SIAM J. OPTIM.
Vol. 29, No. 3, pp. 1958–1987

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) function 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 by-product 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

DOI. 10.1137/18M1209660

1. Introduction.

The *generalized matrix-fractional (GMF)* function was introduced by Burke and Hoheisel 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 expand the number of applications to include all *Ky Fan norms*, matrix *gauge functionals*, and *variational Gram functions* introduced by Jalali, Fazel, and Xiao in [14]. Our analysis includes descriptions of the variational properties of these functions, such as formulas for their convex conjugates and their subdifferentials.

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 $(A, B) \in \mathbb{R}^{\ell\times n} \times \mathbb{R}^{\ell\times m}$ 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., $\varphi : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$ is given by

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

where

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

---

*Received by the editors August 27, 2018; accepted for publication (in revised form) May 13, 2019; published electronically July 23, 2019.
https://doi.org/10.1137/18M1209660
Funding: The first author's research is supported in part by the National Science Foundation under grant DMS-1514559.
†Department of Mathematics, University of Washington, Seattle, WA 98195 (jvburke@uw.edu, yuangao@uw.edu).
‡Department of Mathematics and Statistics, McGill University, Montréal, H3A 0B9 Québec, Canada (tim.hoheisel@mcgill.ca).

A closed-form expression for φ is derived in [5, Theorem 4.1], where it is also shown that φ is smooth on the (nonempty) interior of its domain.

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

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

where $h : \mathbb{S}^n \to \mathbb{R} \cup \{+\infty\}$ is closed, proper, 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 both A and B are zero, p is a weighted nuclear norm where the weights depend on any "square root" of U (see Corollary 4.6). 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 the conjugate p^* and the subdifferential ∂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 5.8) as well as a complete characterization of variational Gram functions [14] and their conjugates. In addition, we show that all variational Gram functions are representable as squares of gauges (cf. Proposition 5.10). Other choices yield weighted sums of Frobenius and nuclear norms [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: Hsieh and Olsen [13] show that (1.3) with $h = \frac{1}{2}\operatorname{tr}(\cdot)$ yields a smoothing approach to optimization problems involving the nuclear norm. More generally, observe that many matrix optimization problems take the form

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

where $f, p : \mathbb{R}^{n\times m} \to \mathbb{R}\cup\{+\infty\}$. 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

$$\min_{(X,V)\in\mathbb{E}} f(X) + \varphi(X, V) + h(V). \tag{1.4}$$

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 a set. In particular, this yields powerful results on variational Gram functions and Ky Fan norms in sections 5.2 and 5.3. We close out with some final remarks in section 6 and supplementary material in Appendix A.

Notation. For a linear transformation L between finite-dimensional linear spaces, we write $\operatorname{rge} L$ and $\ker L$ for its *range* and *kernel*, respectively. For a given choice of bases, every such linear transformation has a matrix representation for some $A \in$

$\mathbb{R}^{\ell\times n}$. Therefore, we also write $\mathrm{rge}\,A$ and $\ker A$ for the *range* and *kernel*, respectively, considering A as a linear map between $\mathbb{R}^n$ and $\mathbb{R}^\ell$. Again, for $A\in\mathbb{R}^{\ell\times n}$, we set

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

and write $\mathrm{Ker}A$ or $\mathrm{Rge}A$ when the choice of r is clear. Observe that $\mathrm{Ker}_1 A=\ker A$, $\mathrm{Rge}_1 A=\mathrm{rge}\,A$, and $(\mathrm{Ker}_r A)^\perp = \mathrm{Rge}_r A^T$. We equip any matrix space with the (Frobenius) inner product $\langle X,Y\rangle := \mathrm{tr}(X^TY)$. The *Moore–Penrose pseudoinverse* [11] of A is denoted by $A^\dagger$. The set of all $n\times n$ symmetric matrices is given by $\mathbb{S}^n$. The positive and negative semidefinite cones 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 $\|\cdot\| := \sqrt{\langle\cdot,\cdot\rangle}$. 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 is denoted by $\mathrm{span}\,S$. The *affine hull* of S, denoted $\mathrm{aff}\,S$, is the intersection of all affine sets containing S, while the *convex hull* of S, denoted $\mathrm{conv}\,S$, is the intersection of all convex sets containing S. Its closure (the *closed convex hull*) is $\overline{\mathrm{conv}}\,S := \mathrm{cl}\,(\mathrm{conv}\,S)$. The *conical* and *convex conical* hull of S are given by $\mathrm{pos}\,S := \{\lambda x\mid x\in S,\ \lambda\ge 0\}$ and $\mathrm{cone}\ S := \{\sum_{i=1}^r\lambda_i x_i\mid r\in\mathbb{N},\ x_i\in S,\ \lambda_i\ge 0\}$, respectively, with $\mathrm{cone}\ S = \mathrm{pos}\,(\mathrm{conv}\,S) = \mathrm{conv}\,(\mathrm{pos}\,S)$. The closure of the latter is $\overline{\mathrm{cone}}\,S := \mathrm{cl}\,(\mathrm{cone}\ S)$.

