VARIATIONAL PROPERTIES OF MATRIX FUNCTIONS VIA THE GENERALIZED MATRIX-FRACTIONAL FUNCTION

JAMES V. BURKE[*], YUAN GAO[†], AND TIM HOHEISEL[‡]

Abstract. We show that many important convex matrix functions can be represented as the partial infimal projection of the generalized matrix fractional (GMF) and a relatively simple convex function. This representation provides conditions under which such functions are closed and proper as well as formulas for the ready computation of both their conjugates and subdifferentials. Special attention is given to support and indicator functions. Particular instances yield all weighted Ky Fan norms and squared gauges on $\mathbb{R}^{n\times m}$, and as an example we show that all variational Gram functions are representable as squares of gauges. Other instances yield weighted sums of the Frobenius and nuclear norms. The scope of applications is large and the range of variational properties and insight is fascinating and fundamental. An important byproduct of these representations is that they lay the foundation for a smoothing approach to many matrix functions on the interior of the domain of the GMF function, which opens the door to a range of unexplored optimization methods.

Key words. convex analysis, infimal projection, matrix-fractional function, support function, gauge function, subdifferential, Ky Fan norm, variational Gram function

AMS subject classifications. 68Q25, 68R10, 68U05

1. Introduction.

The *generalized matrix-fractional (GMF)* function was introduced in [5] where it is shown to unify a number of tools and concepts for matrix optimization including optimal value functions in quadratic programming, nuclear norm optimization, multi-task learning, and, of course, the matrix fractional function. In the present paper we greatly expand the number of applications to include all *Ky Fan norms*, matrix *gauge functionals*, and *variational Gram functions* [14]. Our analysis includes descriptions of the variational properties of these functions such as formulas for their convex conjugates and their subdifferentials.

In what follows, we set $\mathbb{E} := \mathbb{R}^{n\times m} \times \mathbb{S}^n$ where $\mathbb{R}^{n\times m}$ and $\mathbb{S}^n$ are the linear spaces of real $n \times m$ matrices and (real) symmetric $n \times n$ matrices, respectively. Given $(A, B) \in \mathbb{R}^{p\times n} \times \mathbb{R}^{p\times m}$ with $\operatorname{rge} B \subset \operatorname{rge} A$, recall that the GMF function φ is defined as the support function of the graph of the matrix valued mapping $Y \mapsto -\frac{1}{2}YY^T$ over the manifold $\{Y \in \mathbb{R}^{n\times m} \mid AY = B\}$, i.e., $\varphi : \mathbb{E} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ is

$$\varphi(X, V) := \sup \{\langle (Y, W),\, (X, V)\rangle \mid (Y, W) \in \mathcal{D}(A, B)\}, \tag{1.1}$$

where

$$\mathcal{D}(A, B) := \left\{ \left(Y, -\frac{1}{2}YY^T\right) \in \mathbb{E} \;\Big|\; Y \in \mathbb{R}^{n\times m} :\; AY = B \right\}. \tag{1.2}$$

A closed form expression for φ is derived in [5] where it is also shown that φ is smooth on the (nonempty) interior of its domain.

Our study focuses on functions $p : \mathbb{R}^{n\times m} \to \overline{\mathbb{R}}$ representable as the partial infimal projection of the form

$$p(X) := \inf_{V\in\mathbb{S}^n} \varphi(X, V) + h(V), \tag{1.3}$$

[*]Department of Mathematics, University of Washington, Seattle, WA 98195 (jvburke@uw.edu).
Research is supported in part by the National Science Foundation under grant number DMS-1514559.
[†]Department of Applied Mathematics, University of Washington, Seattle, WA 98195 (yuangao@uw.edu).

[‡]McGill University, 805 Sherbrooke St West, Room1114, Montréal, Québec, Canada H3A 0B9 tim.hoheisel@mcgill.ca)

where $h : \mathbb{S}^n \to \overline{\mathbb{R}}$ is convex. Different functions h illuminate different variational properties of the matrix X. For example, when $h := \langle U, \cdot \rangle$ for $U \in \mathbb{S}^n_{++}$, and when both A and B are zero, then p is a weighted nuclear norm where the weights depend on any "square-root" of U (see Corollary 4.7). Among the consequences of the representation (1.3) are conditions under which p is closed and proper as well as formulas for the ready computation of both p^* and ∂p (Section 3). As an application of our general results, we give more detailed explorations in the cases where h is a support function (Section 4) or an indicator function (Section 5). We illustrate these results with specific instances. For example, we obtain all weighted squared gauges on $\mathbb{R}^{n\times m}$, cf. Corollary 5.9, as well as a complete characterization of variational Gram functions [14] and their conjugates. In addition, we show that all variational Gram funtions are representable as squares of gauges, cf. Proposition 5.10. Other choices yield weighted sums of Frobenius and nuclear norms, see [5, Corollary 5.9]. The scope of applications is large and the range of variational properties is fascinating and fundamental.

Beyond the variational results of this paper, there is a compelling but unexplored computational aspect of this representation. Hsieh and Olsen [13] show that (1.3) with $h = \frac{1}{2}\mathrm{tr}\,(\cdot)$ yields a smoothing approach to optimization problems involving the nuclear norm. More generally, observe that many matrix optimization problems often take the form

$$(P) \qquad \min_{X\in\mathbb{R}^{n\times m}} f(X) + p(X),$$

where $f, p : \mathbb{R}^{n\times m} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$. The function f is thought of as the primary objective and is often smooth or convex while p is typically a structure inducing convex function. Using the representation (1.3), the problem (P) can be written as

$$\min_{(X,V)\in\mathbb{E}} f(X) + \varphi(X, V) + h(V).$$

This reformulation allows one to exploit the smoothness of φ on the interior of its domain. For example, if both f and h are smooth, one can employ a damped Newton, or path following approach to solving (P). We emphasize, that this is not the goal or intent of this paper, however, our results provide the basis for future investigations along a variety of such numerical and theoretical avenues.

The paper is organized as follows: In Section 2 we provide the tools from convex analysis and some basic properties of the GMF function. Section 3 contains the general theory for partial infimal projections of the form (1.3). In Section 4 we specify h in (1.3) to be a support function of some closed, convex set $\mathcal{V} \subset \mathbb{S}^n$. In Section 5 we choose h to be the indicator of such set. In particular, this yields powerful results on variational Gram functions and Ky Fan norms, see Section 5.2-5.3. We close out with some final remarks in Section 6 and supplementary material in Section 7.

Notation: For a linear transformation L, we write $\mathrm{rge}\, L$ and $\ker L$ for its *range* and *kernel*, respectively. For $A \in \mathbb{R}^{p\times n}$, we abuse notation somewhat and write $\mathrm{rge}\, A$ and $\ker A$ for its *range* and *kernel*, respectively, when A is considered as a linear transformation between $\mathbb{R}^n$ and $\mathbb{R}^p$. Again, for $A \in \mathbb{R}^{p\times n}$, we set

$$\mathrm{Ker}_r A := \left\{ X \in \mathbb{R}^{n\times r} \mid AX = 0 \right\} = \left\{ X \in \mathbb{R}^{n\times r} \mid \mathrm{rge}\, X \subset \ker A \right\},$$
$$\mathrm{Rge}_r A := \left\{ Y \in \mathbb{R}^{p\times r} \;\middle|\; \exists\, X \in \mathbb{R}^{n\times r} \,:\, Y = AX \right\} = \left\{ Y \in \mathbb{R}^{p\times r} \mid \mathrm{rge}\, Y \subset \mathrm{rge}\, A \right\}$$

and write $\mathrm{Ker} A$ or $\mathrm{Rge} A$ when the choice of r is clear from the context. Observe that $\mathrm{Ker}_1 A = \ker A$, $\mathrm{Rge}_1 A = \mathrm{rge}\, A$, and $(\mathrm{Ker}_r A)^\perp = \mathrm{Rge}_r A^T$, where we equip any matrix space with the (Frobenius) inner product $\langle X, Y\rangle := \mathrm{tr}\,(X^TY)$. The *Moore-Penrose pseudoinverse* of A, see e.g. [11], is denoted by $A^\dagger$. The set of all symmetric matrices of dimension n is given by $\mathbb{S}^n$. The positive and negative semidefinite cone are denoted by $\mathbb{S}^n_+$ and $\mathbb{S}^n_-$, respectively.

For two sets S, T in the same real linear space their *Minkowski sum* is $S + T := \{s + t \mid s \in S,\ t \in T\}$. For $I \subset \mathbb{R}$ we also put $I \cdot S := \{\lambda s \mid \lambda \in I,\ s \in S\}$.

2. Preliminaries.

Tools from convex analysis. Let $(\mathcal{E}, \langle\cdot, \cdot\rangle)$ be a finite-dimensional Euclidean space with induced norm $\|\cdot\| := \sqrt{\langle\cdot, \cdot\rangle}$. E.g. on matrix spaces we use the Frobenius norm induced by the trace operator. The closed ϵ-ball about a point $x \in \mathcal{E}$ is denoted by $B_\epsilon(x)$.

Let $S \subset \mathcal{E}$ be nonempty. The (topological) *closure* and *interior* of S are denoted by $\mathrm{cl}\, S$ and $\mathrm{int}\, S$, respectively. The *(linear) span* of S will be denoted by $\mathrm{span}\, S$. The *convex hull* of S is the set of all convex combinations of elements of S and is denoted by $\mathrm{conv}\, S$. Its closure (the *closed convex hull*) is $\overline{\mathrm{conv}}\, S := \mathrm{cl}\,(\mathrm{conv}\, S)$. The *conical hull* (also *positive hull*) of S is the set

$$\mathrm{pos}\, S := \mathbb{R}_+ \cdot S = \{\lambda x \mid x \in S,\ \lambda \geq 0\}.$$

The *convex conical hull* of S is

$$\mathrm{cone}\, S := \left\{\sum_{i=1}^r \lambda_i x_i \mid r \in \mathbb{N},\ x_i \in S,\ \lambda_i \geq 0\right\}.$$

It is easily seen that $\mathrm{cone}\, S = \mathrm{pos}\,(\mathrm{conv}\, S) = \mathrm{conv}\,(\mathrm{pos}\, S)$. The closure of the latter is $\overline{\mathrm{cone}}\, S := \mathrm{cl}\,(\mathrm{cone}\, S)$. The *affine hull* of S is denoted by $\mathrm{aff}\, S$.

The *relative interior* of a convex set $C \subset \mathcal{E}$ is given by

$$\mathrm{ri}\, C = \{x \in C \mid \exists \varepsilon > 0 :\ B_\varepsilon(x) \cap \mathrm{aff}\, C \subset C\}.$$

The points $x \in \mathrm{ri}\, C$ are characterized through (see, e.g., [2, Section 6.2])

$$\mathrm{pos}\,(C - x) = \mathrm{span}\,(C - x). \tag{2.1}$$

The *polar set* of S is defined by

$$S^\circ := \{v \in \mathcal{E} \mid \langle v, x\rangle \leq 1\ (x \in S)\}.$$

Moreover, we define the *bipolar set* of S by $S^{\circ\circ} := (S^\circ)^\circ$, so that $S^{\circ\circ} = \overline{\mathrm{conv}}\,(S \cup \{0\})$. If $K \subset \mathcal{E}$ is a cone (i.e. $\mathrm{pos}\, K \subset K$) we have

$$K^\circ = \{v \in \mathcal{E} \mid \langle v, x\rangle \leq 0\ (x \in K)\} =: K^-.$$

If $U \subset \mathcal{E}$ is a subspace, U° is the orthogonal subspace $U^\perp$.

The *horizon cone* of $S \subset \mathcal{E}$ is the the closed cone given by

$$S^\infty := \{v \in \mathcal{E} \mid \exists \{\lambda_k\} \downarrow 0,\ \{x_k \in S\} :\ \lambda_k x_k \to v\}.$$

For a cone $K \subset \mathcal{E}$ we have $K^\infty = \mathrm{cl}\, K$. Moreover, for a convex set $C \subset \mathcal{E}$, C^∞ coincides with the *recession cone* of the closure of C, i.e.

$$C^\infty = \{v \mid x + tv \in \mathrm{cl}\, C\ (t \geq 0,\ x \in C)\} = \{y \mid C + y \subset C\}. \tag{2.2}$$

For $f : \mathcal{E} \to \overline{\mathbb{R}}$ its *domain* and *epigraph* are given by

$$\operatorname{dom} f := \{x \in \mathcal{E} \mid f(x) < +\infty\} \quad \text{and} \quad \operatorname{epi} f := \{(x, \alpha) \in \mathcal{E} \times \mathbb{R} \mid f(x) \le \alpha\},$$

respectively. We call f *convex* if its epigraph $\operatorname{epi} f$ is a convex set, and we call it *closed* (or *lower semicontinuous*) if $\operatorname{epi} f$ is closed. If f is proper, we call it *positively homogeneous* if $\operatorname{epi} f$ is a cone, and *sublinear* if $\operatorname{epi} f$ is a convex cone. In what follows we use the following abbreviations:

$$\Gamma(\mathcal{E}) := \{f : \mathcal{E} \to \mathbb{R} \cup \{+\infty\} \mid f \text{ proper, convex}\} \quad \text{and} \quad \Gamma_0(\mathcal{E}) := \{f \in \Gamma(\mathcal{E}) \mid f \text{ closed}\}.$$

The *lower semicontinuous hull* $\operatorname{cl} f$ and the *horizon function* f^∞ of f are defined through the relations

$$\operatorname{cl}(\operatorname{epi} f) = \operatorname{epi} \operatorname{cl} f \quad \text{and} \quad \operatorname{epi} f^\infty = (\operatorname{epi} f)^\infty,$$

respectively. For $f \in \Gamma_0(\mathcal{E})$ the horizon function f^∞ coincides with the *recession function*, see e.g. [15, p. 66], and all of the respective results apply. Note that also the moniker *asymptotic function* is used for the horizon function, see e.g. [1, 10].

The *horizon cone of a function* f is defined as

$$\operatorname{hzn} f := \{x \mid f^\infty(x) \le 0\}.$$

By [15, Theorem 8.7], for $f \in \Gamma_0$, we have

$$\operatorname{hzn} f = \{W \mid f(x) \le \mu\}^\infty \quad (\mu \in \mathbb{R} : \{W \mid h^*(W) \le \mu\} \neq \emptyset).$$

For a convex function $f : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ its *subdifferential* at a point $\bar{x} \in \operatorname{dom} f$ is given by

$$\partial f(\bar{x}) := \{v \in \mathcal{E} \mid f(x) \ge f(\bar{x}) + \langle v, x - \bar{x}\rangle\}.$$

Recall that, for a convex function f, we always have

$$\operatorname{ri}(\operatorname{dom} f) \subset \operatorname{dom} \partial f \subset \operatorname{dom} f,$$

see e.g. [15, p. 227], where $\operatorname{dom} \partial f := \{x \in \mathcal{E} \mid \partial f(x) \neq \emptyset\}$ is the *domain of the subdifferential*.

For some function $f : \mathcal{E} \to \overline{\mathbb{R}}$ its *(Fenchel) conjugate* $f^* : \mathcal{E} \to \overline{\mathbb{R}}$ is given by

$$f^*(y) := \sup_{x \in \mathcal{E}} \{\langle x, y\rangle - f(x)\}.$$

Note that $f \in \Gamma_0(\mathcal{E})$ if and only if $f = f^{**} := (f^*)$. The definition of the conjugate function yields the *Fenchel-Young inequality*

$$f(x) + f^*(y) \ge \langle x, y\rangle \quad (x, y \in \mathcal{E}). \tag{2.3}$$

Given a nonempty set $S \subset \mathcal{E}$, its *indicator function* $\delta_S : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ is given by

$$\delta_S(x) := \begin{cases} 0 & \text{if} \quad x \in S, \\ +\infty & \text{if} \quad x \notin S. \end{cases}$$

The indicator of S is convex if and only if S is a convex set, in which case the *normal cone* of S at $\bar{x} \in S$ is given by

$$N_S(\bar{x}) := \partial \delta_S(\bar{x}) = \{v \in \mathcal{E} \mid \langle v, x - \bar{x}\rangle \le 0 \ (x \in S)\}.$$

The *support function* $\sigma_S : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ and the *gauge function* $\gamma_S : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ of a nonempty set $S \subset \mathcal{E}$ are given respectively by

$$\sigma_S(x) := \sup_{v \in S} \langle v, x \rangle \quad \text{and} \quad \gamma_S(x) := \inf \{t \geq 0 \mid x \in tS\}.$$

Here we use the standard convention that $\inf \emptyset = +\infty$. It is easy to see that

$$\sigma_S = \sigma_{\overline{\mathrm{conv}}\, S} \quad \text{and} \quad \gamma_{\mathrm{conv}\, S} = \gamma_{\overline{\mathrm{conv}}\, S}. \tag{2.4}$$

Moreover, given two (nonempty) sets $S, T \subset \mathcal{E}$ and $x \in \mathcal{E}$, it is easily seen that

$$\sigma_S + \sigma_T = \sigma_{S+T}. \tag{2.5}$$

Suppose $C \subset \mathcal{E}$ is closed and convex. Then its *barrier cone* is defined by $\mathrm{bar}\, C := \mathrm{dom}\, \sigma_C$. The closure of the barrier cone of C and the horizon cone are paired in polarity, i.e.

$$(\mathrm{bar}\, C)^\circ = C^\infty \quad \text{and} \quad \mathrm{cl}\,(\mathrm{bar}\, C) = (C^\infty)^\circ. \tag{2.6}$$

For two functions $f_1, f_2 : \mathcal{E} \to \overline{\mathbb{R}}$, their *infimal convolution* $f_1 \,\Box\, f_2$ is defined by

$$(f_1 \,\Box\, f_2)(x) := \inf_{y \in \mathcal{E}} \{f_1(x-y) + f_2(y)\} \quad (x \in \mathcal{E}).$$

The generalized matrix-fractional function. As noted in the introduction, the GMF function is the support function of $\mathcal{D}(A,B)$ given in (1.2). Hence, we write

$$\sigma_{\mathcal{D}(A,B)}(X,V) = \varphi(X,V) \tag{2.7}$$

and also refer to $\sigma_{\mathcal{D}(A,B)}$ as the GMF function. From [5, 6], we obtain the formula

$$\varphi(X,V) = \begin{cases} \frac{1}{2}\mathrm{tr}\left(\binom{X}{B}^T M(V)^\dagger \binom{X}{B}\right) & \text{if} \quad \mathrm{rge}\binom{X}{B} \subset \mathrm{rge}\, M(V),\ V \in \mathcal{K}_A, \\ +\infty & \text{else,} \end{cases} \tag{2.8}$$

where $(A,B) \in \mathbb{R}^{p \times n} \times \mathbb{R}^{p \times m}$ with $\mathrm{rge}\, B \subset \mathrm{rge}\, A$ and $\mathcal{K}_A$ is the cone of all symmetric matrices that are positive semidefinite with respect to the subspace $\ker A$, i.e.

$$\mathcal{K}_A := \left\{V \in \mathbb{S}^n \;\middle|\; u^T V u \geq 0 \ (u \in \ker A)\right\}, \tag{2.9}$$

and $M(V)^\dagger$ is the Moore-Penrose pseudoinverse of the *bordered matrix*

$$M(V) = \begin{pmatrix} V & A^T \\ A & 0 \end{pmatrix}. \tag{2.10}$$

The *matrix-fractional function* [4, 9] is obtained by setting the matrices A and B to zero.

