{R} \subseteq \liminf_{\alpha\to 0} \widetilde{C}_\alpha$. Let $(u, v, \gamma) \in \widetilde{C} \times \mathbb{R}$, i.e., $\|u\|^2 - \|v\|^2 = 0$ and $\gamma \in \mathbb{R}$. Let $\varepsilon > 0$. We consider three cases:

Case 1: $\gamma = 0$. Then $(u_\alpha, v_\alpha, 0) = (u, v, 0) \in \widetilde{C}_\alpha$ for every α.

Case 2: $\gamma \neq 0$ but $(u, v) = (0, 0)$. If $\alpha\gamma > 0$, take $(u_\alpha, 0, \gamma)$ with $\|u_\alpha\|^2 - 0 = \alpha\gamma$ so that $(u_\alpha, 0, \gamma) \in C_\alpha$; if $\alpha\gamma < 0$, take $(0, v_\alpha, \gamma)$ with $0 - \|v_\alpha\|^2 = \alpha\gamma$ so that $(0, v_\alpha, \gamma) \in C_\alpha$. Then

$$\|(u_\alpha, 0, \gamma) - (0, 0, \gamma)\| = \|u_\alpha\| = \sqrt{|\alpha\gamma|} < \varepsilon,$$

or

$$\|(0, v_\alpha, \gamma) - (0, 0, \gamma)\| = \|v_\alpha\| = \sqrt{|\alpha\gamma|} < \varepsilon,$$

if $|\alpha| < \varepsilon^2/|\gamma|$.

Case 3: $\gamma \neq 0$ and $(u, v) \neq (0, 0)$. Take $\alpha \in \mathbb{R}$ such that

$$|\alpha| < \min\left\{\frac{\varepsilon\|(u, v)\|}{|\gamma|}, \frac{\|(u, v)\|^2}{|\gamma|}\right\},$$

and set

$$\lambda := \frac{\alpha\gamma}{\|(u, v)\|^2}.$$

Then

$$|\lambda| = \frac{|\alpha\gamma|}{\|(u, v)\|^2} < 1.$$

Now set

$$u_\alpha := \sqrt{1+\lambda}\,u, \quad v_\alpha := \sqrt{1-\lambda}\,v.$$

Then

$$\begin{aligned}\|u_\alpha\|^2 - \|v_\alpha\|^2 &= (1+\lambda)\|u\|^2 - (1-\lambda)\|v\|^2 \\ &= \lambda(\|u\|^2 + \|v\|^2) = \alpha\gamma,\end{aligned}$$

so that $(u_\alpha, v_\alpha, \gamma) \in \widetilde{C}_\alpha$ and

$$\begin{aligned}\|(u_\alpha, v_\alpha, \gamma) - (u, v, \gamma)\| &= \sqrt{(\sqrt{1+\lambda}-1)^2\|u\|^2 + (\sqrt{1-\lambda}-1)^2\|v\|^2} \\ &= \sqrt{\frac{\lambda^2}{(1+\sqrt{1+\lambda})^2}\|u\|^2 + \frac{\lambda^2}{(1+\sqrt{1-\lambda})^2}\|v\|^2} \\ &\le \sqrt{\lambda^2(\|u\|^2 + \|v\|^2)} = |\lambda|\|(u,v)\| < \varepsilon.\end{aligned}$$

(ii): This follows from (i) because that $C_\alpha = A\widetilde{C}_\alpha$ and $C \times \mathbb{R} = A(\widetilde{C} \times \mathbb{R})$ and that A is an isometry. See also [9, Theorem 4.26]. ■

Definition 2.5. (graphical limits of mappings) (See [9, Definition 5.32].) For a sequence of set-valued mappings $S^k : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, we say S^k converges graphically to S, in symbols $S^k \xrightarrow{g} S$, if for every $x \in \mathbb{R}^n$ one has

$$\bigcup_{\{x^k \to x\}} \limsup_{k\to\infty} S^k(x^k) \subseteq S(x) \subseteq \bigcup_{\{x^k \to x\}} \liminf_{k\to\infty} S^k(x^k).$$

Fact 2.6. (Rockafellar–Wets) (See [9, Example 5.35].) For closed subsets sets S^k, S of $\mathbb{R}^n$, one has $P_{S^k} \xrightarrow{g} P_S$ if and only if $S^k \to S$.

We are now ready for our main results which we will derive in the next section.

3 Projection onto a rectangular hyperbolic paraboloid

We begin with projections onto rectangular hyperbolic paraboloids. In view of Proposition 2.2(iii), to find P_{C_α} it suffices to find $P_{\widetilde{C}_\alpha}$. That is, for every $(u_0, v_0, \gamma_0) \in X \times X \times \mathbb{R}$, we need to solve:

$$\min_{u,v,\gamma} \quad f(u,v,\gamma) := \|u-u_0\|^2 + \|v-v_0\|^2 + \beta^2|\gamma-\gamma_0|^2 \tag{8a}$$

$$\text{subject to} \quad h(u,v,\gamma) := \|u\|^2 - \|v\|^2 - 2\alpha\gamma = 0. \tag{8b}$$

Theorem 3.1. Let $(u_0, v_0, \gamma_0) \in X \times X \times \mathbb{R}$. Then the following hold:

(i) When $u_0 \neq 0, v_0 \neq 0$, then

$$P_{\widetilde{C}_\alpha}(u_0, v_0, \gamma_0) = \left\{ \left(\frac{u_0}{1+\lambda}, \frac{v_0}{1-\lambda}, \gamma_0 + \frac{\lambda\alpha}{\beta^2} \right) \right\}, \tag{9}$$

where the unique $\lambda \in]-1,1[$ solves the following (essentially) quintic equation

$$g(\lambda) := \frac{(\lambda^2+1)p - 2\lambda q}{(1-\lambda^2)^2} - \frac{2\lambda\alpha^2}{\beta^2} - 2\alpha\gamma_0 = 0, \tag{10}$$

and where $p := \|u_0\|^2 - \|v_0\|^2$ and $q := \|u_0\|^2 + \|v_0\|^2$.