The *relative interior* of a convex set $C\subset\mathcal{E}$, denoted $\mathrm{ri}\,C$, is the interior of C relative to its affine hull. By [2, section 6.2], we have

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

The *polar set* of $S\subset\mathcal{E}$ is defined by $S^\circ := \{v\in\mathcal{E}\mid \langle v,x\rangle\le 1\ (x\in S)\}$, and the *horizon cone* is the closed cone $S^\infty := \{v\in\mathcal{E}\mid \exists\{\lambda_k\}\downarrow 0,\ \{x_k\in S\}: \lambda_k x_k\to v\}$. For a convex set $C\subset\mathcal{E}$, C^∞ coincides with the *recession cone* of the closure of C, i.e.,

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

For $f:\mathcal{E}\to\overline{\mathbb{R}}$ its *domain* and *epigraph* are given by $\mathrm{dom}\,f := \{x\in\mathcal{E}\mid f(x)<+\infty\}$ and $\mathrm{epi}\,f := \{(x,\alpha)\in\mathcal{E}\times\mathbb{R}\mid f(x)\le\alpha\}$, respectively. We say f is *proper* if $f(x)>-\infty$ for all $x\in\mathrm{dom}\,f\neq\emptyset$. We call f *convex* if its epigraph $\mathrm{epi}\,f$ is convex, and *closed* (or *lower semicontinuous*) if $\mathrm{epi}\,f$ is closed. If f is proper, we call it *positively homogeneous* if $\mathrm{epi}\,f$ is a cone, and *sublinear* if $\mathrm{epi}\,f$ is a convex cone. In what follows we use the following abbreviations:

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

The *lower semicontinuous hull* $\mathrm{cl}\,f$ and the *horizon function* f^∞ of f are defined through the relations $\mathrm{cl}\,(\mathrm{epi}\,f) = \mathrm{epi}\,\mathrm{cl}\,f$ and $\mathrm{epi}\,f^\infty = (\mathrm{epi}\,f)^\infty$, respectively. For $f\in\Gamma_0(\mathcal{E})$, f^∞ is also known as the *recession function* [15, Page 66] or the *asymptotic function* [1, 10]. The *horizon cone of a function* f is defined as $\mathrm{hzn}\,f := \{x\mid f^\infty(x)\le 0\}$, and for $f\in\Gamma_0$, we have $\mathrm{hzn}\,f = \{x\mid f(x)\le\mu\}^\infty$ for $\mu\in\mathbb{R}$ such that $\{x\mid f(x)\le\mu\}\neq\emptyset$ [15, Theorem 8.7].

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

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

For $f \in \Gamma_0(\mathcal{E})$, we have $\operatorname{ri}(\operatorname{dom} f) \subset \operatorname{dom} \partial f \subset \operatorname{dom} f$; see, e.g., [15, Page 227], where $\operatorname{dom} \partial f := \{x \in \mathcal{E} \mid \partial f(x) \neq \emptyset\}$ is the *domain of the subdifferential*.

For a function $f : \mathcal{E} \to \overline{\mathbb{R}}$, its *(Fenchel) conjugate* $f^* : \mathcal{E} \to \overline{\mathbb{R}}$ is given by $f^*(y) := \sup_{x \in \mathcal{E}}\{\langle x,\, y\rangle - f(x)\}$, and $f \in \Gamma_0(\mathcal{E})$ if and only if $f = f^{**} := (f^*)^*$ is proper.

Given a nonempty $S \subset \mathcal{E}$, its *indicator function* $\delta_S : \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ is given by $\delta_S(x) = 0$ for $x \in S$ and $+\infty$ otherwise. 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 $N_S(\bar{x}) := \partial \delta_S(\bar{x}) = \{v \in \mathcal{E} \mid \langle v,\, x - \bar{x}\rangle \leq 0 \ (x \in S)\}$. 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 by $\sigma_S(x) := \sup_{v \in S}\langle v,\, x\rangle$ and $\gamma_S(x) := \inf\{t \geq 0 \mid x \in tS\}$, respectively. Here we use the standard convention that $\inf \emptyset = +\infty$.

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

$$(\operatorname{bar} C)^\circ = C^\infty \quad \text{and} \quad \operatorname{cl}(\operatorname{bar} C) = (C^\infty)^\circ. \tag{2.3}$$

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

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

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

$$\varphi(X,V) = \sigma_{\mathcal{D}(A,B)}(X,V) \tag{2.4}$$

and also refer to $\sigma_{\mathcal{D}(A,B)}$ as the GMF function. From [5, Theorem 4.1], we obtain the formula

$$\varphi(X,V) = \begin{cases} \frac{1}{2}\operatorname{tr}\left(\binom{X}{B}^T M(V)^\dagger \binom{X}{B}\right) & \text{if } \operatorname{rge}\binom{X}{B} \subset \operatorname{rge} M(V),\ V \in \mathcal{K}_A, \\ +\infty & \text{else,} \end{cases} \tag{2.5}$$

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

$$\mathcal{K}_A := \{V \in \mathbb{S}^n \mid u^T V u \geq 0 \ (u \in \ker A)\}, \tag{2.6}$$

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

$$M(V) = \begin{pmatrix} V & A^T \\ A & 0 \end{pmatrix}. \tag{2.7}$$

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

The GMF function $\varphi = \sigma_{\mathcal{D}(A,B)}$ appears in Burke and Hoheisel [5] and Burke, Hoheisel, and Gao [6], where it is shown that

$$\begin{aligned} \operatorname{dom}\varphi = \operatorname{dom}\partial\varphi &= \left\{(X,V) \in \mathbb{E} \;\middle|\; \operatorname{rge}\begin{pmatrix} X \\ B \end{pmatrix} \subset \operatorname{rge} M(V),\ V \in \mathcal{K}_A\right\}, \\ \operatorname{int}(\operatorname{dom}\varphi) &= \left\{(X,V) \in \mathbb{E} \;\middle|\; \operatorname{rge}\begin{pmatrix} X \\ B \end{pmatrix} \subset \operatorname{rge} M(V),\ V \in \operatorname{int}\mathcal{K}_A\right\} \neq \emptyset. \end{aligned} \tag{2.8}$$

For a deeper understanding of the support function φ, a description of the closed convex hull of the (nonconvex) set $\mathcal{D}(A,B)$ is critical. An arduous representation of $\overline{\operatorname{conv}}\,\mathcal{D}(A,B)$ was obtained in [5, Proposition 4.3]. A much simpler and more versatile expression was proven in [6, Theorem 2]; see below. The key ingredient in the newer expression is the (closed, convex) cone $\mathcal{K}_A$ defined in (2.6), which reduces to $\mathbb{S}^n_+$ when $A=0$. We briefly summarize the geometric and topological properties of $\mathcal{K}_A$ that are useful to our study. These follow from [6, Proposition 1] (by setting $\mathcal{S}=\ker A$).