A detailed analysis of the GMF function appears in the papers [5, 6]. In particular, it is shown that

$$\begin{aligned} \mathrm{dom}\, \sigma_{\mathcal{D}(A,B)} &= \mathrm{dom}\, \partial\sigma_{\mathcal{D}(A,B)} \\ &= \left\{(X,V) \in \mathbb{R}^{n \times m} \times \mathbb{S}^n \;\middle|\; \mathrm{rge}\begin{pmatrix} X \\ B \end{pmatrix} \subset \mathrm{rge}\, M(V),\ V \in \mathcal{K}_A\right\}. \end{aligned} \tag{2.11}$$

For the study of the convex-analytical properties of the support function $\sigma_{\mathcal{D}(A,B)}$ the computation of the closed convex hull of the (nonconvex) set $\mathcal{D}(A,B)$ has been critical. A representation of $\overline{\text{conv}}\,\mathcal{D}(A,B)$ relying mainly on Carathéodory's theorem was obtained in [5, Proposition 4.3]. A refined and more versatile expression was proven in [6], see below. The key object for this expression is the (closed, convex) cone $\mathcal{K}_A$ defined in (2.9), which reduces to $\mathbb{S}^n_+$ for $A = 0$.

We briefly summarize the geometric and topological properties of $\mathcal{K}_A$ useful to our study, and which follow from [6, Proposition 1] (by setting $\mathcal{S} = \ker A$).

Proposition 2.1. *For $A \in \mathbb{R}^{p\times n}$ let $P \in \mathbb{R}^{n\times n}$ be the orthogonal projection onto $\ker A$ and let $\mathcal{K}_A$ be given by (2.9). Then the following hold:*

a) $\mathcal{K}_A = \{V \in \mathbb{S}^n \mid PVP \succeq 0\}$.

b) $\mathcal{K}_A^\circ = \text{cone}\,\left\{-vv^T \mid v \in \ker A\right\} = \{W \in \mathbb{S}^n \mid W = PWP \preceq 0\}$.

c) $\text{int}\,\mathcal{K}_A = \left\{V \in \mathbb{S}^n \;\middle|\; u^TVu > 0\ (u \in A \setminus \{0\})\right\}$.

The central result from [6] is the following characterization of $\overline{\text{conv}}\,\mathcal{D}(A,B)$.

Theorem 2.2 ([6, Theorem 2]). *Let $\mathcal{D}(A,B)$ be given by (1.2). Then*

$$\overline{\text{conv}}\,\mathcal{D}(A,B) = \Omega(A,B) := \left\{(Y,W) \in \mathbb{E} \;\middle|\; AY = B \text{ and } \frac{1}{2}YY^T + W \in \mathcal{K}_A^\circ\right\}.$$

In particular, Theorem 2.2 in combination with (2.4) implies that $\sigma_{\mathcal{D}(A,B)} = \sigma_{\Omega(A,B)}$, an identity which we will employ throughout.

3. Infimal projections of the generalized matrix-fractional function. We will now focus on infimal projections involving the GMF function. For these purposes consider the function $\psi : \mathbb{E} \to \overline{\mathbb{R}}$, given by

$$\psi(X,V) := \sigma_{\Omega(A,B)}(X,V) + h(V), \tag{3.1}$$

where $h \in \Gamma(\mathbb{S}^n)$ and $\Omega(A,B)$ is given by Theorem 2.2. Our primary objective is the infimal projection of the sum ψ from (3.1) in the variable V, i.e. we analyze the marginal function $p : \mathbb{R}^{n\times m} \to \overline{\mathbb{R}}$ defined by

$$p(X) := \inf_{V\in\mathbb{S}^n} \psi(X,V). \tag{3.2}$$

We lead with an elementary observation.

Lemma 3.1 (Domain of p). *Let p defined by (3.2). Then the following hold:*

a) p is convex.

b) $\text{dom}\,p = \{X \in \mathbb{R}^{n\times m} \mid \exists V \in \mathcal{K}_A \cap \text{dom}\,h : \ \text{rge}\,\binom{X}{B} \subset \text{rge}\,M(V)\}$. *In particular, p is proper if and only if $\text{dom}\,h \cap \mathcal{K}_A$ is nonempty.*

Moreover, if $\text{dom}\,p \neq \emptyset$ then the following hold:

c) If $B = 0$ (e.g. if $A = 0$) then $\text{dom}\,p$ is a subspace, hence relatively open.

d) If $\text{rank}\,A = p$ (full row rank) then $\text{dom}\,p = \mathbb{R}^{n\times m}$, hence open.

Proof. a) The convexity follows from, e.g., [16, Proposition 2.22].

b) The formula for $\text{dom}\,p$ follows from the definition of p and the representation of $\text{dom}\,\sigma_{\Omega(A,B)}$ in (2.11) which also gives the properness exactly when $\text{dom}\,h \cap \mathcal{K}_A \neq \emptyset$.

c) If $B = 0$, note that, $X \in \text{dom}\,p$ if and only if $\text{span}\,\{X\} \subset \text{dom}\,p$. Since $\text{dom}\,p$ is also convex, it is a subspace, see, e.g., [16, Proposition 3.8].

d) The bordered matrix $M(V)$ from (2.10) is invertible if (and only if) $\text{rank}\,A = p$. In this case the condition $\text{rge}\,\binom{X}{B} \subset \text{rge}\,M(V)$ is trivially satisfied for any $X \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{p\times m}$. Therefore the statement follows from b). ∎

The following example shows that the domain of p may not be relatively open (hence not a subspace) if $B \neq 0$, which proves that this assumption in Lemma 3.1 c) is not redundant.

EXAMPLE 3.2 ($\operatorname{dom} p$). *Let* $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ *and* $b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. *Then*

$$\ker A = \operatorname{span}\left\{\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\} \quad \textit{and} \quad \mathcal{K}_A = \left\{\begin{pmatrix} v & w \\ w & u \end{pmatrix} \mid v + u \geq 2w\right\}.$$

Moreover, put $\bar{V} := \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$ *and define*

$$\mathcal{V} := [0,1] \cdot \bar{V} = \left\{\begin{pmatrix} 2w & w \\ w & 0 \end{pmatrix} \mid w \in [0,1]\right\} \subset \mathbb{S}^2.$$

Then $\mathcal{V}$ *is clearly convex and compact. Now let* $h \in \Gamma_0(\mathbb{S}^2)$ *be any function with* $\operatorname{dom} h = \mathcal{V}$ *(e.g.* $h := \delta_{\mathcal{V}}$*). Note that*

$$\operatorname{dom} h \cap \mathcal{K}_A = \mathcal{V}.$$

We hence infer that

$$\begin{aligned}
x \in \operatorname{dom} p &\iff \exists w \in [0,1]: \begin{pmatrix} x \\ b \end{pmatrix} \in \operatorname{rge}\begin{pmatrix} w\bar{V} & A^T \\ A & 0 \end{pmatrix} \\
&\iff \exists w \in [0,1], r, s \in \mathbb{R}^2: \begin{array}{ccc} x & = & w\bar{V}r + A^T s, \\ b & = & Ar \end{array} \\
&\iff \exists w \in [0,1], \lambda, \mu \in \mathbb{R}: x = w\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} + \lambda\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right] + \mu\begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
&\iff \exists w \in [0,1], \gamma \in \mathbb{R}: x = w\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \gamma\begin{pmatrix} 1 \\ 1 \end{pmatrix}.
\end{aligned}$$

Therefore, we find that

$$\operatorname{dom} p = [0,1] \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \operatorname{span}\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\},$$

and hence

$$\operatorname{ri}(\operatorname{dom} p) = (0,1) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \operatorname{span}\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\},$$

so that $\operatorname{dom} p$ *is clearly not relatively open.*

As mentioned above, the former example shows that $\operatorname{dom} p$ may fail to be a subspace if $B \neq 0$. Lemma 3.1 d) and Example 3.17 a), on the other hand, illustrate that $\operatorname{dom} p$ might still be a subspace even if $B \neq 0$, hence the condition $B = 0$ is only sufficient for $\operatorname{dom} p$ to be a subspace (if nonempty).

3.1. ψ, ψ^*, and their subdifferentials. Our study of the infimal projection p given in (3.2) requires a thorough understanding of the properties of the functions ψ, ψ^*, and their subdifferentials. For this we make extensive use of the condition

$$\operatorname{ri}(\operatorname{dom} h) \cap \operatorname{int} \mathcal{K}_A \neq \emptyset,$$

which we refer to as the *conjugate constraint qualification* (CCQ).

LEMMA 3.3 (Conjugate of ψ). *Let* ψ *be given as in* (3.1) *and define*

$$\eta : (Y, W) \in \mathbb{E} \mapsto \inf_{(Y,T) \in \Omega(A,B)} h^*(W - T).$$

Then

$$\operatorname{dom} \eta = \left\{(Y, W) \;\middle|\; AY = B,\ \left(-\frac{1}{2}YY^T + \mathcal{K}_A^\circ\right) \cap (W - \operatorname{dom} h^*) \neq \emptyset\right\} \tag{3.3}$$

and the following hold:

a) ψ *is closed and convex.*
b) If $\operatorname{dom} h \cap \mathcal{K}_A \neq \emptyset$ *then* $\psi, \psi^* \in \Gamma_0(\mathbb{E})$ *with* $\psi^* = \operatorname{cl} \eta$.
c) Under CCQ , we have $\psi^* = \eta$. *Moreover, in this case, the infimum in the definition of* η *is attained on the whole domain, i.e.*

$$\begin{aligned}\mathcal{T}(\bar{Y}, \bar{W}) &:= \operatorname*{argmin}_{(Y,W)} \left\{ h^*(\bar{W}-W) \;\middle|\; (Y,W) \in \Omega(A,B),\, Y = \bar{Y} \right\} \\ &= \left\{ (\bar{Y}, W) \;\middle|\; (\bar{Y}, W) \in \Omega(A,B),\; \psi^*(\bar{Y}, \bar{W}) = h^*(\bar{W}-W) \right\}.\end{aligned} \tag{3.4}$$

is nonempty for all $(\bar{Y}, \bar{W}) \in \operatorname{dom} \psi^*$.
d) Under CCQ, $\operatorname{dom} \partial\psi^* = \{(Y,W) \mid \emptyset \neq \mathcal{T}(Y,W)\}$ *and, for every* $(Y,W) \in \operatorname{dom} \partial\psi^*$*, we have*

$$\partial\psi^*(Y,W) = \left\{ (X,V) \;\middle|\; \begin{array}{l} \exists\, T \in \mathbb{S}^n : \; V \in \partial h^*(W-T) \cap \mathcal{K}_A, \\ \left\langle V, \dfrac{1}{2} Y Y^T + T \right\rangle = 0,\; \operatorname{rge}(X - VY) \subset (\operatorname{Ker} A)^\perp \end{array} \right\}.$$

Proof. Note that $\eta(Y,W) < +\infty$ if and only if there is a $T \in \mathbb{S}^n$ such that $(Y,T) \in \Omega(A,B)$ and $W - T \in \operatorname{dom} h^*$, or equivalently, $AY = B,\; T \in -\frac{1}{2} Y Y^T + \mathcal{K}_A^\circ$ and $T \in W - \operatorname{dom} h^*$, which proves (3.3).

Define $\hat{h} : \mathbb{E} \to \overline{\mathbb{R}}$ by $\hat{h}(X,V) := h(V)$. Then $\operatorname{dom} \hat{h} = \mathbb{R}^{n\times m} \times \operatorname{dom} h$ and $\psi = \sigma_{\Omega(A,B)} + \hat{h}$.

a) The sum of two closed and convex functions is always closed and convex.

b) The sum of two proper functions is proper if and only if the domains of both functions intersect. Here, note that

$$\operatorname{dom} \hat{h} \cap \operatorname{dom} \sigma_{\mathcal{D}(A,B)} \neq \emptyset \quad \Longleftrightarrow \quad \operatorname{dom} h \cap \mathcal{K}_A \neq \emptyset.$$

Therefore, ψ is proper if (and only if) the latter condition holds. Combined with a) this shows ψ is closed, proper, and convex, and hence, so is its conjugate ψ^*.

Moreover, from Theorem 7.1 a) we infer

$$\psi^*(Y,W) = \operatorname{cl} \left(\delta_{\Omega(A,B)} \,\square\, \hat{h}^* \right)(Y,W).$$

Since $\hat{h}^*(Y,W) = \delta_{\{0\}}(Y) + h^*(W)$, we have

$$\delta_{\Omega(A,B)} \,\square\, \hat{h}^*(Y,W) = \inf_{(Y,T)\in\Omega(A,B)} h^*(W-T),$$

which proves $\psi^* = \operatorname{cl} \eta$.

c) We have $\operatorname{ri}(\operatorname{dom} \hat{h}) = \mathbb{R}^{n\times m} \times \operatorname{ri}(\operatorname{dom} h)$. Also, by [5, Theorem 4.1], we have $\operatorname{int}(\operatorname{dom} \sigma_{\Omega(A,B)}) = \{(X,V) \mid V \in \operatorname{int} \mathcal{K}_A\}$. Hence

$$\operatorname{ri}(\operatorname{dom} \hat{h}) \cap \operatorname{ri}(\operatorname{dom} \sigma_{\mathcal{D}(A,B)}) \neq \emptyset \quad \Longleftrightarrow \quad \operatorname{ri}(\operatorname{dom} h) \cap \operatorname{int} \mathcal{K}_A \neq \emptyset. \tag{3.5}$$

Theorem 7.1 a) (applied to $\sigma_{\Omega(A,B)}$ and $\hat{h}$), CCQ, and (3.5) then imply $\psi^* = \eta$ with

$$\emptyset \neq \mathcal{T}(\bar{Y}, \bar{W}) := \operatorname*{argmin}_{(Y,W)} \left\{ h^*(\bar{W} - W) \;\middle|\; (Y,W) \in \Omega(A,B),\; Y = \bar{Y} \right\}.$$

d) Observe that $\partial\sigma^*_{\mathcal{D}(A,B)} = N_{\Omega(A,B)}$ and $\partial\hat{h}^* = \mathbb{R}^{n\times m}\times\partial h^*$. Then part c) and Theorem 7.1 d) (applied to $\sigma_{\Omega(A,B)}$ and $\hat{h}$) yield

$$\begin{aligned}\partial\psi^*(Y,W) &= \left\{(X,V)\;\middle|\; \begin{array}{l}(X,V)\in\partial\sigma^*_{\mathcal{D}(A,B)}(Y_1,W_1)\cap\partial\hat{h}^*(Y_2,W_2),\\ (Y,W)=(Y_1,W_1)+(Y_2,W_2)\end{array}\right\}\\ &= \left\{(X,V)\;\middle|\;\exists T\in\mathbb{R}^{n\times m}:\ (X,V)\in N_{\Omega(A,B)}(Y,T),\ V\in\partial h^*(W-T)\right\}.\end{aligned}$$

The claim follows from the representation for $N_{\Omega(A,B)}(Y,T)$ in [6, Proposition 3]. ◻

We now turn our attention to the subdifferential of ψ which will be used for computing the subdifferential of its infimal projection p.