(ii) When $u_0 = 0, v_0 \neq 0$, we have:

(a) If $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) < -\frac{\|v_0\|^2}{8}$, then

$$P_{\widetilde{C}_\alpha}(0, v_0, \gamma_0) = \left\{ \left(0, \frac{v_0}{1-\lambda}, \gamma_0 + \frac{\lambda\alpha}{\beta^2}\right) \right\}, \tag{11}$$

for a unique $\lambda \in]-1,1[$ that solves the (essentially) cubic equation

$$g_1(\lambda) := \frac{\|v_0\|^2}{(1-\lambda)^2} + \frac{2\lambda\alpha^2}{\beta^2} + 2\alpha\gamma_0 = 0. \tag{12}$$

(b) If $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) \geq -\frac{\|v_0\|^2}{8}$, then

$$P_{\widetilde{C}_\alpha}(0, v_0, \gamma_0) = \left\{ \left(u, \frac{v_0}{2}, \gamma_0 - \frac{\alpha}{\beta^2}\right) \;\middle|\; \|u\| = \sqrt{2\alpha\left(\gamma_0 - \frac{\alpha}{\beta^2}\right) + \frac{\|v_0\|^2}{4}}, u \in X \right\}, \tag{13}$$

which is a singleton if and only if $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) = -\frac{\|v_0\|^2}{8}$.

(iii) When $u_0 \neq 0, v_0 = 0$, we have:

(a) If $\alpha(\gamma_0 + \frac{\alpha}{\beta^2}) > \frac{\|u_0\|^2}{8}$, then

$$P_{\widetilde{C}_\alpha}(u_0, 0, \gamma_0) = \left\{ \left(\frac{u_0}{1+\lambda}, 0, \gamma_0 + \frac{\lambda\alpha}{\beta^2}\right) \right\} \tag{14}$$

for a unique $\lambda \in]-1,1[$ that solves the (essentially) cubic equation

$$g_2(\lambda) := \frac{\|u_0\|^2}{(1+\lambda)^2} - \frac{2\lambda\alpha^2}{\beta^2} - 2\alpha\gamma_0 = 0. \tag{15}$$

(b) If $\alpha(\gamma_0 + \frac{\alpha}{\beta^2}) \leq \frac{\|u_0\|^2}{8}$, then

$$P_{\widetilde{C}_\alpha}(u_0, 0, \gamma_0) = \left\{ \left(\frac{u_0}{2}, v, \gamma_0 + \frac{\alpha}{\beta^2}\right) \;\middle|\; \|v\| = \sqrt{-2\alpha\left(\gamma_0 + \frac{\alpha}{\beta^2}\right) + \frac{\|u_0\|^2}{4}}, v \in X \right\}, \tag{16}$$

which is a singleton if and only if $\alpha(\gamma_0 + \frac{\alpha}{\beta^2}) = \frac{\|u_0\|^2}{8}$.

(iv) When $u_0 = 0, v_0 = 0$, we have:

(a) If $\alpha\gamma_0 > \frac{\alpha^2}{\beta^2}$, then the projection is the non-singleton set

$$P_{\widetilde{C}_\alpha}(0,0,\gamma_0) = \left\{ \left(u, 0, \gamma_0 - \frac{\alpha}{\beta^2}\right) \;\middle|\; \|u\| = \sqrt{2\alpha\left(\gamma_0 - \frac{\alpha}{\beta^2}\right)}, u \in X \right\}. \tag{17}$$

(b) If $|\alpha\gamma_0| \le \frac{\alpha^2}{\beta^2}$, then

$$P_{\widetilde{C}_\alpha}(0,0,\gamma_0) = \{(0,0,0)\}. \tag{18}$$

(c) If $\alpha\gamma_0 < -\frac{\alpha^2}{\beta^2}$, then the projection is the non-singleton set

$$P_{\widetilde{C}_\alpha}(0,0,\gamma_0) = \left\{ \left(0, v, \gamma_0 + \frac{\alpha}{\beta^2}\right) \;\middle|\; \|v\| = \sqrt{-2\alpha\left(\gamma_0 + \frac{\alpha}{\beta^2}\right)}, v \in X \right\}. \tag{19}$$

Proof. Observe that $\nabla f(u,v,\gamma) = (2(u-u_0), 2(v-v_0), 2\beta^2(\gamma-\gamma_0))$ and $\nabla h(u,v,\gamma) = (2u, -2v, -2\alpha)$. Since $\alpha \neq 0$, we have $\forall (u,v,\gamma) \in X \times X \times \mathbb{R}$, $\nabla h(u,v,\gamma) \neq 0$. Using [5, Proposition 4.1.1], we obtain the following KKT optimality conditions of (8):

$$(1+\lambda)u = u_0 \tag{20a}$$
$$(1-\lambda)v = v_0 \tag{20b}$$
$$\beta^2(\gamma - \gamma_0) - \lambda\alpha = 0 \tag{20c}$$
$$\|u\|^2 - \|v\|^2 - 2\alpha\gamma = 0 \tag{20d}$$

where $\lambda \in \mathbb{R}$ is the Lagrange multiplier.

The proofs of (i)–(iv) are presented in Section 3.1–Section 3.4 below.

3.1 Case (i): $u_0 \neq 0, v_0 \neq 0$

Proof. Because $u_0 \neq 0, v_0 \neq 0$, we obtain $\lambda \neq \pm 1$. Solving (20a), (20b) and (20c) gives $u = \frac{u_0}{(1+\lambda)}$, $v = \frac{v_0}{(1-\lambda)}$ and $\gamma = \gamma_0 + \frac{\lambda\alpha}{\beta^2}$. By Proposition 2.3, $1+\lambda > 0$ and $1-\lambda > 0$, i.e., $\lambda \in \,]-1,1[$. Substituting u and v back into equation (20d), we get the (essentially) quintic equation (10). Using also $p < q$ and $q > 0$, we have

$$\begin{aligned}
(\forall \lambda \in \,]-1,1[)\; g'(\lambda) &= \frac{2}{(1-\lambda^2)^3}\big(-q(1+3\lambda^2) + p(\lambda^3+3\lambda)\big) - 2\frac{\alpha^2}{\beta^2} \\
&< \frac{2}{(1-\lambda^2)^3}\big(-q(1+3\lambda^2) + q(\lambda^3+3\lambda)\big) - 2\frac{\alpha^2}{\beta^2}
\end{aligned}$$

$$= \frac{2q(\lambda-1)^3}{(1-\lambda^2)^3} - 2\frac{\alpha^2}{\beta^2} = \frac{-2q}{(1+\lambda)^3} - 2\frac{\alpha^2}{\beta^2}$$
$$< 0;$$

hence, g is strictly decreasing. Moreover, $g(-1) = +\infty$, $g(1) = -\infty$ and g is continuous on $]-1,1[$. Thus, $g(\lambda) = 0$ has unique zero in $]-1,1[$. ■

3.2 Case (ii): $u_0 = 0, v_0 \neq 0$

Proof. When $u_0 = 0$, the objective function is

$$f(u,v,\gamma) = \|u\|^2 + \|v - v_0\|^2 + \beta^2|\gamma - \gamma_0|^2,$$

and the KKT optimality conditions (20) become

$$(1+\lambda)u = 0 \tag{21a}$$
$$(1-\lambda)v = v_0 \tag{21b}$$
$$\gamma = \gamma_0 + \frac{\lambda\alpha}{\beta^2} \tag{21c}$$
$$\|u\|^2 - \|v\|^2 = 2\alpha\gamma. \tag{21d}$$

Then (21a) gives

$$1+\lambda = 0 \ \text{ or } \ u = 0. \tag{22}$$

Because $v_0 \neq 0$, we have $1-\lambda \neq 0$, so that

$$v = \frac{v_0}{1-\lambda}. \tag{23}$$

By Proposition 2.3, $\lambda < 1$.