Proposition 2.1. *For $A\in\mathbb{R}^{\ell\times n}$ let $P\in\mathbb{R}^{n\times n}$ be the orthogonal projection onto $\ker A$ and let $\mathcal{K}_A$ be given by* (2.6). *Then the following hold:*

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

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

(c) $\operatorname{int}\mathcal{K}_A=\{V\in\mathbb{S}^n \mid u^TVu>0\ (u\in\ker A\setminus\{0\})\}$.

The central result in [6] now follows.

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

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

In particular, Theorem 2.2 implies that $\varphi=\sigma_{\mathcal{D}(A,B)}=\sigma_{\Omega(A,B)}$, since $\sigma_S=\sigma_{\overline{\operatorname{conv}}\,S}$ for all subsets S of a Euclidean space. This identity is used throughout.

3. Infimal projections of the generalized matrix-fractional function.

We now focus on infimal projections involving the GMF function. Consider

$$\psi:\mathbb{E}\to\overline{\mathbb{R}},\quad \psi(X,V)=\varphi(X,V)+h(V), \tag{3.1}$$

where $\varphi\in\Gamma_0(\mathbb{E})$ is given in (1.1) and $h\in\Gamma_0(\mathbb{S}^n)$. Our primary object of study is the infimal projection of the sum ψ in the variable V under the standing assumption that $\operatorname{rge}B\subset\operatorname{rge}A$, i.e., $\{Y\in\mathbb{R}^{n\times m} \mid AY=B\}\neq\emptyset$:

$$p:\mathbb{R}^{n\times m}\to\overline{\mathbb{R}},\quad p(X)=\inf_{V\in\mathbb{S}^n}\psi(X,V). \tag{3.2}$$

We lead with some elementary observations.

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

(a) *p is convex.*

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

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

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

(d) *if $\operatorname{rank}A=\ell$ (full row rank) and $\operatorname{dom}h\cap\operatorname{int}\mathcal{K}_A\neq\emptyset$, then $\operatorname{dom}p=\mathbb{R}^{n\times m}$;*

(e) *if $\operatorname{dom}h\cap\mathcal{K}_A\neq\emptyset$ and $(\operatorname{dom}h)^\infty\cap\mathcal{K}_A=\{0\}$, then p is proper, and hence $p\in\Gamma$.*

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

(b) We have $X\in\operatorname{dom}p$ if and only if there is a $V\in\mathbb{S}^n$ such that $(X,V)\in\operatorname{dom}\psi=(\operatorname{dom}\varphi)\cap(\mathbb{R}^{n\times m}\times\operatorname{dom}h)$. Hence, the representation for $\operatorname{dom}p$ follows from that of $\operatorname{dom}\varphi$ in (2.8). This representation for $\operatorname{dom}p$ tells us that $\operatorname{dom}p\neq\emptyset$ implies that $\operatorname{dom}h\cap\mathcal{K}_A\neq\emptyset$. On the other hand, if $V\in\operatorname{dom}h\cap\mathcal{K}_A$, then $(VY,V)\in\operatorname{dom}\psi$ for any $Y\in\mathbb{R}^{n\times m}$ satisfying $AY=B$, and so $(VY,V)\in\operatorname{dom}p\neq\emptyset$.

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

(d) By the description of $\operatorname{int} \mathcal{K}_A$ in Proposition 2.1(c), the assumptions imply that there exists $V \in \operatorname{dom} h \cap \mathcal{K}_A$ such that $M(V)$ is invertible; see [5, Proposition 3.3]. This readily gives the desired statement in view of (b).

(e) By part (b), we have $\operatorname{dom} p \neq \emptyset$. Hence, let $X \in \operatorname{dom} p$, i.e., there is a $V \in \mathcal{K}_A \cap \operatorname{dom} h$ such that $\operatorname{rge}\left(\begin{smallmatrix} X \\ B \end{smallmatrix}\right) \subset \operatorname{rge} M(V)$. If $p(X) = -\infty$, there is a sequence $\{V_k \in \mathbb{S}^n \cap \operatorname{dom} h\}$ with $\{(X, V_k) \in \operatorname{dom} \varphi\}$ such that $\psi(X, V_k) \to -\infty$. This implies that $\varphi(X, V_k) \to -\infty$ or $h(V_k) \to -\infty$. In either case, this tells us that $\|V_k\| \to \infty$ since both φ and h are closed and proper. Consequently, there are a subsequence $J \subset \mathbb{N}$ and a matrix $\widehat{V} \in \mathbb{S}^n$ such that $(V_v/\|V_k\|) \xrightarrow{J} \widehat{V}$. Hence,

$$0 \neq \widehat{V} \in (\operatorname{dom} h \cap \mathcal{K}_A)^\infty = (\operatorname{dom} h)^\infty \cap \mathcal{K}_A,$$

which contradicts the hypothesis. ∎

We give two examples to illustrate the various statements in Lemma 3.1. The first shows that an assumption of the type in part (e) is required to establish that p is proper.

Example 3.2 (p is improper). Let $m = n = 1$, $A = 0$, $B = 0$, and $h(v) = -v$. Since

$$v^\dagger = \begin{cases} \frac{1}{v} & \text{if } v \neq 0, \\ 0 & \text{if } v = 0, \end{cases}$$

we have

$$\varphi(x, v) = \begin{cases} \frac{x^2}{2v} & \text{if } v > 0, \\ 0 & \text{if } v = 0, \\ +\infty & \text{if } v < 0 \end{cases} \quad ((x, v) \in \mathbb{R}^2).$$

Therefore, $p \equiv -\infty$ since

$$p(x) = \inf_{v \in \mathbb{R}} \varphi(x, v) + h(v) = \inf_{v > 0} \left\{ \frac{x^2}{2v} - v \right\} = -\infty \quad (x \in \mathbb{R}).$$

The properness condition given in Lemma 3.1(e) is revisited in Definition 3.10, where it is called the *boundedness primal constraint qualification* (BPCQ). It is the strongest of the constraint qualifications we discuss.