COROLLARY 3.4 (Subdifferential of ψ). *Let ψ be given by (3.1) and $\mathcal{T}(\cdot,\cdot)$ by (3.4). Then the following hold:*

a) If $(\bar{Y},\bar{W})\in\partial\sigma_{\Omega(A,B)}(\bar{X},\bar{V})+\{0\}\times\partial h(\bar{V})$, then $\mathcal{T}(\bar{Y},\bar{W})\neq\emptyset$ and

$$\mathcal{T}(\bar{Y},\bar{W}) = \left\{\bar{T}\in\mathbb{S}^n\;\middle|\;\bar{W}-\bar{T}\in\partial h(\bar{V}),\ (\bar{Y},\bar{T})\in\partial\sigma_{\Omega(A,B)}(\bar{X},\bar{V})\right\}. \tag{3.6}$$

b) Under CCQ we have

$$\operatorname{dom}\partial\psi = \left\{(X,V)\;\middle|\;V\in\operatorname{dom}\partial h\cap\mathcal{K}_A,\ \operatorname{rge}\begin{pmatrix}X\\B\end{pmatrix}\subset\operatorname{rge}M(V)\right\}.$$

Moreover, for all $(\bar{X},\bar{V})\in\operatorname{dom}\partial\psi$ and all $(\bar{Y},\bar{W})\in\partial\psi(\bar{X},\bar{V})$, we have $\mathcal{T}(\bar{Y},\bar{W})\neq\emptyset$ and

$$\begin{aligned}\partial\psi(\bar{X},\bar{V}) &= \partial\sigma_{\Omega(A,B)}(\bar{X},\bar{V})+\{0\}\times\partial h(\bar{V})\\ &= \left\{(\bar{Y},\bar{W})\in\mathbb{E}\;\middle|\;\mathcal{T}(\bar{Y},\bar{W})\neq\emptyset\right\}.\end{aligned} \tag{3.7}$$

Proof. Set $f_1(X,V):=\sigma_{\Omega(A,B)}(X,V)$ and $f_2(X,V):=h(V)$. Then part a) follows from Theorem 7.1 b), and part b) follows from Theorem 7.1 c). ◻

3.2. Infimal projection I. We are now in position to prove our first main result about the infimal projection p defined in (3.2).

THEOREM 3.5 (Conjugate of p and properties under CCQ). *Let p be given by (3.2). Moreover, let $q:\mathbb{R}^{n\times m}\to\overline{\mathbb{R}}$ be given by*

$$q: Y\mapsto \inf_{(Y,-W)\in\Omega(A,B)} h^*(W).$$

Then the following hold:

a) $\operatorname{dom} q = \left\{Y\in\mathbb{R}^{n\times m}\;\middle|\;AY=B,\ \left(\tfrac{1}{2}YY^T-\mathcal{K}_A^\circ\right)\cap\operatorname{dom} h^*\neq\emptyset\right\}$.

b) $p^*=\operatorname{cl} q$, hence $\operatorname{dom} q\subset\operatorname{dom} p^*$.

c) If CCQ holds for p, then we have:

I) $p^*=q$, i.e.

$$p^*(Y) = \inf_{(Y,-W)\in\Omega(A,B)} h^*(W). \tag{3.8}$$

Moreover, for all $Y\in\operatorname{dom} p^*$, the infimum is a minimum, i.e. there exists $W\in\operatorname{dom} h^*$ with $(Y,-W)\in\Omega(A,B)$ such that $p^*(Y)=h^*(W)$. In particular, p^* is closed, proper, and convex with $\operatorname{dom} p^*=\operatorname{dom} q$.

II) $p\in\Gamma_0(\mathbb{R}^{n\times m})$ is finite-valued (hence locally Lipschitz).

Proof. a) Obvious.

b) The expression for p^* (without CCQ) follows from [16, Theorem 11.23 c)] and Lemma 3.3 b). The domain containment is clear as $p^* = \mathrm{cl}\, q \leq q$.

c.I) From [16, Theorem 11.23 c)] we have $p^* = \psi^*(\cdot, 0)$, hence Lemma 3.3 c) gives the claimed statements.

c.II) p is convex by Lemma 3.1 a), and it does not take the value $-\infty$ as p^* is proper by I). To prove the desired statement it therefore suffices to see that $\mathrm{dom}\, p = \mathbb{R}^{n\times m}$. To this end, observe, see Lemma 3.1, that

$$\mathrm{dom}\, p = L(\mathrm{dom}\, \sigma_{\Omega(A,B)} \cap \mathbb{R}^{n\times m} \times \mathrm{dom}\, h),$$

where $L : (X, V) \mapsto X$. By CCQ we have $\mathrm{ri}\,(\mathrm{dom}\, h) \cap \mathrm{int}\, \mathcal{K}_A \neq \emptyset$, hence

$$\begin{aligned} \mathrm{ri}\,(\mathrm{dom}\, \sigma_{\Omega(A,B)} \cap \mathbb{R}^{n\times m} \times \mathrm{dom}\, h) &= \mathrm{int}\,(\mathrm{dom}\, \sigma_{\Omega(A,B)}) \cap \mathbb{R}^{n\times m} \times \mathrm{ri}\,(\mathrm{dom}\, h) \\ &= \mathbb{R}^{n\times m} \times \mathrm{int}\, \mathcal{K}_A \cap \mathbb{R}^{n\times m} \times \mathrm{ri}\,(\mathrm{dom}\, h) \\ &= \mathbb{R}^{n\times m} \times \mathrm{int}\, \mathcal{K}_A \cap \mathrm{ri}\,(\mathrm{dom}\, h), \end{aligned}$$

where we use [5, Theorem 4.1] to represent $\mathrm{int}\,(\mathrm{dom}\, \sigma_{\Omega(A,B)})$. This now gives

$$\mathrm{ri}\,(\mathrm{dom}\, p) = L\left[\mathrm{ri}\,(\mathrm{dom}\, \sigma_{\Omega(A,B)} \cap \mathbb{R}^{n\times m} \times \mathrm{dom}\, h)\right] = \mathbb{R}^{n\times m}. \qquad \square$$

We now take a broader perspective on infimal projection by embedding it into a *pertubation duality framework* in the sense of [16, Theorem 11.39] or [1, Chapter 5].

Given $\bar{X} \in \mathbb{R}^{n\times m}$, we define $\psi_{\bar{X}}$ by

$$\psi_{\bar{X}}(X, V) := \psi(X + \bar{X}, V) \quad ((X, V) \in \mathbb{E}).$$

Moreover define $p_{\bar{X}}$ by

$$p_{\bar{X}}(X) := \inf_{V\in\mathbb{S}^n} \psi_{\bar{X}}(X, V) \quad (X \in \mathbb{R}^{n\times m}). \tag{3.9}$$

Then

$$\psi^*_{\bar{X}}(Y, W) = \psi^*(Y, W) - \langle \bar{X}, Y\rangle \quad ((Y, W) \in \mathbb{E}),$$

see [16, Equation 11(3)]. Defining

$$q_{\bar{X}}(W) := -\sup_Y\{\langle \bar{X}, Y\rangle - \psi^*(Y, W)\} \quad (W \in \mathbb{S}^n), \tag{3.10}$$

then $q_{\bar{X}}$ is a proper (see Lemma 3.7 for its domain) and convex function and we have a natural duality pairing of $p_{\bar{X}}$ and $q_{\bar{X}}$ with weak duality reading

$$p_{\bar{X}}(0) \geq -q_{\bar{X}}(0) \quad (\bar{X} \in \mathbb{R}^{n\times m}).$$

Applying the general pertubation duality to our scenario yields the following result.

PROPOSITION 3.6 (Shifted duality for p). *Let p be defined by (3.2), let $\bar{X} \in \mathrm{dom}\, p$ and $q_{\bar{X}}$ be defined by (3.10). Then the following hold:*

a) If $0 \in \mathrm{ri}\,(\mathrm{dom}\, q_{\bar{X}})$ then $p(\bar{X}) = -q_{\bar{X}}(0) \in \mathbb{R}$, $\mathrm{argmax}\, \psi(\bar{X}, \cdot) \neq \emptyset$, and $\partial q_{\bar{X}}(0) \neq \emptyset$.

b) If $\bar{X} \in \mathrm{ri}\,(\mathrm{dom}\, p)$ then $p(\bar{X}) = -q_{\bar{X}}(0) \in \mathbb{R}$, $\mathrm{argmax}_Y\{\langle \bar{X}, Y\rangle - \psi^*(Y, W)\} \neq \emptyset$, and $\partial p(\bar{X}) \neq \emptyset$.

c) Under either condition $0 \in \operatorname{ri}(\operatorname{dom} q_{\bar X})$ *or* $\bar X \in \operatorname{ri}(\operatorname{dom} p)$*,* p *is lsc at* $\bar X$ *and* $-q_{\bar X}$ *is lsc at* 0*.*

d) We have

$$\left.\begin{array}{rl} & p(\bar X) \\ = & \psi(\bar X, \bar V), \\ = & \langle \bar X, \bar Y\rangle - \psi^*(\bar Y, 0), \\ = & -q_{\bar X}(0) \end{array}\right\} \iff (\bar Y, 0) \in \partial\psi(\bar X, \bar V) \iff (\bar Y, 0) \in \partial \psi^*(\bar X, \bar V).$$

Proof. Let $\bar X \in \operatorname{dom} p$ and observe that

$$p(X + \bar X) = p_{\bar X}(X) \quad (X \in \mathbb{R}^{n\times m}),$$

hence, in particular, $p(\bar X) = p(0) \in \mathbb{R}$. Applying [1, Theorem 5.1.2–5.1.5, Corollary 5.1.2] to the duality pair $p_{\bar X}$ and $q_{\bar X}$ and translating from $p_{\bar X}$ at 0 to p at $\bar X$ gives all the desired statements. ∎

The domain of $q_{\bar X}$ is given below. Here, the set

$$\mathcal{C}(A, B) := \{W \in \mathbb{S}^n \mid \exists Y:\ (Y, W) \in \Omega(A, B)\}, \tag{3.11}$$

which will play a crucial role in what follows, occurs naturally.

LEMMA 3.7 (Domain of $q_{\bar X}$). *Let* $\bar X \in \mathbb{R}^{n\times m}$ *and* $q_{\bar X}$ *defined by* (3.10)*. Then*

$$\operatorname{dom} q_{\bar X} = \mathcal{C}(A, B) + \operatorname{dom} h^*.$$

Proof. a) Using Lemma 3.3, observe that

$$\begin{aligned} q_{\bar X}(W) &= \inf_Y \left\{\psi^*(Y, W) - \langle \bar X, Y\rangle\right\} \\ &= \inf_Y \left\{\eta(Y, W) - \langle \bar X, Y\rangle\right\} \\ &= \inf_{(Y,T)\in\Omega(A,B)} \left\{h^*(W - T) - \langle \bar X, Y\rangle\right\}. \end{aligned}$$

Therefore, we have

$$\operatorname{dom} q_{\bar X} = \{W \mid \exists (Y, T) \in \Omega(A, B) : W - T \in \operatorname{dom} h^*\} = \mathcal{C}(A, B) + \operatorname{dom} h^*.$$

Before we proceed with our analysis, we will discuss various constraint qualifications for the optimization problem defining p in the next section.

3.3. Constraint qualifications. We start our analysis with a result about the set $\mathcal{C}(A, B)$ from (3.11), which was used in Lemma 3.7 to represent the domain of $q_{\bar X}$.

LEMMA 3.8 (Properties of $\mathcal{C}(A, B)$). *Let* $\mathcal{C}(A, B)$ *be as in* (3.11)*. Then we have:*

a) $\mathcal{C}(A, B)$ *is closed and convex with* $\mathcal{C}(A, B)^\infty = \mathcal{K}_A^\circ$*.*

b) $\mathcal{C}(A, B) = \operatorname{dom} \sigma_{\Omega(A,B)}(\bar X, \cdot)^*$ *for all* $\bar X$ *such that* $\sigma_{\Omega(A,B)}(\bar X, \cdot)$ *is proper.*

c) We have

$$\begin{aligned} \operatorname{ri} \mathcal{C}(A, B) &= \left\{W \,\middle|\, \exists Y:\ AY = B,\ \frac{1}{2}YY^T + W \in \operatorname{ri}(\mathcal{K}_A^\circ)\right\} \\ &= \operatorname{ri}(\operatorname{dom} \sigma_{\Omega(A,B)}(\bar X, \cdot)^*) \end{aligned}$$

for all $\bar X$ *such that* $\sigma_{\Omega(A,B)}(\bar X, \cdot)$ *is proper.*

Proof. a) With the linear map $T : (Y, W) \mapsto W$ we have $\mathcal{C}(A,B) = T(\Omega(A,B))$. Therefore $\mathcal{C}(A,B)$ is convex. By [6, Proposition 10] we have $\Omega(A,B)^\infty = \{0\} \times \mathcal{K}_A^\circ$. Therefore, $\ker T \cap \Omega(A,B)^\infty = \{0\}$. Hence [16, Theorem 3.10] gives the rest of a).

b) Apply Corollary 7.2 to $\bar{g} := \sigma_{\Omega(A,B)}(\bar{X}, \cdot)$ to infer that

$$\bar{g}^*(W) = \inf_{Y:(Y,W)\in\Omega(A,B)} \langle -\bar{X}, Y \rangle \quad (W \in \mathbb{S}^n).$$

This proves the claim.

c) Observe that $\operatorname{ri}\mathcal{C}(A,B) = \operatorname{ri} T(\Omega(A,B) = T(\operatorname{ri}\Omega(A,B))$ and use [6, Proposition 8] to get the first representation. The second one follows from b). ∎

We now define the constraint qualifications central to our study. Note that CCQ was already defined earlier.

Definition 3.9 (Constraint qualifications). *Let p be given by (3.2). We say that p satisfies*

i) PCQ *if* $0 \in \operatorname{ri}(\operatorname{dom} h^* + \mathcal{C}(A,B))$*;*
ii) strong PCQ (SPCQ) *if* $0 \in \operatorname{int}(\operatorname{dom} h^* + \mathcal{C}(A,B))$*;*
iii) boundedness PCQ (BPCQ) *if* $\operatorname{dom} h \cap \mathcal{K}_A \neq \emptyset$ *and* $(\operatorname{dom} h)^\infty \cap \mathcal{K}_A = \{0\}$*;*
iv) CCQ *if* $\operatorname{ri}(\operatorname{dom} h) \cap \operatorname{int}\mathcal{K}_A \neq \emptyset$*.*

Note that PCQ stands for *primal constraint qualification* and CCQ for *conjugate constraint qualification*.

The next results clarify the relations between the various constraint qualifications. We lead with characterizations of PCQ and BPCQ.

Lemma 3.10 (Characterizations of (B)PCQ). *Let p be given by (3.2) and let*

$$f_{\bar{X}} := \psi(\bar{X}, \cdot) \quad (\bar{X} \in \mathbb{R}^{n\times m}). \tag{3.12}$$

Let $\bar{X} \in \operatorname{dom} p$. Then the following hold:

a) The following are equivalent:
i) $0 \in \operatorname{ri}(\operatorname{dom} f_{\bar{X}}^*)$*;*
ii) PCQ holds for p;
iii) $\exists Y \in \mathbb{R}^{n\times m} : \ AY = B, \quad \frac{1}{2}YY^T \in \operatorname{ri}(\mathcal{K}_A^\circ + \operatorname{dom} h^*)$.
In addition, similar characterizations of SPCQ hold by substituting the relative interior for the interior.
b) BPCQ holds for p if and only if $\operatorname{dom} h \cap \mathcal{K}_A$ is nonempty and bounded.

Proof. a) Defining $g_{\bar{X}} := \sigma_{\Omega(A,B)}(\bar{X}, \cdot)$, we find that $f_{\bar{X}}^* = \operatorname{cl}(g_{\bar{X}}^* \,\square\, h^*)$ and therefore $\operatorname{ri}(\operatorname{dom} f_{\bar{X}}^*) = \operatorname{ri}(\operatorname{dom} g_{\bar{X}}^* + \operatorname{dom} h^*) = \operatorname{ri}(\mathcal{C}(A,B) + \operatorname{dom} h^*)$, see Lemma 3.8 c). This proves the first two equivalences. The third follows readily from the representation of $\operatorname{ri}(\Omega(A,B))$ from [6, Proposition 8].

b) Follows readily from [16, Theorem 3.5, Proposition 3.9]. ∎

We point out that, under PCQ, Lemma 3.10 shows that the objective functions $\psi(\bar{X}, \cdot)$ ($\bar{X} \in \operatorname{dom} p$) occuring in the definition of p in (3.2) are *weakly coercive* when proper, see [1, Theorem 3.2.1]. The latter reference tells us that the infimum in (3.2) is attained under PCQ if finite, a fact that will be stated again (and derived alternatively) in Theorem 3.14. Under SPCQ, the objective functions $\psi(\bar{X}, \cdot)$ ($\bar{X} \in \operatorname{dom} p$) are *level-bounded* (or *coercive*), in which case the argmin $\psi(\bar{X}, \cdot)$ is nonempty and compact (and clearly convex).

The next result shows the relations between the different notions of PCQ.