Our analysis is divided into the following three situations:

Situation 1: $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) < -\frac{\|v_0\|^2}{8}$.

In view of (22), we analyze two cases.

Case 1: $1+\lambda = 0$, i.e., $\lambda = -1$. By (23), $v = \frac{v_0}{2}$, and then (21d) and (21c) give

$$\|u\|^2 = 2\alpha\gamma + \frac{\|v_0\|^2}{4} = 2\alpha\Big(\gamma_0 - \frac{\alpha}{\beta^2}\Big) + \frac{\|v_0\|^2}{4} < 0,$$

which is absurd.

Case 2: $u = 0$. By (21d), $-\|v\|^2 = 2\alpha\gamma$, together with (23) and (21c), we have

$$g_1(\lambda) := \frac{\|v_0\|^2}{(1-\lambda)^2} + 2\alpha\Big(\gamma_0 + \frac{\lambda\alpha}{\beta^2}\Big) = 0.$$

As

$$g_1'(\lambda) = \frac{2\|v_0\|^2}{(1-\lambda)^3} + \frac{2\alpha^2}{\beta^2} > 0 \quad \text{on} \quad]-\infty,1[,$$

g_1 is strictly increasing on $]-\infty,1[$. Moreover, $g_1(1) = +\infty$ and

$$g_1(-1) = \frac{\|v_0\|^2}{4} + 2\alpha\Big(\gamma_0 - \frac{\alpha}{\beta^2}\Big) < 0.$$

Because g_1 is strictly increasing and continuous, by the Intermediate Value Theorem, there exists a unique $\lambda \in\,]-1,1[$ such that $g_1(\lambda) = 0$. Hence, the possible optimal solution is given by

$$\Big(0, \frac{v_0}{1-\lambda}, \gamma_0 + \frac{\lambda\alpha}{\beta^2}\Big), \tag{24}$$

where $g_1(\lambda) = 0$ and $\lambda \in\,]-1,1[$.

Combining Case 1 and Case 2, we obtain that (24) is the unique projection.

Situation 2: $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) > -\frac{\|v_0\|^2}{8}$.

In view of (22), we consider two cases:

Case 1: $1 + \lambda = 0$, i.e., $\lambda = -1$. By (23), $v = \frac{v_0}{2}$, and then (21d) and (21c) give

$$\|u\|^2 = 2\alpha\gamma + \frac{\|v_0\|^2}{4} = 2\alpha\Big(\gamma_0 - \frac{\alpha}{\beta^2}\Big) + \frac{\|v_0\|^2}{4} > 0.$$

The possible optimal value is attained at

$$\Big(u, \frac{v_0}{2}, \gamma_0 - \frac{\alpha}{\beta^2}\Big) \tag{25}$$

with $\|u\|^2 = 2\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) + \frac{\|v_0\|^2}{4}$ such that

$$f\Big(u, \frac{v_0}{2}, \gamma_0 - \frac{\alpha}{\beta^2}\Big) = 2\alpha\gamma_0 - \frac{\alpha^2}{\beta^2} + \frac{\|v_0\|^2}{2}. \tag{26}$$

Case 2: $u = 0$. By (21d), $-\|v\|^2 = 2\alpha\gamma$, together with (21c), we have

$$g_1(\lambda) := \frac{\|v_0\|^2}{(1-\lambda)^2} + 2\alpha\Big(\gamma_0 + \frac{\lambda\alpha}{\beta^2}\Big) = 0.$$

As

$$g_1'(\lambda) = \frac{2\|v_0\|^2}{(1-\lambda)^3} + \frac{2\alpha^2}{\beta^2} > 0 \quad \text{on }]-\infty, 1[,$$

g_1 is strictly increasing. Observe that

$$g_1(-1) = \frac{\|v_0\|^2}{4} + 2\alpha\Big(\gamma_0 - \frac{\alpha}{\beta^2}\Big) > 0,$$

and $g_1(-\infty) = -\infty$. By the Intermediate Value Theorem, there exists a unique $\lambda \in]-\infty, -1[$ such that $g_1(\lambda) = 0$ because g_1 is strictly increasing and continuous. The possible optimal value is attained at (recall (23))

$$\Big(0, \frac{v_0}{1-\lambda}, \gamma_0 + \frac{\lambda\alpha}{\beta^2}\Big)$$

with

$$f\Big(0, \frac{v_0}{1-\lambda}, \gamma_0 + \frac{\lambda\alpha}{\beta^2}\Big) = \frac{\lambda^2\|v_0\|^2}{(1-\lambda)^2} + \frac{\lambda^2\alpha^2}{\beta^2}, \tag{27}$$

where λ is the unique solution of

$$g_1(\lambda) := \frac{\|v_0\|^2}{(1-\lambda)^2} + 2\alpha\Big(\gamma_0 + \frac{\lambda\alpha}{\beta^2}\Big) = 0 \quad \text{in }]-\infty, -1[. \tag{28}$$

Because both Case 1 and Case 2 may occur, we have to compare possible optimal objective function values, namely, (26) and (27). We claim that Case 1 wins, i.e.,

$$2\alpha\gamma_0 - \frac{\alpha^2}{\beta^2} + \frac{\|v_0\|^2}{2} < \frac{\lambda^2\|v_0\|^2}{(1-\lambda)^2} + \frac{\lambda^2\alpha^2}{\beta^2}. \tag{29}$$

In view of (28), we have

$$0 < \frac{\|v_0\|^2}{(1-\lambda)^2} = -2\alpha\Big(\gamma_0 + \frac{\lambda\alpha}{\beta^2}\Big), \text{ and so } \alpha\Big(\gamma_0 + \frac{\lambda\alpha}{\beta^2}\Big) < 0. \tag{30}$$

To show (29), we shall reformulate it in equivalent forms:

$$\Big(\lambda^2 - \frac{(1-\lambda)^2}{2}\Big)\frac{\|v_0\|^2}{(1-\lambda)^2} + (1+\lambda^2)\frac{\alpha^2}{\beta^2} > 2\alpha\gamma_0,$$

which is

$$\frac{\lambda^2+2\lambda-1}{2}\Big(-2\alpha\Big(\gamma_0+\frac{\lambda\alpha}{\beta^2}\Big)\Big)+(1+\lambda^2)\frac{\alpha^2}{\beta^2}>2\alpha\gamma_0$$

by (30). After simplifications, this reduces to

$$\frac{\alpha^2}{\beta^2}(1+\lambda)^2(1-\lambda)>\alpha\gamma_0(1+\lambda)^2.$$

Since $\lambda+1<0$, this is equivalent to

$$\frac{\alpha^2}{\beta^2}(1-\lambda)>\alpha\gamma_0, \quad \text{i.e.,} \ \alpha\Big(\gamma_0+\frac{\lambda\alpha}{\beta^2}\Big)<\frac{\alpha^2}{\beta^2},$$

which obviously holds because of (30) and $\alpha^2/\beta^2>0$.