The second example shows $\operatorname{dom} p$ may not be relatively open if $B \neq 0$.

Example 3.3 ($\operatorname{dom} p$ is not relatively open). Let $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ and $b = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$. Then $\ker A = \operatorname{span}\left\{\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)\right\}$ and $\mathcal{K}_A = \left\{\left(\begin{smallmatrix} v & w \\ w & u \end{smallmatrix}\right) \mid v + u \geq 2w\right\}$. Moreover, put $\bar{V} := \left(\begin{smallmatrix} 2 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and define $\mathcal{V} := [0,1] \cdot \bar{V} = \left\{\left(\begin{smallmatrix} 2w & w \\ w & 0 \end{smallmatrix}\right) \mid w \in [0,1]\right\} \subset \mathbb{S}^2$. Then $\mathcal{V}$ is convex and compact. Let $h \in \Gamma_0(\mathbb{S}^2)$ be any function with $\operatorname{dom} h = \mathcal{V}$. Note that $\operatorname{dom} h \cap \mathcal{K}_A = \mathcal{V}$. Hence,

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

Therefore, $\operatorname{dom} p = [0,1] \cdot \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) + \operatorname{span}\left\{\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)\right\}$, and hence $\operatorname{ri}(\operatorname{dom} p) = (0,1) \cdot \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) + \operatorname{span}\left\{\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)\right\}$, so that $\operatorname{dom} p$ is clearly not relatively open.

The preceding example shows that $\operatorname{dom} p$ may fail to be a subspace if $B \neq 0$; hence, this assumption in Lemma 3.1(c) is not superfluous. On the other hand, Lemma 3.1(d) and Example 3.18(a) illustrate that the condition $B = 0$ is only sufficient but not necessary for $\operatorname{dom} p$ to be a subspace.

3.1. The functions ψ, ψ^*, and their subdifferentials. The study of the infimal projection p in (3.2) requires an understanding of the properties of the function ψ from (3.1), its conjugate ψ^*, and their subdifferentials. For this we make extensive use of the *conjugate constraint qualification*

$$\text{(CCQ)} \qquad \operatorname{ri}(\operatorname{dom} h) \cap \operatorname{int} \mathcal{K}_A \neq \emptyset.$$

As a direct consequence of the *line segment principle* (cf. [15, Theorem 6.1]), we have

$$\operatorname{ri}(\operatorname{dom} h) \cap \operatorname{int} \mathcal{K}_A \neq \emptyset \quad \iff \quad \operatorname{dom} h \cap \operatorname{int} \mathcal{K}_A \neq \emptyset. \tag{3.3}$$

Lemma 3.4 (conjugate of ψ). *Let ψ be given as in* (3.1) *and define*

$$\eta : (Y, W) \in \mathbb{E} \mapsto \inf_{T \in \mathbb{S}^n} h^*(W - T) + \delta_{\Omega(A,B)}(Y, T). \tag{3.4}$$

Then

$$\begin{aligned} \operatorname{dom} \eta &= \Omega(A, B) + (\{0\} \times \operatorname{dom} h^*) \\ &= \left\{ (Y, W) \,\middle|\, AY = B,\ \left(-\frac{1}{2} YY^T + \mathcal{K}_A^\circ\right) \cap (W - \operatorname{dom} h^*) \neq \emptyset \right\}, \end{aligned} \tag{3.5}$$

and the following hold:

(a) *if $\psi \not\equiv +\infty$, then $\psi \in \Gamma_0(\mathbb{E})$;*

(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.,*

$$\mathfrak{S}(\bar{Y}, \bar{W}) := \operatorname*{argmin}_{T \in \mathbb{S}^n} \{ h^*(\bar{W} - T) \mid (\bar{Y}, T) \in \Omega(A, B) \} \tag{3.6}$$

is nonempty for all $(\bar{Y}, \bar{W}) \in \operatorname{dom} \psi^*$;

(d) *under* (CCQ), $\operatorname{dom} \partial\psi^* = \{(Y, W) \mid \emptyset \neq \mathfrak{S}(Y, W)\}$ *and, for every $(Y, W) \in \operatorname{dom} \partial\psi^*$, we have*

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

Proof. Note that $\eta(Y, W) < +\infty$ if and only if there are $W_1, W_2 \in \mathbb{S}^n$ such that $W = W_1 + W_2$, $(Y, W_1) \in \Omega(A, B)$, and $W_2 \in \operatorname{dom} h^*$, or equivalently, $(Y, W) \in \Omega(A, B) + (\{0\} \times \operatorname{dom} h^*)$, which in turn is equivalent to $AY = B$, $T \in -\frac{1}{2} YY^T + \mathcal{K}_A^\circ$, and $T \in W - \operatorname{dom} h^*$, giving (3.5).

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 $\psi = \varphi + \hat{h} = \sigma_{\Omega(A,B)} + \hat{h}$.

(a) The sum of two closed, proper, convex functions (here φ and $\hat{h}$) is closed and convex. It is proper if and (only) if the sum is not constantly $+\infty$.