LEMMA 3.11. *Let p be given by (3.2). Then the following hold:*

a) BPCQ $\implies$ SPCQ $\implies$ PCQ.

b) If $\mathrm{int}\,(\mathrm{dom}\,h^*) \cap \mathrm{int}\,(-\mathcal{C}(A,B)) \neq \emptyset$ then PCQ and SPCQ are equivalent.

Proof. a) The first implication can be seen as follows: If BPCQ holds then $\mathrm{dom}\, f_{\bar X} \subset \mathrm{dom}\, h \cap \mathcal{K}_A$ is bounded (and nonempty exactly if $\bar X \in \mathrm{dom}\, p$). Therefore $f_{\bar X}$ is level-bounded for all $\bar X \in \mathrm{dom}\, p$, i.e. $0 \in \mathrm{int}\,(\mathrm{dom}\, f^*_{\bar X})$ ($\bar X \in \mathrm{dom}\, p$), see e.g. [16, Theorem 11.8]. In view of Lemma 3.10 a) this implies that SPCQ holds.

The second implication is trivial.

b) Obvious from the definitions. ∎

We now provide characterizations for CCQ.

LEMMA 3.12 (Characterizations of CCQ). *Let p be given by (3.2). Then*
i) $\mathrm{dom}\, h \cap \mathrm{int}\,\mathcal{K}_A \neq \emptyset \iff$ *ii) CCQ holds for p* $\iff$ *iii)* $(-\mathcal{K}_A^\circ) \cap \mathrm{hzn}\, h^* = \{0\}$.

Proof. The first equivalence is a direct consequence of the *line segment principle* (cf. [15, Theorem 6.1]): The fact that ii) implies i) is obvious. For the converse direction let $y \in \mathrm{dom}\, h \cap \mathrm{int}\,\mathcal{K}_A$ and pick $x \in \mathrm{ri}\,(\mathrm{dom}\, h)$. Then $z_\lambda := \lambda x + (1-\lambda) y \in \mathrm{ri}\,(\mathrm{dom}\, h)$ for all $\lambda \in (0,1]$. Letting $\lambda \downarrow 0$ we find that $z_\lambda \in \mathrm{ri}\,(\mathrm{dom}\, h) \cap \mathrm{int}\,\mathcal{K}_A$ for all $\lambda \in (0,1]$ sufficiently small, which proves that $\mathrm{ri}\,(\mathrm{dom}\, h) \cap \mathrm{int}\,\mathcal{K}_A \neq \emptyset$.

The second equivalence can be seen as follows: We apply [15, Corollary 16.2.2] (to $f_1 := h$ and $f_2 := \delta_{\mathcal{K}_A}$). This result tells us that $\mathrm{ri}\,(\mathrm{dom}\, h) \cap \mathrm{int}\,\mathcal{K}_A \neq \emptyset$ if and only if there does not exist a matrix $W \in \mathbb{S}^n$ such that

$$(h^*)^\infty(W) + \sigma_{\mathcal{K}_A}(-W) \le 0 \quad \text{and} \quad (h^*)^\infty(-W) + \sigma_{\mathcal{K}_A}(W) > 0. \tag{3.13}$$

Since $\sigma_{\mathcal{K}_A}(-W) = \delta_{\mathcal{K}_A^\circ}(-W)$, the first of these conditions is equivalent to the condition $W \in (-\mathcal{K}_A^\circ) \cap \mathrm{hzn}\, h^*$. In particular, we can infer that $(-\mathcal{K}_A^\circ) \cap \mathrm{hzn}\, h^* = \{0\}$ gives the inconsistency of (3.13) and thus establishes iii)⇒ii).

The second condition in (3.13) implies $W \neq 0$. Thus, in view of Proposition 2.1 b), $0 \neq -W \in \mathcal{K}_A^\circ \subset \mathbb{S}^n_+$, and hence $W \notin \mathcal{K}_A^\circ$. Thus, every nonzero element of the set $(-\mathcal{K}_A^\circ) \cap \mathrm{hzn}\, h^*$ satisfies (3.13). Thus, the nonexistence of a W satisfying (3.13) implies that $(-\mathcal{K}_A^\circ) \cap \mathrm{hzn}\, h^* = \{0\}$, which altogether proves the result. ∎

We note that for any proper, convex function f we always have $\mathrm{hzn}\, f \subset (\mathrm{dom}\, f)^\infty$ which, in view of Lemma 3.12, implies that the condition

$$(-\mathcal{K}_A^\circ) \cap (\mathrm{dom}\, h^*)^\infty = \{0\} \tag{3.14}$$

is stronger than CCQ. However, we do not use it in our subsequent study.

Moreover, since $\mathcal{K}_A = \mathbb{S}^n$ if (and only if) A has full column rank we have

$$\mathrm{rank}\, A = n \quad \implies \quad \text{CCQ}.$$

3.4. Infimal projection II. We return to our analysis of the infimal projection defining p in (3.2). The following result reveals that the two critical conditions $0 \in \mathrm{ri}\,(\mathrm{dom}\, q_{\bar X})$ and $\bar X \in \mathrm{ri}\,(\mathrm{dom}\, p)$, respectively, that occured in (3.6), embed nicely into our constraint qualifications studied in Section 3.3.

COROLLARY 3.13. *Let p be defined by (3.2), let $\bar X \in \mathrm{dom}\, p$ and $q_{\bar X}$ be defined by (3.10). Then the following hold:*

a) PCQ holds for p if and only if $0 \in \mathrm{ri}\,(\mathrm{dom}\, q_{\bar X})$;

b) If CCQ holds then $\bar X \in \mathrm{ri}\,(\mathrm{dom}\, p)$.

Proof. a) Follows immediately from Lemma 3.7 and the definition of PCQ.
b) Under CCQ we have $\operatorname{dom} p = \mathbb{R}^{n\times m}$, see Theorem 3.5, hence b) follows. ◻

As a consequence of Corollary 3.13 and Proposition 3.6 we can add to the properties of p proven in Theorem 3.5.

THEOREM 3.14 (Properties of p under PCQ). *Let p be defined by (3.2) such that PCQ is satisfied and let $q_{\bar X}$ be given by (3.10). Then the following hold:*

a) $p \in \Gamma_0(\mathbb{R}^{n\times m})$*;*
b) $\operatorname{argmin}_V \psi(\bar X, V) \neq \emptyset \quad (\bar X \in \operatorname{dom} p)$ *(primal attainment);*
c) $p(\bar X) = q_{\bar X}(0) \quad (\bar X \in \operatorname{dom} p)$ *(strong duality).*

Proof. a) Under PCQ, by Corollary 3.13, we have $0 \in \operatorname{ri}(\operatorname{dom} q_{\bar X})$ for all $\bar X \in \operatorname{dom} p$. Hence, by Proposition 3.6 c), p is lsc at $\bar X \in \operatorname{dom} p$. Since p is proper and convex, see Lemma 3.1, this shows that $p \in \Gamma_0$.
b), c) Follows readily from Corollary 3.13 and Proposition 3.6 a). ◻

We note that Theorem 3.14 could have been proven entirely without using the shifted duality framework from Proposition 3.6, but by using the following approach: With the linear projection $L : (X, V) \to X$ which has been used implicitly throughout our study, it can be seen that $p = L\psi$ is a *linear image* in the sense of [15, p. 38]. Then [15, Theorem 9.2] gives all statements from Proposition 3.14. This can be seen after realizing that the constraint qualification from the latter reference, which for $p = L\psi$ reads

$$\psi(0, V) > 0 \quad \text{or} \quad \psi^\infty(0, -V) \leq 0 \quad (V \in \mathbb{S}^n),$$

as $\ker L = \{0\} \times \mathbb{S}^n$, is exactly PCQ, which, however, also takes some effort. For the sake of uniformity, we have chosen to derive Theorem 3.14 from the shifted duality scheme, which will also be serviceable for our subsequent subdifferential analysis.

The next result follows readily from the foregoing analysis.

COROLLARY 3.15. *Let p be given by (3.2). If PCQ and CCQ are satisfied for p then the following hold:*

a) $p \in \Gamma_0(\mathbb{R}^{n\times m})$ *is finite-valued and for all* $\bar X \in \mathbb{R}^{n\times m}$ *there exists* $\bar V$ *such that* $p(\bar X) = \psi(\bar X, \bar V)$.
b) $p^* = q$ *and for all* $\bar Y \in \operatorname{dom} p^*$ *there exists* $\bar W$ *such that* $(\bar Y, \bar W) \in \Omega(A, B)$ *and* $p^*(\bar Y) = h^*(-\bar W)$.

Proof. Follows from Theorem 3.5. ◻

The table below summarizes most of our findings so far. Here $\bar X \in \operatorname{dom} p$ and $\bar Y \in \operatorname{dom} p^*$.

Consequence\Hypothesis	-	PCQ	SPCQ	BPCQ	CCQ	PCQ + CCQ
$p \in \Gamma$	✓	✓	✓	✓	✓	✓
$p \in \Gamma_0$		✓	✓	✓	✓	✓
$p(\bar X) = -q_{\bar X}(0)$		✓	✓	✓	✓	✓
$\operatorname{argmin} \psi(\bar X, \cdot) \neq \emptyset$		✓	✓	✓		✓
$\operatorname{argmin} \psi(\bar X, \cdot)$ compact			✓	✓[1]		✓
$\operatorname{dom} p = \mathbb{R}^{n\times m}$					✓	✓
$p = p^{**}$		✓	✓	✓	✓	✓
$\operatorname{argmin}_{(\bar Y, T) \in \Omega(A,B)} h^*(-T) \neq \emptyset$					✓	✓

In view of Proposition 3.6 b) and Corollary 3.13 one might be inclined to think that using CCQ instead of the pointwise condition $0 \in \mathrm{ri}\,(\mathrm{dom}\, p)$ is excessively strong. However, computing the relative interior of $\mathrm{dom}\, p$ without CCQ is problematic, cf. the derivations in the proof of Theorem 3.5 c.II) under CCQ. Moreover, CCQ is exactly what is needed to establish desirable properties of p^*, see Theorem 3.5 c.I). Hence, we do not consider constraint qualifications weaker than CCQ.

We now turn our attention to subdifferentiation of p.

Proposition 3.16 (Subdifferential of p). *Let p be given by (3.2). Then the following hold:*

a) Under CCQ we have

$$\partial p(\bar{X}) = \underset{Y}{\mathrm{argmax}}\{\langle \bar{X},\, Y\rangle - \inf_{(Y,T)\in\Omega(A,B)} h^*(-T)\}, \tag{3.15}$$

which is nonempty and compact.

b) Under PCQ equation (3.15) holds, and, for $\bar{X} \in \mathrm{dom}\, p$, we have

$$\begin{aligned}\partial p(\bar{X}) &= \{\bar{Y} \mid \exists \bar{V} :\ (\bar{Y}, 0) \in \partial\psi(\bar{X}, \bar{V})\}\\ &= \{\bar{Y} \mid \exists \bar{V} :\ (\bar{X}, \bar{V}) \in \partial\psi^*(\bar{Y}, 0)\}\\ &= \{\bar{Y} \mid \exists \bar{V} : p(\bar{X}) = \psi(\bar{X}, \bar{V}) = \langle \bar{X},\, \bar{Y}\rangle - p^*(\bar{Y})\}.\end{aligned}$$

c) Under PCQ and CCQ, we have

$$\partial p(\bar{X}) = \{Y \mid \exists \bar{V}, \bar{T} :\ -\bar{T} \in \partial h(\bar{V}),\ (Y, \bar{T}) \in \partial\sigma_{\Omega(A,B)}(\bar{X}, \bar{V})\},$$

which is compact and nonempty.

Proof. a) Under CCQ, p is convex and finite-valued (hence closed and proper), therefore (3.15) follows from [15, Theorem 23.5] and the fact that the closure for p^* can be dropped in the argmax problem.

Moreover, we have $\mathrm{dom}\, p = \mathbb{R}^{n\times m}$, which gives the remaining statements in a).

b) Under PCQ we also have that $p \in \Gamma_0$, hence the same reasoning as in a) gives (3.15). We now prove the remainder: For the first identity notice that (see e.g. [10, Chapter D, Corollary 4.5.3])

$$\partial p(\bar{X}) = \{Y \mid (Y, 0) \in \partial\psi(\bar{X}, \bar{V})\} \quad (\bar{V} \in \underset{V}{\mathrm{argmin}}\, \psi(\bar{X}, V)),$$

the latter argmin set being nonempty due to what was argued above. The '$\subset$'-inclusion is hence clear. For the reverse inclusion invoke also [16, Example 10.12] to see that if $(Y, 0) \in \psi(\bar{X}, \bar{V})$ then $\bar{V} \in \mathrm{argmin}_V\, \psi(\bar{X}, V)$.

The second identity in c) is clear from [15, Theorem 23.5] as $\psi \in \Gamma_0(\mathbb{E})$.

The third follows from Proposition 3.6 in combination with Corollary 3.13 and recalling that $\psi^*(\bar{Y}, 0) = p^*(\bar{Y})$.

c) Apply Corollary 3.4 to the first representation in b). ∎

For $\bar{X} \in \mathrm{rbd}\,(\mathrm{dom}\, p)$ the subdifferential $\partial p(\bar{X})$ can be empty. Moreover, it is unbounded if $\bar{X} \notin \mathrm{int}\,(\mathrm{dom}\, p)$. The latter may even occur under BPCQ as the following example shows.

[1] $\mathrm{dom}\, \psi(\bar{X}, \cdot)$ is bounded.

EXAMPLE 3.17. *Let* $A = \left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right)$ *and* $b = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ *so that*

$$\mathcal{K}_A = \{\left(\begin{smallmatrix}v&w\\w&u\end{smallmatrix}\right) \mid u \geq 0\}.$$

Defining $h := \delta_{\mathcal{V}}$ *for*

$$\mathcal{V} := \{\left(\begin{smallmatrix}v&0\\0&u\end{smallmatrix}\right) \mid u \leq 0, v \in [0,1]\}$$

we hence find that

$$\operatorname{dom} h \cap \mathcal{K}_A = \{\left(\begin{smallmatrix}v&0\\0&0\end{smallmatrix}\right) \mid v \in [0,1]\} \quad \text{and} \quad \operatorname{dom} h \cap \operatorname{int} \mathcal{K}_A = \emptyset,$$

so that CCQ is violated but BPCQ (hence (S)PCQ) holds. We find that

$$\begin{aligned} x \in \operatorname{dom} p &\iff \exists V \in \mathcal{V} \cap \mathcal{K}_A : \left(\begin{smallmatrix}x\\b\end{smallmatrix}\right) \in \operatorname{rge} \begin{pmatrix} V & A^T \\ A & 0 \end{pmatrix} \\ &\iff \exists v \in [0,1], r, s \in \mathbb{R}^2 : \begin{array}{rcl} x &=& \left(\begin{smallmatrix}v&0\\0&0\end{smallmatrix}\right) r + \left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right) s, \\ \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) &=& \left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right) r \end{array} \\ &\iff \exists v \in [0,1], \rho, \sigma \in \mathbb{R} : x = \left(\begin{smallmatrix}v&0\\0&0\end{smallmatrix}\right)[\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) + \rho \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)] + \sigma \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) \\ &\iff x \in \operatorname{span}\{\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)\}. \end{aligned}$$

Therefore we have $\operatorname{dom} p = \operatorname{span}\{\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)\}$. *In particular,* $\operatorname{dom} p$ *is a proper subspace of* $\mathbb{R}^2$*, hence relatively open with empty interior. Therefore* $\partial p(x)$ *is nonempty and unbounded for any* $x \in \operatorname{dom} p$.

4. h is a support function.

We now study the case where h is a support function. Concretely, given a closed, convex set $\mathcal{V} \subset \mathbb{S}^n$, we consider the function $p : \mathbb{R}^{n \times m} \to \overline{\mathbb{R}}$ given by

$$p(X) := \inf_{V \in \mathbb{S}^n} \sigma_{\Omega(A,B)}(X, V) + \sigma_{\mathcal{V}}(V). \tag{4.1}$$

Recall that, by *Hörmander's Theorem*, see e.g. [15, Corollary 13.2.1], this covers exactly the cases where h is positively homogeneous (and closed, proper, convex).

We commence by analyzing the constraint qualifications from Section 3.3 in the case that h is a support function. Here, and for the remainder of this section, observe that the choice $h = \sigma_{\mathcal{V}}$ implies that $\operatorname{dom} h = \operatorname{bar} \mathcal{V}$ and $\operatorname{dom} h^* = \mathcal{V}$.