Hence, equation (25) of Case 1 gives the optimal solution.

Situation 3:

$$\alpha\Big(\gamma_0-\frac{\alpha}{\beta^2}\Big)=-\frac{\|v_0\|^2}{8}. \tag{31}$$

We again consider two cases.

Case 1: $1+\lambda=0$, i.e., $\lambda=-1$. By (21b), $v=\frac{v_0}{2}$ and then (21d) and (21c) give

$$\|u\|^2=2\alpha\gamma+\frac{\|v_0\|^2}{4}=2\alpha\Big(\gamma_0-\frac{\alpha}{\beta^2}\Big)+\frac{\|v_0\|^2}{4}=0,$$

so $u=0$. The possible optimal value is attained at

$$\Big(0,\frac{v_0}{2},\gamma_0-\frac{\alpha}{\beta^2}\Big) \tag{32}$$

with

$$f\Big(0,\frac{v_0}{2},\gamma_0-\frac{\alpha}{\beta^2}\Big)=\frac{\|v_0\|^2}{4}+\frac{\alpha^2}{\beta^2}.$$

Case 2: $u=0$. By (21d), $-\|v\|^2=2\alpha\gamma$, together with (21c), we have

$$g_2(\lambda):=\frac{\|v_0\|^2}{(1-\lambda)^2}+2\alpha\Big(\gamma_0+\frac{\lambda\alpha}{\beta^2}\Big)=0.$$

By (31),

$$g_2(-1)=\frac{\|v_0\|^2}{4}+2\alpha\Big(\gamma_0-\frac{\alpha}{\beta^2}\Big)=0.$$

As

$$g_2'(\lambda) = \frac{2\|v_0\|^2}{(1-\lambda)^3} + \frac{2\alpha^2}{\beta^2} > 0 \quad \text{on} \quad]-\infty, 1[,$$

g_2 is strictly increasing and continuous on $]-\infty, 1[$, so $\lambda = -1$ is the unique solution in $]-\infty, 1[$. Then the possible optimal value is attained at

$$\Big(0, \frac{v_0}{2}, \gamma_0 - \frac{\alpha}{\beta^2}\Big)$$

with

$$f\Big(0, \frac{v_0}{2}, \gamma_0 - \frac{\alpha}{\beta^2}\Big) = \frac{\|v_0\|^2}{4} + \frac{\alpha^2}{\beta^2}. \tag{33}$$

Therefore, Case 1 and Case 2 give exactly the same solution. The optimal solution is given by (32), and it can be recovered by (25), the optimal solution of Situation 2. ■

3.3 Case (iii): $u_0 \neq 0, v_0 = 0$

Proof. The minimization problem now is

$$\text{minimize} \quad f(u, v, \gamma) = \|u_0 - u\|^2 + \|v\|^2 + \beta^2 |\gamma_0 - \gamma|^2 \tag{34a}$$

$$\text{subject to} \quad \|u\|^2 - \|v\|^2 = 2\alpha\gamma. \tag{34b}$$

Rewrite it as

$$\text{minimize} \quad f(u, v, \gamma) = \|v\|^2 + \|u_0 - u\|^2 + \beta^2 |\gamma_0 - \gamma|^2 \tag{35a}$$

$$\text{subject to} \quad \|v\|^2 - \|u\|^2 = 2(-\alpha)\gamma. \tag{35b}$$

Luckily, we can apply Section 3.2 for the point $(0, u_0, \gamma_0)$ and parameter $-\alpha$. More precisely, when $-\alpha(\gamma_0 - \frac{-\alpha}{\beta^2}) < -\frac{\|u_0\|^2}{8}$, the optimal solution to (35) is

$$\Big(0, \frac{u_0}{1-\tilde{\lambda}}, \gamma_0 + \frac{\tilde{\lambda}(-\alpha)}{\beta^2}\Big)$$

where $\tilde{g}_2(\tilde{\lambda}) = 0$, $\tilde{\lambda} \in]-1, 1[$, and

$$\tilde{g}_2(\tilde{\lambda}) = \frac{\|u_0\|^2}{(1-\tilde{\lambda})^2} + 2(-\alpha)\Big(\gamma_0 - \frac{\tilde{\lambda}\alpha}{\beta^2}\Big) = 0.$$

Put $\lambda = -\tilde{\lambda}$. Simplifications give: when $\alpha(\gamma_0 + \frac{\alpha}{\beta^2}) > \frac{\|u_0\|^2}{8}$, the optimal solution to (35) is

$$\Big(0, \frac{u_0}{1+\lambda}, \gamma_0 + \frac{\lambda\alpha}{\beta^2}\Big) \tag{36}$$

where $g_2(\lambda) = 0$, $\lambda \in \,]-1,1[$, and

$$g_2(\lambda) := \tilde{g}_2(-\lambda) = \frac{\|u_0\|^2}{(1+\lambda)^2} - 2\alpha\Big(\gamma_0 + \frac{\lambda\alpha}{\beta^2}\Big) = 0.$$

Switching the first and second components in (36) gives the optimal solution to (34).