(b) The sum of two proper functions is proper if and only if the domains of both functions intersect. By (2.8), we have $\mathrm{dom}\,\hat{h}\cap\mathrm{dom}\,\varphi \neq \emptyset$ if and only if $\mathrm{dom}\,h\cap\mathcal{K}_A \neq \emptyset$. Therefore, ψ is proper if (and only if) the latter condition holds. Combined with (a) this shows $\psi \in \Gamma_0(\mathbb{E})$, and so $\psi^* \in \Gamma_0(\mathbb{E})$. Moreover, by Theorem A.1(a), $\psi^*(Y,W) = \mathrm{cl}\,(\delta_{\Omega(A,B)} \,\Box\, \hat{h}^*)(Y,W)$. Since $\hat{h}^*(Y,W) = \delta_{\{0\}}(Y) + h^*(W)$, $(\delta_{\Omega(A,B)} \,\Box\, \hat{h}^*)(Y,W) = \inf_{(Y,T)\in\Omega(A,B)} h^*(W-T) = \eta(Y,W)$, proving $\psi^* = \mathrm{cl}\,\eta$.

(c) By [5, Theorem 4.1], $\mathrm{int}\,(\mathrm{dom}\,\varphi) = \{(X,V) \mid V \in \mathrm{int}\,\mathcal{K}_A\}$ and, by definition, $\mathrm{ri}\,(\mathrm{dom}\,\hat{h}) = \mathbb{R}^{n\times m} \times \mathrm{ri}\,(\mathrm{dom}\,h)$. Hence,

$$\mathrm{ri}\,(\mathrm{dom}\,\hat{h}) \cap \mathrm{ri}\,(\mathrm{dom}\,\varphi) \neq \emptyset \quad \Longleftrightarrow \quad \mathrm{ri}\,(\mathrm{dom}\,h) \cap \mathrm{int}\,\mathcal{K}_A \neq \emptyset. \tag{3.7}$$

Theorem A.1(a) (applied to φ and $\hat{h}$), (CCQ), and (3.7) imply $\psi^* = \eta$ with

$$\emptyset \neq \mathcal{T}(\bar{Y},\bar{W}) := \underset{(Y,T),(0,W)\in\mathbb{E}}{\mathrm{argmin}} \{h^*(W) \mid (Y,T) \in \Omega(A,B),\ Y = \bar{Y},\ \bar{W} = W + T\}. \tag{3.8}$$

Since

$$\begin{aligned} &\mathfrak{S}(\bar{Y},\bar{W}) = \{T \in \mathbb{S}^n \mid [(\bar{Y},T),(0,\bar{W}-T)] \in \mathcal{T}(\bar{Y},\bar{W})\} \quad \text{and} \\ &\mathcal{T}(\bar{Y},\bar{W}) = \{[(\bar{Y},T),(0,\bar{W}-T)] \mid T \in \mathfrak{S}(\bar{Y},\bar{W})\}, \end{aligned} \tag{3.9}$$

we have $\mathfrak{S}(\bar{Y},\bar{W}) \neq \emptyset$ if and only if $\mathcal{T}(\bar{Y},\bar{W}) \neq \emptyset$.

(d) Observe that $\partial\varphi^* = N_{\Omega(A,B)}$ and $\partial\hat{h}^* = \mathbb{R}^{n\times m} \times \partial h^*$ with $\mathrm{dom}\,\partial\hat{h}^* = \{0\} \times \mathrm{dom}\,\partial h^*$. Then part (c) and Theorem A.1(d) (applied to φ and $\hat{h}$) yield

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

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

COROLLARY 3.5 (subdifferential of ψ). *Let ψ be given by* (3.1)*, and $\mathfrak{S}$ by* (3.6). *Then the following hold.*

(a) *If $(\bar{Y},\bar{W}) \in \partial\varphi(\bar{X},\bar{V}) + (\{0\} \times \partial h(\bar{V}))$, then $\mathfrak{S}(\bar{Y},\bar{W}) \neq \emptyset$ and*

$$\mathfrak{S}(\bar{Y},\bar{W}) = \{T \in \mathbb{S}^n \mid \bar{W} - T \in \partial h(\bar{V}),\ (\bar{Y},T) \in \partial\varphi(\bar{X},\bar{V})\}, \tag{3.10}$$

where $\partial\varphi$ is described in [6, Corollary 3.2].

(b) *Under* (CCQ) *we have*

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

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

$$\begin{aligned} \partial\psi(\bar{X},\bar{V}) &= \partial\varphi(\bar{X},\bar{V}) + (\{0\} \times \partial h(\bar{V})) \\ &= \{(\bar{Y},\bar{W}) \in \mathbb{E} \mid \mathfrak{S}(\bar{Y},\bar{W}) \neq \emptyset\}. \end{aligned} \tag{3.11}$$

Proof. Set $f_1(X,V) := \varphi(X,V)$ and $f_2(X,V) := h(V)$, so that the mapping $\mathcal{T}$ in Theorem A.1 is given by (3.8). Then, using (3.9), part (a) follows from Theorem A.1(b), and part (b) follows from Theorem A.1(c). ◻

3.2. Infimal projection I. Let the infimal projection p be as given in (3.2). We are now in position to give a formula for p^* under (CCQ).

THEOREM 3.6 (conjugate of p and properties under (CCQ)). *Let p be given by* (3.2)*. Moreover, let $\eta_0 : \mathbb{R}^{n\times m} \to \overline{\mathbb{R}}$ be given by*

$$\eta_0 : Y \mapsto \inf_{(Y,-W)\in\Omega(A,B)} h^*(W). \tag{3.12}$$

Then the following hold.