LEMMA 4.1 (Constraint qualifications for (4.1)). *Let* p *be given by* (4.1). *Then the following hold:*

a) (CCQ) The conditions

$$\operatorname{bar} \mathcal{V} \cap \operatorname{int} \mathcal{K}_A \neq \emptyset, \tag{4.2}$$

$$\mathcal{V}^\infty \cap (-\mathcal{K}_A^\circ) = \{0\}, \tag{4.3}$$

$$\operatorname{cl}(\operatorname{bar} \mathcal{V}) - \mathcal{K}_A = \mathbb{S}^n \tag{4.4}$$

are each equivalent to CCQ for p.

b) (PCQ) PCQ holds for p *if and only if*

$$\operatorname{pos}(\mathcal{C}(A,B) + \mathcal{V}) = \operatorname{span}(\mathcal{C}(A,B) + \mathcal{V}). \tag{4.5}$$

c) (BPCQ) The conditions

$$\operatorname{bar} \mathcal{V} \cap \mathcal{K}_A \neq \emptyset \quad \text{and} \quad \operatorname{cl}(\operatorname{bar} \mathcal{V}) \cap \mathcal{K}_A = \{0\}, \tag{4.6}$$

$$\operatorname{bar} \mathcal{V} \cap \mathcal{K}_A \neq \emptyset \text{ is bounded}, \tag{4.7}$$

$$\operatorname{bar} \mathcal{V} \cap \mathcal{K}_A \neq \emptyset \quad \text{and} \quad \mathcal{V}^\infty + \mathcal{K}_A^\circ = \mathbb{S}^n \tag{4.8}$$

are each equivalent to BPCQ for p*, hence imply* (4.5).

Proof. Observe that with $h = \sigma_{\mathcal{V}}$ we have $\operatorname{dom} h = \operatorname{bar} \mathcal{V}$ and $\operatorname{hzn} h^* = \mathcal{V}^\infty$.
a) (4.2) is condition i) in Lemma 3.12 for $h = \sigma_{\mathcal{V}}$, while (4.3) is condition iii).
Employing [3, Section 3.3, Exercise 16]) we have

$$(4.3) \iff \operatorname{cl}(\operatorname{bar} \mathcal{V} - \mathcal{K}_A) = \mathbb{S}^n.$$

This completes the proof of a).
b) This is just an application of (2.1).
c) Using (2.6), we see that (4.6) is exactly BPCQ (for $h = \sigma_{\mathcal{V}}$), while the equivalence to (4.7) follows from Lemma 3.10 b). The equivalence of (4.8) to the former follows from the fact that

$$(4.6) \iff \operatorname{cl}(\mathcal{V}^\infty + \mathcal{K}_A^\circ) = \mathbb{S}^n,$$

see [3, Section 3.3, Exercise 16]), where the closure can be dropped by interpreting [15, Theorem 6.3] accordingly. ∎

By the additivity of support functions, see (2.5), we find that

$$p(X) = \inf_{V \in \mathbb{S}^n} \sigma_\Sigma(X, V) \quad (X \in \mathbb{R}^{n\times m}), \tag{4.9}$$

where

$$\Sigma := \Sigma(A, B, \mathcal{V}) := \Omega(A, B) + \{0\} \times \mathcal{V} \subset \mathbb{E}. \tag{4.10}$$

This facilitates some of the analysis.

PROPOSITION 4.2. *Let p be given by (4.1). Then the following hold:*

a) $p \in \Gamma_0$ (i.e. $p = p^{**}$) under any of the conditions in (4.2)-(4.4) or (4.5). In particular this holds under any condition (4.6)-(4.8). Under any of the conditions (4.2)-(4.4) p also finite-valued.

b) $p^* = \delta_{\operatorname{cl}\Sigma}(\cdot, 0)$ where the closure is superfluous (i.e. Σ is closed), in particular, under any condition (4.2)-(4.4).

Proof. a) Follows respectively from Lemma 4.1, Theorem 3.5 c) and Theorem 3.14.
b) By [16, Exercise 3.12], Σ is closed if $(-\mathcal{K}_A^\circ) \cap \mathcal{V}^\infty = \{0\}$, i.e. under any condition in (4.2)-(4.4), see Lemma 4.1 a). The rest follows from [16, Proposition 11.23 (c)]. ∎

We are now interested in computing refined representations for the conjugate of p given by (4.1).

COROLLARY 4.3. *Consider the function p from (4.1) with $\mathcal{V} \subset \mathbb{S}^n$ nonempty, closed and convex. Under any condition (4.2)-(4.4) we have*

$$p^* = \delta_{\Xi(A,B)}$$

where

$$\begin{aligned}\Xi(A, B) &:= \{Y \mid \exists W \in \mathcal{V} : (Y, -W) \in \Omega(A, B)\}\\ &= \left\{ Y \;\middle|\; AY = B,\ \left(\frac{1}{2}YY^T - \mathcal{K}_A^\circ\right) \cap \mathcal{V} \neq \emptyset \right\}.\end{aligned}$$

In particular, we have $p = \sigma_{\Xi(A,B)}$ which is finite-valued.

Proof. By Theorem 3.5 c) and Lemma 4.1 we find that

$$p^*(Y) = \inf_{(Y,-W)\in\Omega(A,B)} \delta_{\mathcal{V}}(W) = \begin{cases} 0 & \text{if} \quad \exists W \in \mathcal{V} : (Y,-W) \in \Omega(A,B), \\ +\infty & \text{else}, \end{cases}$$ ◻

which shows that $p^* = \delta_\Xi(A,B)$. The fact about p follows from Proposition 4.2 a).

4.1. The case $B = 0$. We now consider the case when $B = 0$. Recall from [6, Theorem 11] that this implies that $\sigma_{\Omega(A,0)}$ is a gauge function. Similarly, if $0 \in \mathcal{V}$, then $\sigma_{\mathcal{V}}$ is also a gauge, in fact, $\sigma_{\mathcal{V}} = \gamma_{\mathcal{V}^\circ}$, cf. [16, Example 11.19].

This combination of assumptions has interesting consequences when the geometries of the sets $\mathcal{V}$ and $-\mathcal{K}_A^\circ$ are compatible in the following sense.

Definition 4.4 (Cone compatible gauges). *Given a closed, convex cone $K \subset \mathcal{E}$, we define an ordering on $\mathcal{E}$ by $x \preceq_K y$ if and only if $y - x \in K$. A gauge γ on $\mathcal{E}$ is said to be compatible with this ordering if and only if*

$$\gamma(x) \le \gamma(y) \quad \textit{whenever} \quad 0 \preceq_K x \preceq_K y.$$

The following lemma provides a characterization of cone compatible gauges.

Lemma 4.5 (Cones and compatible gauges). *Let $0 \in C \subset \mathcal{E}$ be a closed, convex set, and let $K \subset \mathcal{E}$ be a closed, convex cone. Then γ_C is compatible with the ordering $\preceq_K$ if and only if*

$$K \cap (y - K) \subset C \quad (y \in K \cap C). \tag{4.11}$$

Proof. Note that, for $y \in K$, we have

$$K \cap (y - K) = \{x \mid 0 \preceq_K x \preceq_K y\}.$$

Suppose that γ_C is compatible with K, and let $y \in C \cap K$. If $x \in K \cap (y - K)$, then $\gamma_C(x) \le \gamma_C(y) \le 1$, and, consequently, $K \cap (y - K) \subset C$.

Next suppose (4.11) holds, and let $x, y \in \mathcal{E}$ be such that $0 \preceq_K x \preceq_K y$. Then, $y \in K$ and $x \in K \cap (y - K)$. We need to show that $\gamma_C(x) \le \gamma_C(y)$. If $\gamma_C(y) = +\infty$, this is trivially the case, so we may as well assume that $\gamma_C(y) =: \bar{t} < +\infty$. If $\bar{t} > 0$, then $\bar{t}^{-1} y \in C \cap K$ and $\bar{t}^{-1} x \in K \cap (\bar{t}^{-1} y - K) \subset C$. Hence, $\gamma_C(\bar{t}^{-1} y) = 1 \ge \gamma_C(\bar{t}^{-1} x)$, and so, $\gamma_C(x) \le \gamma_C(y)$ as desired. In turn, if $\bar{t} = 0$, then $ty \in K \cap C$ $(t > 0)$, so that $tx \in K \cap (ty - K) \subset C$ $(t > 0)$, i.e., $x \in C^\infty$ and so $\gamma_C(x) = 0$. ◻

Corollary 4.6 (Infimal projection with a gauge function). *Let p be given by (4.1) where $\mathcal{V}$ is a nonempty, closed, convex subset of $\mathbb{S}^n$. Suppose that $B = 0$. Then the following hold:*

a) Under any of the conditions (4.2)-(4.4) we have

$$p^* = \delta_{\left\{Y \,\middle|\, AY=0,\ \exists W\in\mathcal{V} : AW=0,\ \frac{1}{2}YY^T \preceq W\right\}}. \tag{4.12}$$

b) If $0 \in \mathcal{V}$ and $\gamma_{\mathcal{V}}$ is compatible with the ordering induced by $-\mathcal{K}_A^\circ$ then

$$\begin{aligned} p^*(Y) &= \delta_{\left\{Y \,\middle|\, AY=0,\ \gamma_{\mathcal{V}}\left(\frac{1}{2}YY^T\right)\le 1\right\}}(Y) \\ &= \delta_{(-\mathcal{K}_A^\circ)\cap\mathcal{V}}\left(\frac{1}{2}YY^T\right). \end{aligned} \tag{4.13}$$

Proof. a) Follows readily from Corollary 4.3 by setting $B = 0$ and using the representation of $\mathcal{K}_A$ in Proposition 2.1.
b) First observe that $-\mathcal{K}_A^\circ = \left\{W \in \mathbb{S}^n_+ \mid \operatorname{rge} W \subset \ker A\right\}$, see Proposition 2.1 b), recall that $\operatorname{rge} Y = \operatorname{rge} YY^T$ $(Y \in \mathbb{R}^{n\times m})$ and $V \in \mathcal{V}$ if and only if $\gamma_{\mathcal{V}}(V) \le 1$. Exploiting these facts, we see that

$$
\begin{aligned}
& AY = 0,\ \exists W \in \mathcal{V} : AW = 0, \frac{1}{2}YY^T \preceq W \\
\iff & AY = 0,\ \exists W \in \mathcal{V} :\ \gamma_{\mathcal{V}}(W) \ge \gamma_{\mathcal{V}}\left(\frac{1}{2}YY^T\right) \\
\iff & AY = 0,\ \gamma_{\mathcal{V}}\left(\frac{1}{2}YY^T\right) \le 1 \\
\iff & AY = 0,\ \frac{1}{2}YY^T \in \mathcal{V} \\
\iff & \operatorname{rge} YY^T \subset \ker A,\ \frac{1}{2}YY^T \in \mathcal{V} \\
\iff & \frac{1}{2}YY^T \in (-\mathcal{K}_A^\circ) \cap \mathcal{V}.
\end{aligned}
$$

Therefore b) follows from a). ◻

Linear functionals are special instances of support functions. We hence obtain the following remarkable result as a consequence of our more general analysis above. Here $\|\cdot\|_*$ denotes the nuclear norm[2].

COROLLARY 4.7 (h linear). *Let $p : \mathbb{R}^{n\times m} \to \overline{\mathbb{R}}$ be defined by*

$$p(X) = \inf_{V \in \mathbb{S}^n} \sigma_{\Omega(A,0)}(X, V) + \langle \bar{U}, V\rangle$$

for some $\bar{U} \in \mathbb{S}^n_+ \cap \operatorname{Ker}_n A$ and $C(\bar{U}) := \left\{Y \,\middle|\, \frac{1}{2}YY^T \preceq \bar{U}\right\}$. Then we have:

a) $p^* = \delta_{C(\bar{U})\cap \operatorname{Ker}_n A}$ is closed, proper, convex.

b) $p = \sigma_{C(\bar{U})\cap \operatorname{Ker}_n A} = \gamma_{C(\bar{U})^\circ + \operatorname{Rge}_n A^T}$ is sublinear, finite-valued, nonnegative and symmetric (i.e. a seminorm).

c) If $\bar{U} \succ 0$ with $2\bar{U} = LL^T$ $(L \in \mathbb{R}^{n\times n})$ and $A = 0$ then

$$p = \sigma_{C(\bar{U})} = \|L^T(\cdot)\|_*,$$

i.e. p is a norm with $C(\bar{U})^\circ$ as its unit ball and $\gamma_{C(\bar{U})}$ as its dual norm.

d) If $\bar{U}$ is positive definite, $C(\bar{U})$ and $C(\bar{U})^\circ$ are compact, convex, symmetric[3] with 0 in their interior, thus $\operatorname{pos} C(\bar{U}) = \operatorname{pos} C(\bar{U})^\circ = \mathbb{S}^n$.

Proof. a) Observe that $h := \langle \bar{U}, \cdot\rangle = \sigma_{\{\bar{U}\}}$. Hence the machinery from above applies with $\mathcal{V} = \{\bar{U}\}$. As $\mathcal{V}$ is bounded, CCQ is trivially satisfied (cf. (4.2)-(4.4)) and the representation of p^* follows from Corollary 4.6 a).
b) We have

$$
\begin{aligned}
p &= p^{**} \\
&= \sigma_{C(\bar{U})\cap \operatorname{Ker}_n A} \\
&= \gamma_{(C(\bar{U})\cap \operatorname{Ker}_n A)^\circ} \\
&= \gamma_{\operatorname{cl}(C(\bar{U})^\circ + \operatorname{Rge}_n A^T)} \\
&= \gamma_{C(\bar{U})^\circ + \operatorname{Rge}_n A^T}.
\end{aligned}
$$

[2]For a matrix T the nuclear norm $\|T\|_*$ is the sum of its singular values.
[3]We say the set $S \subset \mathcal{E}$ symmetric if $S = -S$.

As CCQ holds, the first identity is due to Proposition 4.2. The second uses a), the third follows from [15, Theorem 14.5]. The sublinearity of p is clear. The finite-valuedness follows from Proposition 4.2. Since $0 \in C(\bar{U})$ the nonnegativity follows as well, and the symmetry is due to the symmetry of $C(\bar{U})$.

c) Consider the case $\bar{U} = \frac{1}{2}I$: By part a), we have $p^* = \delta_{\{Y \mid YY^T \preceq I\}}$. Observe that

$$\left\{Y \;\middle|\; YY^T \preceq I\right\} = \{Y \mid \|Y\|_2 \leq 1\} =: \mathbb{B}_\Lambda$$

is the closed unit ball of the spectral norm. Therefore, $p = \sigma_{\mathbb{B}_\Lambda} = \|\cdot\|_{\mathbb{B}_\Lambda^\circ} = \|\cdot\|_*$.

To prove the general case suppose that $2\bar{U} = LL^T$. Then it is clear that $C(\bar{U}) = \left\{Y \;\middle|\; L^{-1}Y \in C(\frac{1}{2}I)\right\}$, and therefore

$$\begin{aligned} p(X) &= \sigma_{C(\bar{U})}(X) \\ &= \sup_{Y: L^{-1}Y \in C(\frac{1}{2}I)} \langle Y, X\rangle \\ &= \sup_{L^{-1}Y \in C(\frac{1}{2}I)} \left\langle L^{-1}Y, L^T X\right\rangle \\ &= \sigma_{C(\frac{1}{2}I)}(L^T X) \\ &= \|L^T X\|_*. \end{aligned}$$

Here the first identity is due to part b) (with $A = 0$) and the last one follows from the special case considered above.

d) Follows from c) using [15, Theorem 15.2]. ∎

We point out that Corollary 4.7 generalizes the nuclear norm smoothing result by Hsieh and Olsen [13, Lemma 1] and complements [5, Theorem 5.7]

5. h is an indicator function.

We now suppose that the function h in (3.1) is given by $h := \delta_{\mathcal{V}}$ for some nonempty, closed, and convex set $\mathcal{V} \in \mathbb{S}^n$, i.e., in this section, the infimal projection $p : \mathbb{R}^{n\times m} \to \overline{\mathbb{R}}$ is given by

$$p(X) = \inf_{V \in \mathbb{S}^n} \sigma_{\mathcal{D}(A,B)}(X, V) + \delta_{\mathcal{V}}(V). \tag{5.1}$$

We first want to discuss the constraint qualifications from Section 3.3 in this particular case. Here, and for the remainder of this section, observe that the choice $h = \delta_{\mathcal{V}}$ implies that $\operatorname{dom} h = \mathcal{V}$ and $\operatorname{dom} h^* = \operatorname{bar} \mathcal{V}$.

LEMMA 5.1 (Constraint qualifications for (5.1)). *Let p be given by (5.1). Then the following hold:*

a) (CCQ) The conditions

$$\mathcal{V} \cap \operatorname{int} \mathcal{K}_A \neq \emptyset, \tag{5.2}$$

$$\overline{\operatorname{cone}}\, \mathcal{V} - \mathcal{K}_A = \mathbb{S}^n \tag{5.3}$$

are each equivalent to CCQ for p.

b) (PCQ) The PCQ holds for p if and only if

$$\operatorname{pos} \mathcal{C}(A, B) + \operatorname{bar} \mathcal{V} = \operatorname{span}(\mathcal{C}(A, B) + \operatorname{bar} \mathcal{V}). \tag{5.4}$$

c) (BPCQ) The qualification conditions

$$\mathcal{V} \cap \mathcal{K}_A \neq \emptyset \quad \text{and} \quad \mathcal{V}^\infty \cap \mathcal{K}_A = \{0\}, \tag{5.5}$$

$$\mathcal{V} \cap \mathcal{K}_A \neq \emptyset \text{ is bounded}, \tag{5.6}$$

$$\mathcal{V} \cap \mathcal{K}_A \neq \emptyset \quad \text{and} \quad \operatorname{bar} \mathcal{V} + \mathcal{K}_A^\circ = \mathbb{S}^n \tag{5.7}$$

are each equivalent to BPCQ for p, hence imply (5.4).