When $-\alpha(\gamma_0 - \frac{-\alpha}{\beta^2}) \geq -\frac{\|u_0\|^2}{8}$, the optimal solution to (35) is

$$\Big(v, \frac{u_0}{2}, \gamma_0 - \frac{-\alpha}{\beta^2}\Big)$$

with

$$\|v\|^2 = 2(-\alpha)\Big(\gamma_0 - \frac{-\alpha}{\beta^2}\Big) + \frac{\|u_0\|^2}{4}.$$

That is, when $\alpha(\gamma_0 + \frac{\alpha}{\beta^2}) \leq \frac{\|u_0\|^2}{8}$, the optimal solution to (35) is

$$\Big(v, \frac{u_0}{2}, \gamma_0 + \frac{\alpha}{\beta^2}\Big) \tag{37}$$

with

$$\|v\|^2 = -2\alpha\Big(\gamma_0 + \frac{\alpha}{\beta^2}\Big) + \frac{\|u_0\|^2}{4}.$$

Switching the first and second components in (37) gives the optimal solution to (34). ■

3.4 Case (iv): $u_0 = v_0 = 0$

Proof. The objective function is $f(u,v,\gamma) = \|u\|^2 + \|v\|^2 + \beta^2|\gamma - \gamma_0|^2$, and the KKT optimality conditions (20) become

$$(1+\lambda)u = 0, \tag{38a}$$
$$(1-\lambda)v = 0, \tag{38b}$$
$$\gamma = \gamma_0 + \frac{\lambda\alpha}{\beta^2}, \tag{38c}$$
$$\|u\|^2 - \|v\|^2 = 2\alpha\gamma. \tag{38d}$$

We shall consider three cases:

(i) $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) > 0$; hence, $\gamma_0 - \frac{\alpha}{\beta^2} \neq 0$.
(ii) $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) = 0$; hence, $\gamma_0 - \frac{\alpha}{\beta^2} = 0$.

(iii) $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) < 0$; hence, $\gamma_0 - \frac{\alpha}{\beta^2} \neq 0$.

For each item (i)–(iii), we will apply (38):
Case 1: $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) > 0$. By (38a), we have $\lambda = -1$ or $u = 0$. We consider two subcases.

Subcase 1: $\lambda = -1$. Using (38b), (38c) and (38d), we obtain $v = 0$, $\gamma = \gamma_0 - \frac{\alpha}{\beta^2}$, and

$$\|u\|^2 = 2\alpha\Big(\gamma_0 - \frac{\alpha}{\beta^2}\Big). \tag{39}$$

Therefore, the candidate for the solution is $(u, 0, \gamma_0 - \frac{\alpha}{\beta^2})$ with u given by (39) and its objective function value is

$$f\Big(u, 0, \gamma_0 - \frac{\alpha}{\beta^2}\Big) = 2\alpha\Big(\gamma_0 - \frac{\alpha}{\beta^2}\Big) + 0 + \beta^2\Big(\frac{-\alpha}{\beta^2}\Big)^2 = 2\alpha\gamma_0 - \frac{\alpha^2}{\beta^2}. \tag{40}$$

Subcase 2: $u = 0$. Using (38b)–(38d), we obtain $-\|v\|^2 = 2\alpha\gamma, \gamma = \gamma_0 + \frac{\lambda\alpha}{\beta^2}$ and $(1 - \lambda)v = 0$. We have to consider two further cases: $1 - \lambda = 0$ or $v = 0$.

(i) $v = 0$. We get $-(0)^2 = 2\alpha\gamma \Rightarrow \gamma = 0$ because $\alpha \neq 0$. This gives a possible solution $(0, 0, 0)$ with function value

$$f(0, 0, 0) = \|u\|^2 + \|v\|^2 + \beta^2|\gamma - \gamma_0|^2 = \beta^2\gamma_0^2. \tag{41}$$

(ii) $\lambda = 1$. We have $\gamma = \gamma_0 + \frac{\alpha}{\beta^2}$ and $-\|v\|^2 = 2\alpha(\gamma_0 + \frac{\alpha}{\beta^2})$. So, $0 \leq \|v\|^2 = -2\alpha(\gamma_0 + \frac{\alpha}{\beta^2})$. However,

$$-2\alpha\Big(\gamma_0 + \frac{\alpha}{\beta^2}\Big) = \underbrace{-2\alpha\Big(\gamma_0 - \frac{\alpha}{\beta^2}\Big)}_{<0} - \frac{4\alpha^2}{\beta^2} < 0 \tag{42}$$

because $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) > 0$. This contradiction shows $\lambda = 1$ does not happen.

We now compare objective function values (40) and (41):

$$2\alpha\gamma_0 - \frac{\alpha^2}{\beta^2} < \beta^2\gamma_0^2 \Leftrightarrow \beta^2\gamma_0^2 + \frac{\alpha^2}{\beta^2} - 2\alpha\gamma_0 > 0 \Leftrightarrow \Big(\beta\gamma_0 - \frac{\alpha}{\beta}\Big)^2 > 0 \Leftrightarrow \beta^2\Big(\gamma_0 - \frac{\alpha}{\beta^2}\Big)^2 > 0,$$

which holds because $\gamma_0 - \frac{\alpha}{\beta^2} \neq 0$. Hence, the optimal solution is $(u, 0, \gamma_0 - \frac{\alpha}{\beta^2})$ with $\|u\| = \sqrt{2\alpha(\gamma_0 - \frac{\alpha}{\beta^2})}$. That is,

$$P_{\widetilde{C}_2}(0, 0, \gamma_0) = \Big\{\Big(u, 0, \gamma_0 - \frac{\alpha}{\beta^2}\Big) \;\Big|\; \|u\| = \sqrt{2\alpha\Big(\gamma_0 - \frac{\alpha}{\beta^2}\Big)}\Big\}.$$

Case 2: $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) = 0$; hence, $\gamma_0 - \frac{\alpha}{\beta^2} = 0$. By (38a), we have two subcases to consider.
Subcase 1: $\lambda = -1$. We have $v = 0$, $\gamma = \gamma_0 - \frac{\alpha}{\beta^2} = 0$, $\|u\|^2 = 2\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) = 0$. The possible solution is $(0,0,0)$.
Subcase 2: $u = 0$. We have $-\|v\|^2 = 2\alpha\gamma$ and $\gamma = \gamma_0 + \frac{\lambda\alpha}{\beta^2}$. By (38b), $v = 0$ or $\lambda = 1$. This requires us to consider two further cases. For $v = 0$, we get $\gamma = 0$, which gives a possible solution $(0,0,0)$. For $\lambda = 1$, we get $\gamma = \gamma_0 + \frac{\alpha}{\beta^2}$, $\|v\|^2 = -2\alpha(\gamma_0 + \frac{\alpha}{\beta^2}) = \frac{-4\alpha^2}{\beta^2} < 0$, which is impossible, i.e., $\lambda = 1$ does not happen.

Both **Subcase 1** and **Subcase 2** give the same solution $(0,0,0)$. Therefore, we have the optimal solution is $(0,0,0)$, when $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) = 0$; equivalently, when $\gamma_0 = \frac{\alpha}{\beta^2}$.