(a) *We have*

$$\begin{aligned}\operatorname{dom}\eta_0 &= \left\{ Y \in \mathbb{R}^{n\times m} \;\middle|\; AY = B,\ \left(-\frac{1}{2}YY^T + \mathcal{K}_A^\circ\right) \cap (-\operatorname{dom} h^*) \neq \emptyset \right\}\\ &= \{Y \in \mathbb{R}^{n\times m} \mid (Y, 0) \in \operatorname{dom}\eta\},\end{aligned}$$

where η is defined in (3.4)*.*

(b) *If $\operatorname{dom} h \cap \mathcal{K}_A \neq \emptyset$, then $p^* = \operatorname{cl}\eta_0$; hence, $\operatorname{dom}\eta_0 \subset \operatorname{dom} p^*$.*

(c) *If* (CCQ) *holds for p, then $\operatorname{dom} p = \mathbb{R}^{n\times m}$ and the following hold.*

(I) *$p^* = \eta_0$, i.e.,*

$$p^*(Y) = \inf_{(Y,-W)\in\Omega(A,B)} h^*(W). \tag{3.13}$$

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 convex under (CCQ) *if and only if it is proper, which is the case if and only if*

$$\begin{aligned}\emptyset \neq \operatorname{dom}\psi^*(\cdot, 0) &= \{Y \mid \exists W \in \operatorname{dom} h^* : (Y, -W) \in \Omega(A, B)\}\\ &= \{Y \mid (Y, 0) \in \Omega(A, B) + (\{0\} \times \operatorname{dom} h^*)\},\end{aligned}$$

with $\operatorname{dom} p^* = \operatorname{dom}\psi^*(\cdot, 0) = \operatorname{dom}\eta_0$.

(II) *Either p is (convex) finite valued (hence $p \in \Gamma_0(\mathbb{R}^{n\times m})$) or $p \equiv -\infty$. The former is the case if and only if $\operatorname{dom}\psi^*(\cdot, 0) \neq \emptyset$.*

Proof. (a) This follows from the definition of η_0. Also note that $\eta_0 = \eta(\cdot, 0)$.

(b) By Lemma 3.4(b), $\psi^* \in \Gamma_0(\mathbb{E})$ with $\psi^* = \operatorname{cl}\eta$ and η defined in (3.4). Hence, by [16, Theorem 11.23(c)], $p^* = \psi^*(\cdot, 0)$, which establishes the given representation. The domain containment is clear as $p^* = \operatorname{cl}\eta_0 \leq \eta_0$.

(c) Observe that $\operatorname{dom} p = L(\operatorname{dom}\varphi \cap \mathbb{R}^{n\times m} \times \operatorname{dom} h)$, where $L : (X, V) \mapsto X$; see Lemma 3.1. By (CCQ), we have $\operatorname{ri}(\operatorname{dom} h) \cap \operatorname{int}\mathcal{K}_A \neq \emptyset$; hence,

$$\begin{aligned}\operatorname{ri}(\operatorname{dom}\varphi \cap (\mathbb{R}^{n\times m} \times \operatorname{dom} h)) &= \operatorname{int}(\operatorname{dom}\varphi) \cap (\mathbb{R}^{n\times m} \times \operatorname{ri}(\operatorname{dom} h))\\ &= (\mathbb{R}^{n\times m} \times \operatorname{int}\mathcal{K}_A) \cap (\mathbb{R}^{n\times m} \times \operatorname{ri}(\operatorname{dom} h))\\ &= \mathbb{R}^{n\times m} \times (\operatorname{int}\mathcal{K}_A \cap \operatorname{ri}(\operatorname{dom} h)),\end{aligned}$$

where we use [5, Theorem 4.1] to represent $\operatorname{int}(\operatorname{dom}\varphi)$. This now gives

$$\operatorname{ri}(\operatorname{dom} p) = L\left[\operatorname{ri}(\operatorname{dom}\varphi \cap \mathbb{R}^{n\times m} \times \operatorname{dom} h)\right] = \mathbb{R}^{n\times m}.$$

(I) As in part (b), $p^* = \psi^*(\cdot, 0)$. Hence, Lemma 3.4(c) gives the identity $p^* = \eta_0$ under (CCQ) as well as the attainment statement. Since ψ^* is closed, proper, convex (under (CCQ)) by Lemma 3.4(b), $\psi^*(\cdot,0)$ is too, if and only if $\mathrm{dom}\,\psi^*(\cdot,0)\neq\emptyset$, and so the statements about $p^*=\psi^*(\cdot,0)$ follow.

(II) We have $\mathrm{dom}\,p=\mathbb{R}^{n\times m}$. By [15, Corollary 7.2.3] this implies that either $p\equiv-\infty$ or p is finite valued, which shows the first statement. For the second statement, again, as $\mathrm{dom}\,p=\mathbb{R}^{n\times m}$, observe that the convex function p is finite valued if and only if it is proper, which is true if and only if p^* is proper, so (I) gives the desired statement. ∎

Observe that Example 3.2 shows that the condition $\emptyset\neq\mathrm{dom}\,\psi^*(\cdot,0)$ is essential in parts (I) and (II) of Theorem 3.6. Indeed, in this example, $p\equiv-\infty$ so $\mathrm{dom}\,p=\mathbb{R}$, while $h=\sigma_{\{-1\}}$ and $h^*=\delta_{\{-1\}}$, $\psi^*(\cdot,0)=p^*\equiv\infty$, and (CCQ) is satisfied.

We now broaden our perspective of the infimal projection by embedding it into a *perturbation duality framework* in the sense of [16, Theorem 11.39] or the development in [1, Chapter 5]. Given $\bar X\in\mathbb{R}^{n\times m}$, define $f_{\bar X}$ by

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

and $p_{\bar X}$ by

$$p_{\bar X}(X):=\inf_{V\in\mathbb{S}^n}f_{\bar X}(X,V)\quad(X\in\mathbb{R}^{n\times m}).\tag{3.14}$$

Then $f^*_{\bar X}(Y,W)=\psi^*(Y,W)-\langle\bar X,\,Y\rangle$ $((Y,W)\in\mathbb{E})$ [16, equation 11(3)]. Define

$$q_{\bar X}(W):=-\sup_{Y}\{\langle\bar X,\,Y\rangle-\psi^*(Y,W)\}\quad(W\in\mathbb{S}^n).\tag{3.15}$$

Then $q_{\bar X}$ is a convex function that pairs in duality with $p_{\bar X}$ satisfying the weak duality $p_{\bar X}(0)\geq-q_{\bar X}(0)$ $(\bar X\in\mathbb{R}^{n\times m})$. Applying the general perturbation duality to our scenario yields the following result.