Proof. a) First, observe that , with $h = \delta_{\mathcal{V}}$, condition i) in Lemma 3.12 is exactly (5.2). By the same lemma this is equivalent to

$$\operatorname{hzn}\sigma_{\mathcal{V}} \cap (-\mathcal{K}_A^\circ) = \{0\}.$$

Moreover, as $\sigma_{\mathcal{V}} = \sigma_{\mathcal{V}}^\infty$, we have

$$\operatorname{hzn}\sigma_{\mathcal{V}} = \{V \mid \sigma_{\mathcal{V}}(V) \le 0\} = \mathcal{V}^-.$$

Invoking [3, Section 3.3, Exercise 16 (a)] implies that

$$\operatorname{hzn}\sigma_{\mathcal{V}} \cap (-\mathcal{K}_A^\circ) = \{0\} \iff \operatorname{cl}(\overline{\operatorname{cone}}\,\mathcal{V} - \mathcal{K}_A) = \mathbb{S}^n,$$

where the closure in the latter statement can clearly be dropped, e.g. by interpreting [15, Theorem 6.3] accordingly.

b) Use (2.1) to infer that PCQ holds for p if and only if

$$\operatorname{pos}(\mathcal{C}(A,B)) + \operatorname{bar} V = \operatorname{pos}(\mathcal{C}(A,B) + \bar{V}) = \operatorname{span}(\mathcal{C}(A,B) + \operatorname{bar}\mathcal{V}).$$

c) The equivalences of BPCQ, (5.5), and (5.6) are clear. Since $\mathcal{V}^\infty$ and $\operatorname{cl}(\operatorname{bar}\mathcal{V})$ are paired in polarity, see (2.6), [3, Section 3.3, Exercise 16 (a)] implies that

$$\mathcal{V}^\infty \cap \mathcal{K}_A = \{0\} \iff \operatorname{cl}(\operatorname{bar}\mathcal{V} + \mathcal{K}_A^\circ) = \mathbb{S}^n,$$

where the closure in the latter statement can be dropped as in a). This establishes all equivalences. ∎

The following result provides sufficient conditions for the occurence of $p = p^{**}$ when p is given as in (5.1), i.e. in the case that h is an indicator function.

COROLLARY 5.2. *Let p be given by (5.1). Then $p \in \Gamma_0(\mathbb{R}^{n\times m})$ (i.e. $p = p^{**}$) under any of the conditions in (5.2)-(5.7). Under condition (5.2)-(5.3) it is also finite-valued.*

Proof. Follows from Lemma 5.1 and Theorem 3.5 c) and Theorem 3.14, respectively. ∎

We treat the case $A = 0$ and $B = 0$ separately as we will use it in Section 5.2.

COROLLARY 5.3. *Let p be given as in (5.1) and assume that $A = 0$ and $B = 0$ and such that $\mathcal{V} \cap \mathbb{S}^n_+$ is nonempty. Then we have*

$$PCQ \iff \mathbb{S}^n_+ + \operatorname{bar}\mathcal{V} = \mathbb{S}^n \iff BPCQ.$$

Moreover, $p \in \Gamma_(\mathbb{R}^{n\times m})$, i.e. $p = p^{**}$ under any of following conditions:

i) $\mathcal{V} \cap \mathbb{S}^n_{++} \neq \emptyset$ (CCQ);

ii) $\mathcal{V} \cap \mathbb{S}^n_+ \neq \emptyset$ is bounded (or equivalently $\mathbb{S}^n_+ + \operatorname{bar}\mathcal{V} = \mathbb{S}^n$) ((B/S)PCQ).

Under condition i) p is also finite-valued.

Proof. For the first statement notice that $\mathcal{C}(0,0) = \mathbb{S}^n_- = \mathcal{K}_0^\circ$ and invoke Lemma 5.1. The rest follows from Corollary 5.2 and Lemma 5.1. ∎

To compute the conjugate p^*, instead of using Theorem 3.5, a direct derivation relying on [5, Theorem 3.2] yields a powerful result.

THEOREM 5.4 (Infimal projection with an indicator function). *Let p be given by (5.1). Then its conjugate $p^*:\mathbb{R}^{n\times m}\to\overline{\mathbb{R}}$ is given by*

$$p^*(Y)=\frac{1}{2}\sigma_{\mathcal{V}\cap\mathcal{K}_A}\left(YY^T\right)+\delta_{\{Y\mid AY=B\}}(Y).$$

In particular, for $A=0$ and $B=0$ we obtain

$$p^*(Y)=\frac{1}{2}\sigma_{\mathcal{V}\cap\mathbb{S}^n_+}\left(YY^T\right).$$

Proof. By (2.7), we have

$$\begin{aligned}
p^*(Y)&=\sup_X\left[\langle X,Y\rangle-\inf_V\sigma_{\mathcal{D}(A,B)}(X,V)+\delta_{\mathcal{V}}(V)\right]\\
&=\sup_V\sup_X\left[\langle X,Y\rangle-\sigma_{\mathcal{D}(A,B)}(X,V)-\delta_{\mathcal{V}}(V)\right]\\
&=\sup_{V\in\mathcal{V}\cap\mathcal{K}_A}\ \sup_{\operatorname{rge}\binom{X}{B}\subset\operatorname{rge}M(V)}\operatorname{tr}\left(-\frac{1}{2}\begin{pmatrix}X\\B\end{pmatrix}^T M(V)^\dagger\begin{pmatrix}X\\B\end{pmatrix}+Y^TX\right)
\end{aligned}$$

for $Y\in\mathbb{R}^{n\times m}$. Since $\operatorname{rge}\binom{X}{B}\subset\operatorname{rge}M(V)$, we can make the substitution $M(V)\binom{U}{W}=\binom{X}{B}$, to obtain

$$\begin{aligned}
p^*(Y)&=\sup_{V\in\mathcal{V}\cap\mathcal{K}_A}\ \sup_{\substack{U,W\\AU=B}}\operatorname{tr}\left(-\frac{1}{2}\begin{pmatrix}U\\W\end{pmatrix}^T M(V)\begin{pmatrix}U\\W\end{pmatrix}+Y^T(VU+A^TW)\right)\\
&=\sup_{V\in\mathcal{V}\cap\mathcal{K}_A}-\sum_{i=1}^m\inf_{\substack{u_i,w_i\\Au_i=b_i}}\left(\frac{1}{2}\begin{pmatrix}u_i\\w_i\end{pmatrix}^T M(V)\begin{pmatrix}u_i\\w_i\end{pmatrix}-y_i^TVu_i-w_i^TAy_i\right)\\
&=\sup_{V\in\mathcal{V}\cap\mathcal{K}_A}-\sum_{i=1}^m\inf_{\substack{u_i,w_i\\Au_i=b_i}}\left(\frac{1}{2}u_i^TVu_i-\langle Vy_i,u_i\rangle+\langle w_i,b_i-Ay_i\rangle\right)\\
&=\sup_{V\in\mathcal{V}\cap\mathcal{K}_A}-\sum_{i=1}^m\left[\inf_{Au_i=b_i}\left(\frac{1}{2}u_i^TVu_i-\langle Vy_i,u_i\rangle\right)+\inf_{w_i}\left(\langle w_i,b_i-Ay_i\rangle\right)\right]\\
&=\delta_{\{Z\mid AZ=B\}}(Y)+\sup_{V\in\mathcal{V}\cap\mathcal{K}_A}-\sum_{i=1}^m\inf_{Au_i=b_i}\left(\frac{1}{2}u_i^TVu_i-\langle Vy_i,u_i\rangle\right),
\end{aligned}$$

where the final equality follows since $\delta_{\{y\mid b_i-Ay_i\}}(y_i)=\sup_{w_i}\langle w_i,b_i-Ay_i\rangle$ $(i=1,\dots,m)$. By hypothesis $\operatorname{rge}B\subset\operatorname{rge}A$, and so, by [5, Theorem 3.2]

$$-\frac{1}{2}\begin{pmatrix}Vy_i\\b_i\end{pmatrix}^T M(V)^\dagger\begin{pmatrix}Vy_i\\b_i\end{pmatrix}=\inf_{Au_i=b_i}\left(\frac{1}{2}u_i^TVu_i-\langle Vy_i,u_i\rangle\right)\quad(i=1,\dots,m),$$

Therefore, when $AY = B$, we have

$$\begin{aligned} p^*(Y) &= \sup_{V\in\mathcal{V}\cap\mathcal{K}_A} -\sum_{i=1}^m -\frac{1}{2}\begin{pmatrix} Vy_i \\ b_i\end{pmatrix}^T M(V)^\dagger \begin{pmatrix} Vy_i \\ b_i\end{pmatrix} \qquad \left(\begin{array}{c}\text{where } Ay_i = b_i \text{ so} \\ \binom{Vy_i}{b_i} = M(V)\binom{y_i}{0}\end{array}\right) \\ &= \sup_{V\in\mathcal{V}\cap\mathcal{K}_A} \frac{1}{2}\sum_{i=1}^m \left(M(V)\begin{pmatrix} y_i \\ 0\end{pmatrix}\right)^T M(V)^\dagger \left(M(V)\begin{pmatrix} y_i \\ 0\end{pmatrix}\right)^T \\ &= \sup_{V\in\mathcal{V}\cap\mathcal{K}_A} \frac{1}{2}\sum_{i=1}^m \begin{pmatrix} y_i \\ 0\end{pmatrix}^T M(V)\begin{pmatrix} y_i \\ 0\end{pmatrix}^T \\ &= \sup_{V\in\mathcal{V}\cap\mathcal{K}_A} \frac{1}{2}\sum_{i=1}^m y_i^T V y_i \\ &= \sup_{V\in\mathcal{V}\cap\mathcal{K}_A} \frac{1}{2}\mathrm{tr}\,(Y^T V Y), \end{aligned}$$

which proves the general expression for p^*. The case $A = 0, B = 0$ follows readily. ◻

We now study the subdifferential of p given by (5.1).

COROLLARY 5.5. *Let p be given by (5.1). If $\mathcal{V}\cap \mathrm{int}\,\mathcal{K}_A \neq \emptyset$ (CCQ) then*

$$\partial p(\bar{x}) = \underset{Y}{\mathrm{argmax}}\{\langle \bar{X},\, Y\rangle - \inf_{(Y,T)\in\Omega(A,B)} \sigma_{\mathcal{V}}(-T)\}$$

is nonempty and compact for all $\bar{X} \in \mathrm{dom}\,p$. If, in addition, $\mathrm{pos}\,\mathcal{C}(A,B) + \mathrm{bar}\,\mathcal{V} = \mathrm{span}\,(\mathcal{C}(A,B) + \mathrm{bar}\,\mathcal{V})$ (PCQ), then

$$\partial p(\bar{X}) = \left\{\bar{Y} \;\middle|\; \exists \bar{V},\bar{T}:\; -\bar{T} \in N_{\mathcal{V}}(\bar{V}),\; (\bar{Y},\bar{T}) \in \partial\sigma_{\Omega(A,B)}(\bar{X},\bar{V})\right\}$$

is nonempty and compact for all $\bar{X} \in \mathbb{R}^{n\times m}$.

Proof. Follows readily from Proposition 3.16 in combination with Lemma 5.1. ◻

5.1. $B = 0$ and $0 \in \mathcal{V}$.. We now consider the important special case of p given by (5.1) where $0 \in \mathcal{V}$ and $B = 0$. In this case p turns out to be a squared gauge function, see Corollary 5.9. We start with a technical lemma.

LEMMA 5.6. *Let $C, K \subset \mathbb{E}$ be nonempty, convex with K being a cone. Then $(C + K)^\circ = C^\circ \cap K^\circ$. If $C + K$ is closed with $0 \in C$, then $(C^\circ \cap K^\circ)^\circ = C + K$. In particular, the set $C + K$ is closed if C and K are closed and $K \cap (-C^\infty) = \{0\}$.*

Proof. Clearly, $C^\circ \cap K^\circ \subset (C + K)^\circ$. Conversely, if $z \in (C + K)^\circ$, then $\langle z,\, x + ty\rangle \leq 1$ for all $x \in C$, $y \in K$, and $t > 0$. Multiplying this inequality by t^{-1} and letting $t \to \infty$, we see that $z \in K^\circ$. By letting $t \downarrow 0$, we see that $z \in C^\circ$.

Now assume that $C + K$ is closed with $0 \in C$. Then $C + K$ is closed and convex with $0 \in C + K$. Hence, by [15, Theorem 14.5], $C + K = (C + K)^{\circ\circ} = (C^\circ \cap K^\circ)^\circ$.

The final statement of the lemma follows from [15, Corollary 9.1.1]. ◻

The first main result in this section is concerned with a representation of the conjugate p^* under the standing assumptions.

COROLLARY 5.7 (The gauge case I). *Let p be given by (5.1) with $0 \in \mathcal{V}$ and $B = 0$ and let P be the orthogonal projection onto* $\ker A$. *Moreover, let*

$$\mathcal{S} := \{W \in \mathbb{S}^n \mid \mathrm{rge}\,W \subset \ker A\} = \{W \in \mathbb{S}^n \mid W = PWP\}\,.^4$$

Then the following hold:

[4] Here we consider $\mathcal{S} = \mathbb{S}^n \cap \mathrm{Ker}_n A$ as a subset in the space $\mathbb{S}^n$.

a) We have

$$p^*(Y) = \frac{1}{2}\sigma_{(\mathcal{V}\cap\mathcal{K}_A)+\mathcal{S}^\perp}\left(YY^T\right) = \frac{1}{2}\gamma_{(\mathcal{V}\cap\mathcal{K}_A)^\circ\cap\mathcal{S}}\left(YY^T\right)$$

where $\mathcal{S}^\perp = \{V \in \mathbb{S}^n \mid PVP = 0\}$. *In particular,* p^* *is positively homogeneous of degree 2.*

b) If $\mathcal{V}^\circ + \mathcal{K}_A^\circ$ *is closed (e.g. when* $\mathcal{K}_A^\circ \cap -(\mathrm{cone}\,\mathcal{V})^\circ = \{0\}$*) then*

$$p^*(Y) = \frac{1}{2}\gamma_{(\mathcal{V}^\circ\cap\mathcal{S})+\mathcal{K}_A^\circ}\left(YY^T\right), \tag{5.8}$$

where $\mathrm{dom}\, p^* = \left\{Y \;\middle|\; YY^T \in \mathrm{cone}\,\mathcal{V}^\circ \cap \mathcal{S} + \mathcal{K}_A^\circ\right\}$.

Proof. a) We have

$$\begin{aligned} p^*(Y) =& \frac{1}{2}\sigma_{\mathcal{V}\cap\mathcal{K}_A}\left(YY^T\right) + \delta_{\{Y \mid AY=0\}} \\ =& \frac{1}{2}\sigma_{\mathcal{V}\cap\mathcal{K}_A}\left(YY^T\right) + \frac{1}{2}\delta_{\mathcal{S}}\left(YY^T\right) \\ =& \frac{1}{2}\sigma_{\mathcal{V}\cap\mathcal{K}_A}\left(YY^T\right) + \frac{1}{2}\sigma_{\mathcal{S}^\perp}\left(YY^T\right) \\ =& \frac{1}{2}\sigma_{(\mathcal{V}\cap\mathcal{K}_A)+\mathcal{S}^\perp}\left(YY^T\right) \\ =& \frac{1}{2}\gamma_{(\mathcal{V}\cap\mathcal{K}_A)^\circ\cap\mathcal{S}}\left(YY^T\right). \end{aligned}$$

Here the first equality uses Theorem 5.4, the second equality follows from the fact that $\mathrm{rge}\, Y = \mathrm{rge}\, YY^T$, the third can be seen from [16, Example 7.4], the fourth uses (2.5), and the final equivalence follows from [15, Theorem 14.5] and Lemma 5.6.

b) If $\mathcal{V}^\circ + \mathcal{K}_A^\circ$ is closed, then Lemma 5.6 also tells us that $(\mathcal{V}\cap\mathcal{K}_A)^\circ = \mathcal{V}^\circ + \mathcal{K}_A^\circ$. Since $\mathcal{K}_A^\circ \subset \mathcal{S}$, see Lemma 2.1 b), we have

$$(\mathcal{V}^\circ + \mathcal{K}_A^\circ) \cap \mathcal{S} = (\mathcal{V}^\circ \cap \mathcal{S}) + \mathcal{K}_A^\circ$$

which, using a), gives the first equivalence in (5.8). ∎

Our final goal is to show that p, under the standing assumption in this section, is a squared gauge. To this end, the next result is key.

LEMMA 5.8. *Let* $0 \in C \subset \mathcal{E}$ *be closed and convex and define* $q : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ *through* $q(x) := \frac{1}{2}\gamma_C^2(x)$. *Then* $q^* = \frac{1}{2}\gamma_{C^\circ}^2$.

Proof. Apply [16, Proposition 11.21] with $\theta = \frac{1}{2}(\cdot)^2$. ∎

We are now in a position to prove the last result of this section announced earlier. Here we denote by $\mathbb{B}_F$ the (closed) unit ball in the Frobenius norm.