Case 3: $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) < 0$. In view of (38a), we have $\lambda = -1$ or $u = 0$. We show that $\lambda = -1$ can't happen. Indeed, when $\lambda = -1$, by (38b)–(38c), we have $v = 0$, $\gamma = \gamma_0 - \frac{\alpha}{\beta^2}$, and $0 \le \|u\|^2 = 2\alpha\gamma = 2\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) < 0$, which is impossible. Therefore, we consider only the case $u = 0$. Then (38b)–(38d) yield $\|v\|^2 = -2\alpha\gamma$, $\gamma = \gamma_0 + \frac{\lambda\alpha}{\beta^2}$, and $(1-\lambda)v = 0$, which requires us to consider two further cases.
Subcase 1: $v = 0$. Then $\gamma = 0$. The possible optimal solution is $(0,0,0)$ and its objective function value is

$$f(0,0,0) = \|u\|^2 + \|v\|^2 + \beta^2|\gamma - \gamma_0|^2 = \beta^2\gamma_0^2. \tag{43}$$

Subcase 2: $\lambda = 1$. Then $u = 0$, $\gamma = \gamma_0 + \frac{\alpha}{\beta^2}$, and $-\|v\|^2 = 2\alpha(\gamma_0 + \frac{\alpha}{\beta^2})$. We consider three additional cases based on the sign of $\alpha(\gamma_0 + \frac{\alpha}{\beta^2})$.

(i) $\alpha(\gamma_0 + \frac{\alpha}{\beta^2}) > 0$. This case never happens because the relation $0 \ge -\|v\|^2 = 2\alpha(\gamma_0 + \frac{\alpha}{\beta^2}) > 0$ is absurd.
(ii) $\alpha(\gamma_0 + \frac{\alpha}{\beta^2}) = 0$. As $\alpha \neq 0$, we have $\gamma_0 + \frac{\alpha}{\beta^2} = 0$. This gives $\gamma = 0, u = 0$ and $v = 0$. So the possible optimal solution is $(0,0,0)$.
(iii) $\alpha(\gamma_0 + \frac{\alpha}{\beta^2}) < 0$. We have $\gamma_0 + \frac{\alpha}{\beta^2} \neq 0$. The possible optimal solution is $(0, v, \gamma_0 + \frac{\alpha}{\beta^2})$ with $\|v\| = \sqrt{-2\alpha(\gamma_0 + \frac{\alpha}{\beta^2})}$ and function value

$$f\left(0, v, \gamma_0 + \frac{\alpha}{\beta^2}\right) = -2\alpha\left(\gamma_0 + \frac{\alpha}{\beta^2}\right) + \beta^2\left(\frac{\alpha}{\beta^2}\right)^2 = -2\alpha\gamma_0 - \frac{\alpha^2}{\beta^2}. \tag{44}$$

Both (i) and (ii) imply that $(0,0,0)$ from **Subcase 1** is the only optimal solution, when $\alpha^2/\beta^2 > \alpha\gamma_0 \ge -\alpha^2/\beta^2$.

When $\alpha\gamma_0 < -\frac{\alpha^2}{\beta^2}$, both **Subcase 1** and **Subcase 2** happen. We have to compare objectives (43) and (44). We claim $f(0,v,\gamma_0+\frac{\alpha}{\beta^2}) < f(0,0,0)$. Indeed, this is equivalent to

$$-2\alpha\gamma_0 - \frac{\alpha^2}{\beta^2} < \beta^2\gamma_0^2 \Leftrightarrow \beta^2\gamma_0^2 + 2\alpha\gamma_0 + \frac{\alpha^2}{\beta^2} > 0 \Leftrightarrow \Big(\beta\gamma_0 + \frac{\alpha}{\beta}\Big)^2 > 0 \Leftrightarrow \beta^2\Big(\gamma_0 + \frac{\alpha}{\beta^2}\Big)^2 > 0$$

which holds because $\gamma_0 + \frac{\alpha}{\beta^2} \neq 0$. Therefore, the optimal solution is $(0,v,\gamma_0+\frac{\alpha}{\beta^2})$ with $\|v\| = \sqrt{-2\alpha(\gamma_0+\frac{\alpha}{\beta^2})}$, i.e.,

$$P_{\widetilde{C}_2}(0,0,\gamma_0) = \left\{ \Big(0,v,\gamma_0 + \frac{\alpha}{\beta^2}\Big) \;\Big|\; \|v\| = \sqrt{-2\alpha\Big(\gamma_0 + \frac{\alpha}{\beta^2}\Big)} \right\}$$

when $\alpha\gamma_0 < \frac{-\alpha^2}{\beta^2}$. ■

Altogether, Section 3.1–Section 3.4 conclude the proof of Theorem 3.1. ■

Let us illustrate Theorem 3.1.

Example 3.2. Suppose that $X = \mathbb{R}$, $\alpha = 5$, and $\beta = 1$. Writing z instead of γ, we note that $\widetilde{C}_\alpha$ turns into the set

$$S := \{(x,y,z) \in \mathbb{R}^3 \mid x^2 - y^2 = 10z\} = \operatorname{gra}\big((x,y) \mapsto \tfrac{1}{10}(x^2-y^2)\big).$$

Let us now compute $P_S(x_0,y_0,z_0)$ for various points.

(i) Suppose that $(x_0,y_0,z_0) = (2,-3,4)$.
In view of Theorem 3.1(i), we set $p := |x_0|^2 - |y_0|^2 = 2^2 - (-3)^2 = -5$ and $q := |x_0|^2 + |y_0|^2 = 2^2 + (-3)^2 = 13$. Following (10), we consider the equation

$$\frac{(\lambda^2+1)p - 2\lambda q}{(1-\lambda^2)^2} - \frac{2\lambda\alpha^2}{\beta^2} - 2\alpha\gamma_0 = -\frac{5\lambda^2+26\lambda+5}{(1-\lambda^2)^2} - 50\lambda - 40 = 0$$

which has $\lambda = -0.52416$ as its unique (approximate) root in $]-1,1[$. Using (9) now yields

$$\begin{aligned} P_S(x_0,y_0,z_0) &= \left\{\Big(\frac{x_0}{1+\lambda}, \frac{y_0}{1-\lambda}, z_0 + \frac{\lambda\alpha}{\beta^2}\Big)\right\} \\ &= \big\{(4.20311, -1.96830, 1.37919)\big\}. \end{aligned}$$

This is depicted in Fig. 1 with the green arrow.