PROPOSITION 3.7 (shifted duality for p). *Let p be defined by* (3.2)*, and let $\bar X\in\mathrm{dom}\,p$ and $q_{\bar X}$ be defined by* (3.15)*. Then the following hold.*

(a) *If $0\in\mathrm{ri}\,(\mathrm{dom}\,q_{\bar X})$, then $p(\bar X)=-q_{\bar X}(0)\in\mathbb{R}$, $\mathrm{argmin}\,\psi(\bar X,\cdot)\neq\emptyset$, 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 $0\in\mathrm{ri}\,(\mathrm{dom}\,q_{\bar X})$ or $\bar X\in\mathrm{ri}\,(\mathrm{dom}\,p)$, p is l.s.c. at $\bar X$ and $-q_{\bar X}$ is l.s.c. at 0.*

(d) *We have*

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

Proof. Let $\bar X\in\mathrm{dom}\,p$ and observe that $p(X+\bar X)=p_{\bar X}(X)$ $(X\in\mathbb{R}^{n\times m})$; hence, in particular, $p(\bar X)=p_{\bar X}(0)\in\mathbb{R}$. Moreover, notice that ψ, and hence $f_{\bar X}$ is proper (and hence in Γ_0) as, by assumption, $\bar X\in\mathrm{dom}\,p$ exists. Applying the results [1, Theorems 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 $q_{\bar X}$ plays a key role in interpreting this result in a given setting. Below we provide a useful representation of this domain using the set

$$\Omega_2(A,B):=\{W\in\mathbb{S}^n\mid\exists Y:(Y,W)\in\Omega(A,B)\}.\tag{3.16}$$

LEMMA 3.8 (domain of $q_{\bar X}$). *Let $\bar X \in \mathbb{R}^{n\times m}$ and $q_{\bar X}$ defined by* (3.15). *Then* $\mathrm{dom}\, q_{\bar X} = \Omega_2(A,B) + \mathrm{dom}\, h^*$.

Proof. Using Lemma 3.4, 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,

$$\begin{aligned} \mathrm{dom}\, q_{\bar X} &= \{W \in \mathbb{S}^n \mid \exists (Y,T) \in \Omega(A,B) : W - T \in \mathrm{dom}\, h^*\} \\ &= \Omega_2(A,B) + \mathrm{dom}\, h^*. \end{aligned}$$ ◻

We now discuss various constraint qualifications for p.

3.3. Constraint qualifications. We start our analysis with a result about the set $\Omega_2(A,B)$ from (3.16), which was used in Lemma 3.8 to represent the domain of $q_{\bar X}$.

LEMMA 3.9 (properties of $\Omega_2(A,B)$). *Let $\Omega_2(A,B)$ be as in* (3.16). *Then we have the following:*

(a) *$\Omega_2(A,B)$ is closed and convex with $\Omega_2(A,B)^\infty = \mathcal{K}_A^\circ$.*
(b) *$\Omega_2(A,B) = \mathrm{dom}\,\varphi(\bar X,\cdot)^*$ for all $\bar X \in \mathbb{R}^{n\times m}$ such that $\varphi(\bar X,\cdot)$ is proper.*
(c) *For all $\bar X$ such that $\varphi(\bar X,\cdot)$ is proper,*

$$\begin{aligned} \mathrm{ri}\,\Omega_2(A,B) &= \left\{ W \;\middle|\; \exists Y : AY = B,\ \frac{1}{2}YY^T + W \in \mathrm{ri}\,(\mathcal{K}_A^\circ)\right\} \\ &= \mathrm{ri}\,(\mathrm{dom}\,\varphi(\bar X,\cdot)^*). \end{aligned}$$

Proof. (a) With the linear map $T : (Y,W) \mapsto W$ we have $\Omega_2(A,B) = T(\Omega(A,B))$. Therefore, $\Omega_2(A,B)$ is convex. By [6, Proposition 10], we have $\Omega(A,B)^\infty = \{0\}\times\mathcal{K}_A^\circ$, and so $\ker T \cap \Omega(A,B)^\infty = \{0\}$, giving the remainder of (a) by [16, Theorem 3.10].

(b) Recall from [5, Theorem 4.1] that $\mathrm{int}\,(\mathrm{dom}\,\sigma_\varphi) = \{(X,V) \in \mathbb{E} \mid V \in \mathrm{int}\,\mathcal{K}_A\}$. Thus, we can apply Proposition A.2 to $\bar g := \varphi(\bar X,\cdot)$ 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 $\mathrm{ri}\,\Omega_2(A,B) = \mathrm{ri}\,T(\Omega(A,B)) = T(\mathrm{ri}\,\Omega(A,B))$ 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 the conjugate constraint qualification (CCQ) was introduced in section 3.1.