COROLLARY 5.9 (The gauge case II). *Let* p *be as in Theorem 5.4 with* $0 \in \mathcal{V}$ *and* $B = 0$. *For* $P \in \mathbb{R}^{n\times n}$ *the orthogonal projector on* $\ker A$, *define the (closed, convex) sets*

$$\mathcal{V}_A^{1/2} := \left\{L \in \mathbb{R}^{n\times n} \;\middle|\; LL^T \in P(\mathcal{V}\cap\mathcal{K}_A)P\right\}, \quad \mathcal{F} := \left\{LZ \;\middle|\; L \in \mathcal{V}_A^{1/2},\ Z \in \mathbb{B}_F\right\},$$

and the subspace $\mathcal{U} := \mathrm{Ker}_m A$.[5] *Then*

$$p = \frac{1}{2}\gamma^2_{\mathcal{F}+\mathcal{U}^\perp} \quad \textit{and} \quad p^* = \frac{1}{2}\gamma^2_{\mathcal{F}^\circ\cap\mathcal{U}}.$$

[5] Hence $\mathcal{U}^\perp = \mathrm{Rge}_m A^T$.

In particular, for $A = 0$ and $\mathcal{F} := \left\{ LZ \;\middle|\; LL^T \in \mathcal{V} \cap \mathbb{S}^n_+,\ Z \in \mathbb{B}_F \right\}$ we obtain

$$p = \frac{1}{2}\gamma^2_{\mathcal{F}} \quad \text{and} \quad p^* = \gamma^2_{\mathcal{F}^\circ}.$$

Proof. For all $Y \in \mathbb{R}^{n\times m}$, by Theorem 5.4 and the definition of $\mathcal{U}$, we have

$$p^*(Y) = \frac{1}{2}\sigma_{\mathcal{V}\cap\mathcal{K}_A}(YY^T) + \delta_{\mathcal{U}}(Y) = \frac{1}{2}\sup_{V\in\mathcal{V}\cap\mathcal{K}_A}\left\langle PVP,\, YY^T\right\rangle + \delta_{\mathcal{U}}(Y).$$

In turn, by the definitions of $\mathcal{V}_A^{1/2}$ and the Frobenius norm, the latter equals

$$\frac{1}{2}\sup_{L\in\mathcal{V}_A^{1/2}}\left\langle LL^T,\, YY^T\right\rangle + \delta_{\mathcal{U}}(Y) = \frac{1}{2}\sup_{L\in\mathcal{V}_A^{1/2}}\|L^TY\|_F^2 + \delta_{\mathcal{U}}(Y).$$

On the other hand, by the monotonicity and continuity of $t \in \mathbb{R}_+ \mapsto t^2$ as well as the self-duality of the Frobenius norm, we find that the latter can be written as

$$\frac{1}{2}\left[\sup_{L\in\mathcal{V}_A^{1/2}}\|L^TY\|_F\right]^2 + \delta_{\mathcal{U}}(Y) = \frac{1}{2}\left[\sup_{(Z,L)\in\mathbb{B}_F\times\mathcal{V}_A^{1/2}}\left\langle L^TY,\, Z\right\rangle\right]^2 + \delta_{\mathcal{U}}(Y).$$

This, however, using the definition of $\mathcal{F}$ and the convention $(+\infty)^2 = +\infty$, we can rewrite as

$$\frac{1}{2}\sigma_{\mathcal{F}}(Y)^2 + \delta_{\mathcal{U}}(Y) = \frac{1}{2}\left[\sigma_{\mathcal{F}}(Y) + \delta_{\mathcal{U}}(Y)\right]^2.$$

All in all, using the latter, [16, Example 11.4], (2.5), and [16, Example 11.19] and the polar cone calculus from, e.g., [3, p. 70], we conclude that

$$p^*(Y) = \frac{1}{2}\left[\sigma_{\mathcal{F}}(Y) + \delta_{\mathcal{U}}(Y)\right]^2 = \frac{1}{2}\left[\sigma_{\mathcal{F}}(Y) + \sigma_{\mathcal{U}^\perp}(Y)\right]^2 = \frac{1}{2}\sigma^2_{\mathcal{F}+\mathcal{U}^\perp}(Y) = \frac{1}{2}\gamma^2_{\mathcal{F}^\circ\cap\mathcal{U}}(Y).$$

∎

This proves the representation for p^*; the one for p then follows from Lemma 5.8.

5.2. Variational Gram Functions. Given a closed, convex set $\mathcal{V} \subset \mathbb{S}^n$ we define

$$\Omega_{\mathcal{V}} : \mathbb{R}^{n\times m} \to \overline{\mathbb{R}}, \quad \Omega_{\mathcal{V}}(Y) := \frac{1}{2}\sigma_{\mathcal{V}\cap\mathbb{S}^n_+}(YY^T). \tag{5.9}$$

These kinds of functions are called *variational Gram function (VGF)* and have received some attention lately in the machine learning community due to their orthogonality promoting properties when used as penalty functions, cf. [14].

Note that our definition explicitly intersects $\mathcal{V}$ with the positive semidefinite cone $\mathbb{S}^n_+$ while in the analysis in [14] a standing assumption is that $\Omega_{\mathcal{V}} = \Omega_{\mathcal{V}\cap\mathbb{S}^n}$. These (equivalent) conventions guarantee that $\Omega_{\mathcal{V}}$ is convex. We also scale by $\frac{1}{2}$ to have more elegant formulas.

Our first result follows readily from our above analysis and refines [14, Proposition 4] about the conjugate of a VGF.

Proposition 5.10 (Conjugate of VGFs and VGFs as Squared Gauges). *Let $\Omega_{\mathcal{V}}$ be given by (5.9). Under either of the following assumptions*

i) $\mathcal{V} \cap \mathbb{S}^n_{++} \neq \emptyset$,

ii) $\mathcal{V}\cap\mathbb{S}^n_+\neq\emptyset$ *bounded (or equivalently* $\mathbb{S}^n_+ + \mathrm{bar}\,\mathcal{V} = \mathbb{S}^n$*),*
we have

$$\Omega^*_{\mathcal{V}}(X) = \inf_V \sigma_\Omega(X,V) + \delta_{\mathcal{V}}(V) = \frac{1}{2}\inf_{\substack{V\in\mathcal{V}\cap\mathbb{S}^n_+:\\ \mathrm{rge}\,X\subset\mathrm{rge}\,V}} \mathrm{tr}\,\left(X^TV^\dagger X\right) \quad (X\in\mathbb{R}^{n\times m}).$$

Under i), $\Omega^*_{\mathcal{V}}$ *is finite-valued, and under ii),* $\Omega_{\mathcal{V}}$ *is finite-valued. In addition, if* $0\in\mathcal{V}$ *we also have*

$$\Omega_{\mathcal{V}} = \frac{1}{2}\gamma^2_{\mathcal{F}^\circ} \quad \textit{and} \quad \Omega^*_{\mathcal{V}} = \frac{1}{2}\gamma^2_{\mathcal{F}}$$

with $\mathcal{F} = \left\{LZ \;\middle|\; LL^T\in\mathcal{V}\cap\mathbb{S}^n_+,\ Z\in\mathbb{B}_F\right\}$.

Proof. Using Theorem 5.4, Corollary 5.3 and the function p occuring there, we have $\Omega^*_{\mathcal{V}} = p^{**} = p$. The rest is clear from the definition of p and the matrix-fractional function as well as the respective results from Section 5, in particular Corollary 5.9 for the last statement. ∎

Next we are interested in the subdifferential of a VGF in the sense of (5.9). Although, by our definition, a VGF is always convex, we take the *convex-composite* perspective, see e.g. [7], since essentially a VGF is simply the composition of a closed, proper, convex function $\sigma_{\mathcal{V}\cap\mathbb{S}^n_+}$ and a nonlinear map $H: Y\mapsto YY^T$. It turns out, that the *basic constraint qualification* for $\Omega_{\mathcal{V}} = \frac{1}{2}\sigma_{\mathcal{V}\cap\mathbb{S}^n_+}\circ H$, which reads

$$N_{\mathrm{dom}\,\sigma_{\mathcal{V}\cap\mathbb{S}^n_+}}(\bar{Y}\bar{Y}^T)\cap(\mathrm{Ker}_n\bar{Y}^T) = \{0\} \quad (\bar{Y}\in\mathrm{dom}\,\Omega_{\mathcal{V}}), \tag{5.10}$$

and which is essential for full subdifferential calculus of convex-composites, is intimately linked with condition ii) in Corollary 5.3.

LEMMA 5.11 (BCQ for VGF). *Let* $\Omega_{\mathcal{V}}$ *be given by* (5.9) *and assume that* $\mathbb{S}^n_+\cap\mathcal{V}\neq\emptyset$*. Then the following are equivalent:*

i) There exists $\bar{Y}\in\mathrm{dom}\,\Omega_{\mathcal{V}}$ *such that* (5.10) *holds;*

ii) $\mathcal{V}^\infty\cap\mathbb{S}^n_+ = \{0\}$ *(or equivalently* $\mathcal{V}\cap\mathbb{S}^n_+$ *is bounded);*

iii) (5.10) *holds at every* $\bar{Y}\in\mathrm{dom}\,\Omega_{\mathcal{V}}$*.*

Proof. 'i)⇒ii)': Assume ii) were violated, i.e. there exists $0\neq W\in(\mathcal{V}\cap\mathbb{S}^n_+)^\infty = \mathcal{V}^\infty\cap\mathbb{S}^n_+$. Moreover, by assumption there exists $\bar{V}\in\mathbb{S}^n_+\cap\mathcal{V}$. By the properties of the horizon cone of closed, convex sets, see (2.2), we have

$$V_t := \bar{V} + tW \in \mathcal{V}\cap\mathbb{S}^n_+ \quad (t>0). \tag{5.11}$$

Now, take any $\bar{Y}\in\mathrm{dom}\,\Omega_{\mathcal{V}}$. Then, for all $t>0$, we have

$$\begin{aligned} +\infty &> \Omega_{\mathcal{V}}(\bar{Y}) \\ &= \sup_{V\in\mathbb{S}^n_+\cap\mathcal{V}} \left\langle V,\, \bar{Y}\bar{Y}^T\right\rangle \\ &\geq \left\langle V_t,\, \bar{Y}\bar{Y}^T\right\rangle \\ &\geq t\left\langle W,\, \bar{Y}\bar{Y}^T\right\rangle. \end{aligned}$$

Since $W\succeq 0$, we have $\left\langle \bar{Y}\bar{Y}^T,\, W\right\rangle = \mathrm{tr}\,(\bar{Y}^TW\bar{Y})\geq 0$. In view of the above chain of inequalities this implies $\left\langle W,\, \bar{Y}\bar{Y}^T\right\rangle = 0$ and as $W,\ \bar{Y}\bar{Y}^T\succeq 0$ this gives $W\bar{Y}\bar{Y}^T = 0$. Since $\mathrm{rge}\,\bar{Y} = \mathrm{rge}\,\bar{Y}\bar{Y}^T$ this implies $W\bar{Y} = 0$ or, equivalently, $\bar{Y}^TW = 0$. Therefore, we have $0\neq W\in(\mathcal{V}\cap\mathbb{S}^n_+)^\infty\cap(\mathrm{Ker}_n\bar{Y}^T)$. Now, observe that $N_{\mathrm{dom}\,\sigma_{\mathcal{V}\cap\mathbb{S}^n_+}}(Z) =$

$(\mathcal{V}\cap\mathbb{S}^n_+)^\infty$ for any $Z\in\operatorname{dom}\sigma_{\mathcal{V}\cap\mathbb{S}^n_+}$, see e.g. [16]. This shows that (5.10) is violated at $\bar{Y}$. Since $\bar{Y}\in\operatorname{dom}\Omega_{\mathcal{V}}$ was chosen arbitrarily, this establishes the desired implication.

'ii)⇒iii)': If $\mathcal{V}\cap\mathbb{S}^n_+$ is bounded, then $\operatorname{dom}\sigma_{\mathcal{V}\cap\mathbb{S}^n_+}=\mathbb{S}^n$, and hence $N_{\operatorname{dom}\sigma_{\mathcal{V}\cap\mathbb{S}^n_+}}(\bar{Y}\bar{Y}^T)=\mathbb{S}^n$ for every $\bar{Y}\in\operatorname{dom}\Omega_{\mathcal{V}}$, which gives the desired implication.

'iii)⇒i)': Obvious. ∎

We now derive the formula for the subdifferential of the VGF from (5.9).

Proposition 5.12. *Let $\Omega_{\mathcal{V}}$ be given by (5.9). Then*

$$\partial\Omega_{\mathcal{V}}(\bar{Y})\supset\left\{\bar{V}\bar{Y}\;\middle|\;\bar{V}\in\mathcal{V}\cap\mathbb{S}^n_+:\;\left\langle\bar{V},\,\bar{Y}\bar{Y}^T\right\rangle=\Omega_{\mathcal{V}}(\bar{Y})\right\}\quad(\bar{Y}\in\operatorname{dom}\Omega_{\mathcal{V}}).$$

If $\mathbb{S}^n_+\cap\mathcal{V}$ is nonempty and bounded, equality holds and $\operatorname{dom}\Omega_{\mathcal{V}}=\mathbb{R}^{n\times m}$.

Proof. Combine Lemma 5.11 with [16, Theorem 10.6], [16, Corollary 8.25] and the fact that for $H:Y\to YY^T$ we have $\nabla H(Y)^*V=2VY$ for all $(Y,V)\in\mathbb{E}$. ∎

We next consider an example.

Example 5.13 (Failure of subdifferential calculus for VGF). *Let $\mathcal{V}:=\operatorname{pos}\{I\}\subset\mathbb{S}^n$, put $m:=1$ and let $H:Y\mapsto YY^T$. Then clearly condition i) in Proposition 5.10 holds, but condition ii) and hence the BCQ (5.10) fails. We have*

$$\sigma_{\mathcal{V}\cap\mathbb{S}^n_+}(W)=\sup_{\alpha\geq 0}\alpha\operatorname{tr}(W)=\delta_{\{U\in\mathbb{S}^n\,|\,\operatorname{tr}(U)\leq 0\}}(W)\quad(W\in\mathbb{S}^n).\tag{5.12}$$

Hence, we obtain $\operatorname{dom}\Omega_{\mathcal{V}}=\{0\}$ and $\nabla H(0)^*\partial\sigma_{\mathcal{V}\cap\mathbb{S}^n_+}(0)=\{0\}$. On the other hand, we have $\Omega_{\mathcal{V}}=\frac{1}{2}\sigma_{\mathcal{V}\cap\mathbb{S}^n_+}\circ H=\delta_{\{0\}}$. Therefore, we have

$$\partial\Omega_{\mathcal{V}}(0)=N_{\{0\}}(0)=\mathbb{R}^{n\times m}\supsetneq\{0\}=\nabla H(0)^*\partial\sigma_{\mathcal{V}\cap\mathbb{S}^n_+}(0).$$

Example 5.13 establishes various things: First, it shows that condition i) in Proposition 5.10 does not yield equality in the subdifferential formula for VGFs. It also illustrates that equality in the subdifferential formula may fail tremendously in the absence of BCQ, even for a convex-composite which is, in fact, convex.

Much effort is made in [14] to compute the conjugate of a (convex) VGF, cf. [14, Proposition 7] and its proof. A slightly refined version of the latter result follows readily from our analysis.

Proposition 5.14 (Subdifferential of $\Omega^*_{\mathcal{V}}$). *Let $\Omega_{\mathcal{V}}$ be given by (5.9) and assume that that $0\in\operatorname{ri}(\mathcal{C}+\operatorname{bar}\mathcal{V})$. Under either of the following assumptions*

i) $\mathcal{V}\cap\mathbb{S}^n_{++}\neq\emptyset$,

ii) $\mathcal{V}\cap\mathbb{S}^n_+\neq\emptyset$ and $\mathcal{V}^\infty\cap\mathbb{S}^n_+=\{0\}$ (or equivalently $\mathcal{V}\cap\mathbb{S}^n_+\neq\emptyset$ bounded),

for any $\bar{X}\in\mathbb{R}^{n\times m}$ where $\bar{X}$ is finite we have

$$\partial\Omega^*_{\mathcal{V}}(\bar{X})=\left\{\bar{Y}\;\middle|\;\begin{array}{l}\exists\bar{V}\in\mathcal{V}\cap\mathbb{S}^n_+:\operatorname{rge}\bar{X}\subset\operatorname{rge}\bar{V},\\ \Omega^*_{\mathcal{V}}(\bar{X})=\dfrac{1}{2}\operatorname{tr}\left(\bar{X}^T\bar{V}^\dagger\bar{X}\right)=\left\langle\bar{X},\,\bar{Y}\right\rangle-\Omega_{\mathcal{V}}(\bar{Y})\end{array}\right\}$$

with $\operatorname{dom}\partial\Omega^*_{\mathcal{V}}=\operatorname{dom}\Omega^*_{\mathcal{V}}$.

Proof. Using Corollary 5.3 and the function p occuring there, we have $\Omega^*_{\mathcal{V}}=p$ and $\Omega_{\mathcal{V}}=p^*$ under either i) or ii). The subdifferential formula follows then from Proposition 3.16 (see in particular the third identity in c)).