(ii) Suppose that $(x_0,y_0,z_0) = (0,-3,3)$.
In view of Theorem 3.1(ii), we evaluate $\alpha(z_0 - \frac{\alpha}{\beta^2}) = 5(3-5) = -10 < -\frac{9}{8} = -\frac{|y_0|^2}{8}$ and we are thus in case (ii)(a). In view of (12), we consider the equation

$$\frac{|y_0|^2}{(1-\lambda)^2} + \frac{2\lambda\alpha^2}{\beta^2} + 2\alpha z_0 = \frac{9}{(1-\lambda)^2} + 50\lambda + 30 = 0$$

which has $\lambda = -0.66493$ as its unique (approximate) root in $]-1,1[$. Using (11) now yields

$$P_S(x_0,y_0,z_0) = \left\{\left(0, \frac{v_0}{1-\lambda}, \gamma_0 + \frac{\lambda\alpha}{\beta^2}\right)\right\} = \left\{\left(0, -1.80187, -0.32467\right)\right\}.$$

This is depicted in Fig. 1 with a single blue arrow.

(iii) Suppose that $(x_0,y_0,z_0) = (0,\sqrt{32},6) = (0,5.65685,6)$.
In view of Theorem 3.1(ii), we evaluate $\alpha(z_0 - \frac{\alpha}{\beta^2}) = 5(6-5) = 5 > -4 = -\frac{32}{8} = -\frac{|y_0|^2}{8}$ and we are thus in case (ii)(b). We compute

$$\sqrt{2\alpha\left(z_0 - \frac{\alpha}{\beta^2}\right) + \frac{|y_0|^2}{4}} = \sqrt{10(6-5) + \frac{32}{4}} = \sqrt{18}$$

and now (13) yields

$$P_S(x_0,y_0,z_0) = \left\{\left(x, \frac{y_0}{2}, z_0 - \frac{\alpha}{\beta^2}\right) \;\middle|\; |x| = \sqrt{2\alpha\left(z_0 - \frac{\alpha}{\beta^2}\right) + \frac{|y_0|^2}{4}}, u \in \mathbb{R}\right\} = \left\{\left(\pm\sqrt{18}, \sqrt{8}, 1\right)\right\} = \left\{\left(\pm 4.24264, 2.82843, 1\right)\right\}.$$

This is depicted in Fig. 1 with double blue arrows.

(iv) Suppose that $(x_0,y_0,z_0) = (0,0,6)$.
In view of Theorem 3.1(iv), we have $\alpha z_0 = 5(6) = 30 > 25 = \frac{\alpha^2}{\beta^2}$ and we are thus in case (iv)(a). We compute

$$\sqrt{2\alpha\left(\gamma_0 - \frac{\alpha}{\beta^2}\right)} = \sqrt{10(6-5)} = \sqrt{10}$$

and now (17) yields

$$P_S(x_0,y_0,z_0) = \left\{\left(x, 0, z_0 - \frac{\alpha}{\beta^2}\right) \;\middle|\; |x| = \sqrt{2\alpha\left(z_0 - \frac{\alpha}{\beta^2}\right)}, u \in \mathbb{R}\right\} = \left\{\left(\pm\sqrt{10}, 0, 1\right)\right\} = \left\{\left(\pm 3.16228, 0, 1\right)\right\}.$$

This is depicted in Fig. 1 with double black arrows.

(v) Suppose that $(x_0,y_0,z_0) = (0,0,4)$.
In view of Theorem 3.1(iv), we have $|\alpha z_0| = |5(4)| = 20 < 25 = \frac{\alpha^2}{\beta^2}$ and we are thus in case (iv)(b). Therefore,

$$P_S(x_0,y_0,z_0) = \{(0,0,0)\}.$$

This is depicted in Fig. 1 with a single black arrow.

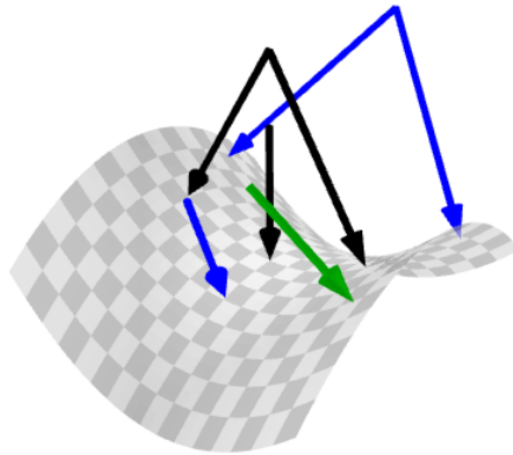

Figure 1: Visualization of the 5 projections from Example 3.2.

4 Further results

Recall that

$$C_\alpha = \{(x,y,\gamma) \in X \times X \times \mathbb{R} \mid \langle x,y\rangle = \alpha\gamma\},$$

and this is the representation more natural to use in Deep Learning (see [6]). Armed with Theorem 3.1, the projection onto C_α now readily obtained:

Theorem 4.1. Let $(x_0,y_0,\gamma_0) \in X \times X \times \mathbb{R}$. Then the following hold:

(i) If $x_0 \neq \pm y_0$, then

$$P_{C_\alpha}(x_0,y_0,\gamma_0) = \left\{\left(\frac{x_0 - \lambda y_0}{1-\lambda^2}, \frac{y_0 - \lambda x_0}{1-\lambda^2}, \gamma_0 + \frac{\lambda\alpha}{\beta^2}\right)\right\}$$

for a unique $\lambda \in \left]-1,1\right[$ that solves the (essentially) quintic equation

$$g(\lambda) := \frac{(\lambda^2+1)p - 2\lambda q}{(1-\lambda^2)^2} - \frac{2\lambda\alpha^2}{\beta^2} - 2\alpha\gamma_0 = 0,$$

where $p := 2\langle x_0,y_0\rangle$ and $q := \|x_0\|^2 + \|y_0\|^2$.

(ii) If $y_0 = -x_0 \neq 0$, then we have the following:

a) When $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) < -\frac{\|x_0\|^2}{4}$, then

$$P_{C_\alpha}(x_0,-x_0,\gamma_0) = \left\{\left(\frac{x_0}{1-\lambda}, \frac{-x_0}{1-\lambda}, \gamma_0 + \frac{\lambda\alpha}{\beta^2}\right)\right\}$$

for a unique $\lambda \in]-1,1[$ that solves

$$g_1(\lambda) := \frac{2\|x_0\|^2}{(1-\lambda)^2} + \frac{2\lambda\alpha^2}{\beta^2} + 2\alpha\gamma_0 = 0.$$

b) When $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) \geq -\frac{\|x_0\|^2}{4}$, then

$$P_{C_\alpha}(x_0, -x_0, \gamma_0) = \\ \left\{ \left(\frac{x_0}{2} + \frac{u}{\sqrt{2}}, -\frac{x_0}{2} + \frac{u}{\sqrt{2}}, \gamma_0 - \frac{\alpha}{\beta^2} \right) \,\middle|\, \|u\| = \sqrt{2\alpha\Big(\gamma_0 - \frac{\alpha}{\beta^2}\Big) + \frac{\|x_0\|^2}{2}},\ u \in X \right\}$$

which is a singleton if and only if $\alpha(\gamma_0 - \frac{\alpha}{\beta^2}) = -\frac{\|x_0\|^2}{4}$.

(iii) If $y_0 = x_0 \neq 0$, then we have the following:

a) When $\alpha(\gamma_0 + \frac{\alpha}{\beta^2}) > \frac{\|x_0\|^2}{4}$, then

$$P_{C_\alpha}(x_0, x_0, \gamma_0) = \left\{ \left(\frac{x_0}{1+\lambda}, \frac{x_0}{1+\lambda}, \gamma_0 + \frac{\lambda\alpha}{\beta^2} \right) \right\}$$

for a unique $\lambda \in]-1,1[$ that solves the (essentially) cubic equation

$$g_2(\lambda) := \frac{2\|x_0\|^2}{(1+\lambda)^2} - \frac{2\lambda\alpha^2}{\beta^2} - 2\alpha\gamma_0 = 0.$$

b) When $\alpha(\gamma_0 + \frac{\alpha}{\beta^2}) \leq \frac{\|x_0\|^2}{4}$, then

$$P_{C_\alpha}(x_0, x_0, \gamma_0) = \\ \left\{ \left(\frac{x_0}{2} - \frac{v}{\sqrt{2}}, \frac{x_0}{2} + \frac{v}{\sqrt{2}}, \gamma_0 + \frac{\alpha}{\beta^2} \right) \,\middle|\, \|v\| = \sqrt{-2\alpha\Big(\gamma_0 + \frac{\alpha}{\beta^2}\Big) + \frac{\|x_0\|^2}{2}},\ v \in X \right\}$$

which is a singleton if and only if $\alpha(\gamma_0 + \frac{\alpha}{\beta^2}) = \frac{\|x_0\|^2}{4}$.

(iv) If $x_0 = y_0 = 0$, then we have the following:

a) When $\alpha\gamma_0 > \frac{\alpha^2}{\beta^2}$, then the projection is the non-singleton set

$$P_{C_\alpha}(0,0,\gamma_0) = \left\{ \left(\frac{u}{\sqrt{2}}, \frac{u}{\sqrt{2}}, \gamma_0 - \frac{\alpha}{\beta^2} \right) \,\middle|\, \|u\| = \sqrt{2\alpha\Big(\gamma_0 - \frac{\alpha}{\beta^2}\Big)},\ u \in X \right\}.$$

b) When $|\alpha\gamma_0| \leq \frac{\alpha^2}{\beta^2}$, then

$$P_{C_\alpha}(0,0,\gamma_0) = \big\{ (0,0,0) \big\}.$$

c) When $\alpha\gamma_0 < -\frac{\alpha^2}{\beta^2}$, then the projection is the non-singleton set

$$P_{C_\alpha}(0,0,\gamma_0) = \Big\{\Big(-\frac{v}{\sqrt{2}}, \frac{v}{\sqrt{2}}, \gamma_0 + \frac{\alpha}{\beta^2}\Big) \,\Big|\, \|v\| = \sqrt{-2\alpha\Big(\gamma_0 + \frac{\alpha}{\beta^2}\Big)}, v \in X\Big\}.$$

Proof. With

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}}\mathrm{Id} & -\frac{1}{\sqrt{2}}\mathrm{Id} & 0 \\ \frac{1}{\sqrt{2}}\mathrm{Id} & \frac{1}{\sqrt{2}}\mathrm{Id} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

in mind, by Proposition 2.2(iii) we have

$$\begin{aligned} P_{C_\alpha}[x_0, y_0, \gamma_0]^\intercal &= A\, P_{\widetilde{C}_\alpha} A^\intercal [x_0, y_0, \gamma_0]^\intercal \\ &= A\, P_{\widetilde{C}_\alpha}\Big[\frac{x_0 + y_0}{\sqrt{2}}, \frac{-x_0 + y_0}{\sqrt{2}}, \gamma_0\Big]^\intercal. \end{aligned}$$

Hence (i)–(iv) follow by applying Theorem 3.1. ■

Remark 4.2. Theorem 4.1(i) was given in [6, Appendix B] without a rigorous mathematical justification.

It is interesting to ask what happens when $\alpha \to 0$.

Theorem 4.3. Suppose that $X = \mathbb{R}^n$. Then $P_{C_\alpha} \xrightarrow{g} P_{C\times\mathbb{R}} = P_C \times \mathrm{Id}$ and $P_{\widetilde{C}_\alpha} \xrightarrow{g} P_{\widetilde{C}\times\mathbb{R}} = P_{\widetilde{C}} \times \mathrm{Id}$ when $\alpha \to 0$.

Proof. Apply Proposition 2.4 and Fact 2.6. ■

Remark 4.4. The projection onto the cross C, P_C, has been given in [2].

Acknowledgments

HHB and XW were supported by NSERC Discovery Grants. MKL was partially supported by a SERB-UBC Fellowship and NSERC Discovery Grants of HHB and XW.

References

[1] Jean-Pierre Aubin and Hélène Frankowska, *Set-valued analysis*, Springer Science & Business Media, 2009.

[2] Heinz H Bauschke, Manish Krishan Lal, and Xianfu Wang, *The projection onto the cross*, Set-Valued and Variational Analysis **30** (2022), no. 3, 997–1009.

[3] Heinz H Bauschke, Manish Krishan Lal, and Xianfu Wang, *Projections onto hyperbolas or bilinear constraint sets in Hilbert spaces*, Journal of Global Optimization **86** (2023), no. 1, 25–36.

[4] F. Bernard and L. Thibault, *Prox-regular functions in Hilbert spaces*, Journal of Mathematical Analysis and Applications **303** (2005), no. 1, 1–14.

[5] Dimitri P Bertsekas, *Nonlinear programming*, Journal of the Operational Research Society **48** (1997), no. 3, 334–334.

[6] Veit Elser, *Learning without loss*, Fixed Point Theory and Algorithms for Sciences and Engineering **2021** (2021), no. 1, 12.

[7] Erwin Kreyszig, *Introductory functional analysis with applications*, Vol. 17, John Wiley & Sons, 1991.

[8] Boris Odehnal, Hellmuth Stachel, and Georg Glaeser, *The universe of quadrics*, Springer Nature, 2020.

[9] R. Tyrrell Rockafellar and Roger J. B. Wets, *Variational Analysis*, Vol. 317, Springer, Berlin.