The fact that $\operatorname{dom}\partial\Omega^*_{\mathcal{V}}=\operatorname{dom}\Omega^*_{\mathcal{V}}$ is due to the fact that the latter is a subspace, hence relatively open, cf. Lemma 3.1 c). ∎

5.3. VGFs and squared Ky Fan norms. For $p \geq 1$, $1 \leq k \leq \min\{m, n\}$, the *Ky Fan (p,k)-norm* [12, Ex. 3.4.3] of a matrix $X \in \mathbb{R}^{n\times m}$ is defined as

$$\|X\|_{p,k} = \left(\sum_{i=1}^{k} \sigma_i^p\right)^{1/p},$$

where σ_i are the singular values of X sorted in nonincreasing order. In particular, the $(p, \min\{m, n\})$-norm is the Schatten-p norm and the $(1, k)$-norm is the standard Ky Fan k-norm, see [12]. For $1 \leq p \leq \infty$, denote the closed unit ball for $\|\cdot\|_{p,k}$ by $\mathbb{B}_{p,k} := \{X \mid \|X\|_{p,k} \leq 1\}$. For $1 \leq p \leq \infty$, define $s := p/2$. Then, for $2 \leq p \leq \infty$, we have

$$\begin{aligned}\frac{1}{2}\|X\|_{p,k}^2 &= \frac{1}{2}\left[\sum_{i=1}^{k}(\sigma_i^2)^s\right]^{1/s}\\ &= \frac{1}{2}\|XX^T\|_{s,k} = \frac{1}{2}\sigma_{\mathbb{B}^\circ_{s,k}}(XX^T) = \frac{1}{2}\sigma_{\mathbb{B}^\circ_{s,k}\cap\mathbb{S}^n_+}(XX^T)\\ &= \frac{1}{2}\Omega_{\mathbb{B}^\circ_{s,k}}(X),\end{aligned}$$

where the first equality follows from the definition of s, the second from the definition of the singular values, the third from properties of gauges and their polars, the fourth from the equivalence $\langle V, XX^T\rangle = \sum_{j=1}^m x_j^T V x_j$ with the x_j's the columns of X, and the final from (5.9). For the Schatten norms, where $k = \min\{n, m\}$ we have $\mathbb{B}^\circ_{s,k} = \mathbb{B}_{\hat{s},k}$, where $\hat{s}$ satisfies $\frac{1}{s} + \frac{1}{\hat{s}} = 1$, see [11]. For other values of k, the representation of $\mathbb{B}^\circ_{s,k}$ can be significantly more complicated, e.g. see [8].

6. Final remarks. In this paper we studied partial infimal projections of the generalized matrix-fractional function with a closed, proper, convex function $h : \mathbb{S}^n \to \overline{\mathbb{R}}$. Sufficient conditions for closedness and properness as well as representations of both the conjugate and the subdifferential of the infimal projections are given, along with the essential constraint qualifications. Particular emphasis was given in the instances where the function h is a support or an indicator function of a closed, convex set in $\mathbb{S}^n$. As a special case of support functions, infimal projections with suitable linear functionals yielded smoothing variational representations for the family of scaled nuclear norms. In the indicator case, it was shown that, under appropriate assumptions, the infimal projection is positively homogeneous of degree two, in fact, a squared gauge. Moreover, in a special case, it was proven that the conjugate of the infimal projection coincides with a variational Gram function (VGF) of the underlying set. Thus we were able to easily establish a variational calculus for VGFs as a consequence of our more general analysis. In addition, we made a connection with Ky Fan norms.

7. Appendix. In what follows we use the *direct sum* of functions $f_i \in \mathcal{E}$ which is defined by

$$\oplus_{i=1}^m f_i : \mathcal{E}^m \to \mathbb{R}\cup\{+\infty\}, \quad \oplus_{i=1}^m f_i(x_1, \dots, x_m) = \sum_{i=1}^{m} f_i(x_i).$$

THEOREM 7.1 (Extended sum rule). *Let $f_i \in \Gamma_0(\mathcal{E})$ $(i = 1, \dots, m)$ and set $f := \sum_{i=1}^m f_i$. Then the following hold:*

a) (Attouch-Brézis) It holds that $f^* = \mathrm{cl}\,(f_1^* \,\Box\, f_2^* \,\Box \cdots \Box\, f_m^*)$*. Under the qualification condition*

$$\bigcap_{i=1}^{m} \mathrm{ri}\,(\mathrm{dom}\, f_i) \neq \emptyset \tag{7.1}$$

we have $f^* = f_1^* \,\Box\, f_2^* \,\Box \cdots \Box\, f_m^*$ *which is closed, proper and convex and*

$$\emptyset \neq \mathcal{T}(z) := \mathrm{argmin}\left\{ \sum_{i=1}^{m} f_i^*(z^i) \;\middle|\; z = \sum_{i=1}^{m} z^i \right\} \quad (z \in \mathrm{dom}\, f^*).$$

b) If $\bar{z} \in \sum_{i=1}^m \partial f_i(\bar{x})$*, then* $\mathcal{T}(\bar{z}) \neq \emptyset$ *and*

$$\mathcal{T}(\bar{z}) = \left\{ (z^1, \dots, z^m) \;\middle|\; \bar{z} = \sum_{i=1}^{m} z^i,\ z^i \in \partial f_i(\bar{x}),\ i = 1, \dots, m \right\}.$$

c) Under (7.1) we have $\partial f = \sum_{i=1}^m \partial f_i$*,* $\mathrm{dom}\,\partial f = \bigcap_{i=1}^m \mathrm{dom}\,\partial f_i$*, and*

$$\begin{aligned} \partial f(\bar{x}) &= \left\{ \sum_{i=1}^{m} z^i \;\middle|\; z^i \in \partial f_i(\bar{x}), i = 1, \dots, m \right\} \quad (\bar{x} \in \mathrm{dom}\,\partial f) \\ &= \left\{ \bar{z} \;\middle|\; (z^1, \dots, z^m) \in \mathcal{T}(\bar{z}) \ \ \textit{and} \ \ z^i \in \partial f_i(\bar{x})\ i = 1, \dots m \right\}. \end{aligned}$$

d) Under (7.1), $f^* = f_1^* \,\Box\, f_2^* \,\Box \cdots \Box\, f_m^*$*,* $\mathrm{dom}\,\partial f^* = \{ z \mid \emptyset \neq \mathcal{T}(z) \} \neq \emptyset$*, and*

$$\partial f^*(\bar{z}) = \left\{ \bigcap_{i=1}^{m} \partial f_i^*(z^i) \;\middle|\; \bar{z} = \sum_{i=1}^{m} z^i \right\} \quad (\bar{z} \in \mathrm{dom}\,\partial f^*).$$

Proof. a) See [15, Theorem 16.4]).
b) Let $L : \mathcal{E}^m \to \mathcal{E}$ be defined by $L(z^1, \dots, z^m) = \sum_{i=1}^m z^i$. Then its adjoint $L^* : \mathcal{E} \to \mathcal{E}^m$ is given by $L^*(x) = (x, \dots, x)$ $(x \in \mathcal{E})$. Let $\bar{z} \in \sum_{i=1}^m \partial f_i(\bar{x})$, and take any $z^i \in \partial f_i(\bar{x})$ $(i = 1, \dots, m)$ such that $\bar{z} = \sum_{i=1}^m z^i$. By Proposition [15, Theorem 23.5], $\bar{x} \in \partial f_i^*(z^i)$ $(i = 1, \dots, m)$. Hence, by [15, Theorem 23.8, 23.9] and [2, Proposition 16.8] we obtain

$$0 \in \mathrm{rge}\, L^* + \partial f_1^*(z^1) \times \cdots \times \partial f_m^*(z^m) \subset \partial(\delta_{\{0\}}\,(L(\cdot) - \bar{z}) + \oplus_{i=1}^m f_i^*)(z^1, \dots, z^m).$$

Hence, $(z^1, \dots, z^m) \in \mathcal{T}(\bar{z})$. This establishes that

$$\emptyset \neq \left\{ (z^1, \dots, z^m) \;\middle|\; \bar{z} = \sum_{i=1}^{m} z^i,\ z^i \in \partial f_i(\bar{x}),\ i = 1, \dots, m \right\} \subset \mathcal{T}(\bar{z}).$$

To see the reverse inclusion, let $(z^1, \dots, z^m) \in \mathcal{T}(\bar{z})$. By assumption and again [15, Theorem 23.8], we have $\bar{z} \in \sum_{i=1}^m \partial f_i(\bar{x}) \subset \partial f(\bar{x})$. By Proposition [15, Theorem 23.5] and the fact that $f^*(\bar{z}) = \sum_{i=1}^m f_i^*(z^i)$, we have

$$\sum_{i=1}^{m} \left\langle z^i, \bar{x} \right\rangle = \langle \bar{z}, \bar{x} \rangle = f^*(\bar{z}) + f(\bar{x}) = \sum_{i=1}^{m} (f_i^*(z^i) + f_i(\bar{x})),$$

so that

$$0 = \sum_{i=1}^{m} (f_i^*(z^i) + f_i(\bar{x}) - \left\langle z^i, \bar{x} \right\rangle).$$

By the Fenchel-Young inequality, $f_i^*(z^i) + f_i(\bar{x}) - \langle z^i, \bar{x} \rangle \geq 0$ $(i = 1, \ldots, m)$, hence equality must hold for each $i = 1, \ldots, m$, or equivalently $z^i \in \partial f_i(\bar{x})$ $(i = 1, \ldots, m)$. This establishes the reverse inclusion.

c) The first two consequences follow from [15, Theorem 23.8]. For the third, the first equivalence simply follows from the fact that $\partial f = \sum_{i=1}^m \partial f_i$. To see the second equivalence, let $\bar{z} \in \partial f(\bar{x})$. Then, by part b), $\mathcal{T}(\bar{z}) \neq \emptyset$, and, for every $(z^1, \ldots, z^m) \in \mathcal{T}(\bar{z})$, we have $z^i \in \partial f_i(\bar{x}),\ i = 1, \ldots, m$. Hence,

$$\partial f(\bar{x}) \subset \left\{ \bar{z} \;\middle|\; (z^1, \ldots, z^m) \in \mathcal{T}(\bar{z}),\ z^i \in \partial f_i(\bar{x}),\ i = 1, \ldots, m \right\}.$$

The reverse inclusion follows from the first equivalence.

d) By part a), $f^* = f_1^* \,\square\, f_2^* \,\square \cdots \square\, f_m^*$ is closed, proper, convex, and $\mathcal{T}(z) \neq \emptyset$ for all $z \in \operatorname{dom} f^*$.

Let us first suppose that $\bar{z} \in \operatorname{dom} \partial f^* \subset \operatorname{dom} f^*$, then $\mathcal{T}(\bar{z}) \neq \emptyset$. Let $\bar{x} \in \partial f^*(\bar{z})$. By [15, Theorem 23.5], $\bar{z} \in \partial f(\bar{x})$. By part c), this is equivalent to the existence of $z^i \in \partial f_i(\bar{x})$ such that $\bar{z} = \sum_{i=1}^m z^i$, which, by [15, Theorem 23.5], is equivalent to $\bar{x} \in \left\{ \bigcap_{i=1}^m \partial f_i^*(z^i) \;\middle|\; \bar{z} = \sum_{i=1}^m z^i \right\}$. Hence $\partial f^*(\bar{z}) \subset \left\{ \bigcap_{i=1}^m \partial f_i^*(z^i) \;\middle|\; \bar{z} = \sum_{i=1}^m z^i \right\}$.

On the other hand, let $\bar{x} \in \left\{ \bigcap_{i=1}^m \partial f_i^*(z^i) \;\middle|\; \bar{z} = \sum_{i=1}^m z^i \right\}$. Then, by [15, Theorem 23.5] we have $\bar{z} \in \partial f(\bar{x})$. But then, again by [15, Theorem 23.5], $\bar{x} \in \partial f^*(\bar{y})$. Finally, suppose that $(z^1, \ldots, z^m) \in \mathcal{T}(\bar{z}) \neq \emptyset$. Then, as in part a), $0 \in \operatorname{rge} L^* + \partial f_1^*(z^1) \times \cdots \times \partial f_m^*(z^m)$, or equivalently, there is an $\bar{x}$ such that $\bar{x} \in \bigcap_{i=1}^m \partial f_1^*(z^i)$ with $\bar{z} = \sum_{i=1}^m z^i$, i.e., $\bar{x} \in \partial f^*(\bar{z})$. This completes the proof. ∎

An interesting consequence of Proposition 7.1 a) is the following result.

COROLLARY 7.2 (Partial conjugates). *Let $f \in \Gamma(\mathcal{E}_1 \times \mathcal{E}_2)$ and $\bar{x} \in \mathcal{E}_1$ such that $\bar{g} := f(\bar{x}, \cdot)$ is proper. Then $\bar{g}^*$ is the closure of the function*

$$w \mapsto \inf_{z:(z,w) \in \operatorname{dom} f^*} \{ f^*(z, w) - \langle \bar{x}, z \rangle \}.$$

If $\bar{x} \in \operatorname{ri} L(\operatorname{dom} f)$, where $L : (x, v) \mapsto x$, then the closure can be dropped.

Proof. We use Proposition 7.1 a) throughout: Observe that

$$\begin{aligned}
\bar{g}^*(w) &= \sup_v \{ \langle v, w \rangle - f(\bar{x}, w) \} \\
&= \sup_{(x,v)} \{ \langle (x, v), (0, w) \rangle - (f + \delta_{\{\bar{x}\} \times \mathcal{E}_2})(x, v) \} \\
&= (f + \delta_{\{\bar{x}\} \times \mathcal{E}_2})^*(0, w) \\
&= \operatorname{cl} (f^* \,\square\, \sigma_{\{\bar{x}\} \times \mathcal{E}_2})(0, w).
\end{aligned}$$

Now notice that $\sigma_{\{\bar{x}\} \times \mathcal{E}_2} = \langle \bar{x}, \cdot \rangle \oplus \delta_{\{0\}}$. Hence

$$\begin{aligned}
(f^* \,\square\, \sigma_{\{\bar{x}\} \times \mathcal{E}_2})(0, w) &= \inf_{(z,u)} \{ f^*(z, u) + \langle \bar{x}, 0 - z \rangle + \delta_{\{0\}}(w - u) \} \\
&= \inf_{z:(z,w) \in \operatorname{dom} f^*} \{ f^*(z, w) - \langle \bar{x}, z \rangle \}.
\end{aligned}$$

This proves the first statement. Note that the closure can be dropped if $\operatorname{ri}(\operatorname{dom} f)$ and $\operatorname{ri}(\operatorname{dom} \delta_{\{\bar{x}\} \times \mathcal{E}_2}) = \{\bar{x}\} \times \mathcal{E}_2$ intersect, which is equivalent to the condition stated.

This concludes the proof.

REFERENCES

[1] A. Auslender and M. Teboulle: *Asymptotic Cones and Functions in Optimization and Variational Inequalities.* Springer Monographs in Mathematics, Springer, New York 2003.
[2] H.H. Bauschke and P.L. Combettes, *Convex analysis and Monotone Operator Theory in Hilbert Spaces.* CMS Books in Mathematics, Springer-Verlag, 2011.
[3] J.M. Borwein and A.S. Lewis: *Convex Analysis and Nonlinear Optimization. Theory and Examples.* CMS Books in Mathematics, Springer-Verlag, New York, 2000.
[4] S. Boyd and L. Vandenbergh: *Convex Optimization.* Cambridge University Press, 2004.
[5] J. V. Burke and T. Hoheisel, *Matrix support functionals for inverse problems, regularization, and learning.* SIAM Journal on Optimization 25, 2015, pp. 1135–1159.
[6] J. V. Burke, Y. Gao and T. Hoheisel: *Convex Geometry of the Generalized Matrix-Fractional Function.* SIAM Journal on Optimzation, to appear.
[7] J.V. Burke and R.A. Poliquin: *Optimality conditions for non-finite valued convex composite functions.* Mathematical Programming 57, 1992, pp. 103–120.
[8] X. V. Doan and S. Vavasis: *Finding the largest low-rank clusters with Ky Fan 2-k-norm and ℓ_1-norm.* arXiv:1403.5901, 2015.
[9] J. Dattorro: *Convex Optimization & Euclidean Distance Geometry.* M$\varepsilon\beta oo$ Publishing USA, Version 2014.04.08, 2005.
[10] J.-B. Hiriart-Urrruty and C. Lemaréchal: *Fundamentals of Convex Analysis.* Grundlehren Text Editions, Springer, Berlin, Heidelberg, 2001.
[11] R.A. Horn and C.R. Johnson: *Matrix Analysis.* Cambridge University Press, New York, N.Y., 1985.
[12] R.A. Horn and C. R. Johnson: *Topics in Matrix Analysis.* Cambridge University Press, New York, N.Y., 1991.
[13] C.-J. Hsieh and P. Olsen: *Nuclear Norm Minimization via Active Subspace Selection.* JMLR W&CP 32 (1), 2014, pp. 575-583.
[14] A. Jalali, M. Fazel, and L. Xiao: *Variational Gram functions: Convex analysis and optimization.* SIAM Journal on Optimization 27(4), 2017, pp. 2634–2661.
[15] R.T. Rockafellar, *Convex analysis*, Princeton University Press, 1970.
[16] R.T. Rockafellar and R.J.-B. Wets, *Variational analysis*, vol. 317, Springer, 1998.