DEFINITION 3.10 (constraint qualifications). *Let p be given by* (3.2). *We say that p satisfies the following:*

(i) *the* primal constraint qualification (PCQ) *if* $0 \in \mathrm{ri}\,(\Omega_2(A,B) + \mathrm{dom}\, h^*)$*;*
(ii) *the* strong PCQ (SPCQ) *if* $0 \in \mathrm{int}\,(\Omega_2(A,B) + \mathrm{dom}\, h^*)$*;*
(iii) *the* boundedness PCQ (BPCQ) *if* $\mathrm{dom}\, h \cap \mathcal{K}_A \neq \emptyset$ *and* $(\mathrm{dom}\, h)^\infty \cap \mathcal{K}_A = \{0\}$*;*
(iv) *the* CCQ *if* $\mathrm{ri}\,(\mathrm{dom}\, h) \cap \mathrm{int}\,\mathcal{K}_A \neq \emptyset$*;*

(v) *the* strong CCQ (SCCQ) *if the CCQ is satisfied and* $\emptyset \neq \operatorname{dom}\psi^*(\cdot,0)$, *or equivalently,*

$$\begin{aligned}\emptyset \neq \Xi(A,B) &:= \left\{ Y\in\mathbb{R}^{n\times m} \;\middle|\; AY=B,\ \frac{1}{2}YY^T \in \operatorname{dom} h^* + \mathcal{K}_A^\circ \right\} \\ &= \{Y\in\mathbb{R}^{n\times m} \mid (Y,0)\in\Omega(A,B)+(\{0\}\times\operatorname{dom} h^*)\}.\end{aligned} \tag{3.17}$$

Theorem 3.6 and Lemma 3.8, respectively, give the following useful implications:

$$\begin{aligned}\text{CCQ} &\implies \operatorname{dom} p^* = \Xi(A,B),\\ \text{SCCQ} &\implies \operatorname{dom} p^* = \Xi(A,B)\neq\emptyset.\end{aligned} \tag{3.18}$$

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

Lemma 3.11 (characterizations of (B)PCQ). *Let p be given by* (3.2) *and $\bar X \in \operatorname{dom} p$, and set*

$$\psi_{\bar X} := \psi(\bar X,\cdot) \quad (\bar X\in\mathbb{R}^{n\times m}). \tag{3.19}$$

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

(b) *The BPCQ holds for p if and only if* $\operatorname{dom} h\cap\mathcal{K}_A$ *is nonempty and bounded.*

Proof. (a) Defining $\varphi_{\bar X} := \varphi(\bar X,\cdot)$, we find that $\varphi^*_{\bar X} = \operatorname{cl}(\varphi^*_{\bar X}\,\square\, h^*)$, and therefore $\operatorname{ri}(\operatorname{dom}\psi^*_{\bar X}) = \operatorname{ri}(\operatorname{dom}\varphi^*_{\bar X}+\operatorname{dom} h^*) = \operatorname{ri}(\Omega_2(A,B)+\operatorname{dom} h^*)$; see Lemma 3.9(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) This follows readily from [16, Theorem 3.5, Proposition 3.9]. ∎

We point out that, under the PCQ, Lemma 3.11 shows that the objective functions $\psi(\bar X,\cdot)$ ($\bar X\in\operatorname{dom} p$) in the definition of p in (3.2) are *weakly coercive* [1, Definition 3.2.1] when proper; see [1, Theorem 3.2.1]. This tells us that the infimum in (3.2) is attained under the PCQ if it is finite [1, Proposition 3.2.2, Theorem 3.4.1], a fact that is restated (and alternatively derived) in Theorem 3.15. Under the SPCQ, the objective functions $\psi(\bar X,\cdot)$ ($\bar X\in\operatorname{dom} p$) are *level-bounded* (or *coercive*), in which case the $\operatorname{argmin}\psi(\bar X,\cdot)$ is nonempty and compact (and clearly convex). Finally, it was shown in Lemma 3.1(e) that p is closed, proper, convex under the BPCQ.

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

Lemma 3.12. *Let p be given by* (3.2). *Then the following hold.*

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

(b) *If* $\operatorname{int}(\operatorname{dom} h^*)\cap\operatorname{int}(-\Omega_2(A,B))\neq\emptyset$, *the PCQ and SPCQ are equivalent.*

Proof. (a) The first implication can be seen as follows: if the BPCQ holds, then $\operatorname{dom}\psi_{\bar X}\subset\operatorname{dom} h\cap\mathcal{K}_A$ is bounded (and nonempty exactly if $\bar X\in\operatorname{dom} p$). Therefore, $\psi_{\bar X}$ is level-bounded for all $\bar X\in\operatorname{dom} p$, i.e., $0\in\operatorname{int}(\operatorname{dom}\psi^*_{\bar X})$ ($\bar X\in\operatorname{dom} p$); see, e.g., [16, Theorem 11.8]. In view of Lemma 3.11(a) this implies that the SPCQ holds. The second implication is trivial.

(b) This follows directly from the definitions. ∎

We now provide characterizations for the CCQ.

LEMMA 3.13 (characterizations of the CCQ). *Let p be given by* (3.2). *Then*

$$\text{(i) } \operatorname{dom} h \cap \operatorname{int} \mathcal{K}_A \neq \emptyset \iff \text{(ii) } \textit{CCQ holds for } p \iff \text{(iii) } (-\mathcal{K}_A^\circ) \cap \operatorname{hzn} h^* = \{0\}.$$

Proof. The first equivalence was previously observed in (3.3). 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 $\operatorname{ri}(\operatorname{dom} h) \cap \operatorname{int} \mathcal{K}_A \neq \emptyset$ if and only if there does not exist a matrix $W \in \mathbb{S}^n$ such that

$$(h^*)^\infty(W) + \sigma_{\mathcal{K}_A}(-W) \leq 0 \quad \text{and} \quad (h^*)^\infty(-W) + \sigma_{\mathcal{K}_A}(W) > 0. \tag{3.20}$$

Since $\sigma_{\mathcal{K}_A}(-W) = \delta_{\mathcal{K}_A^\circ}(-W)$, the first of these conditions is equivalent to the condition $W \in (-\mathcal{K}_A^\circ) \cap \operatorname{hzn} h^*$. In particular, we can infer that $(-\mathcal{K}_A^\circ) \cap \operatorname{hzn} h^* = \{0\}$ gives the inconsistency of (3.20) and thus establishes (iii) $\implies$ (ii).

The second condition in (3.20) 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 $(-\mathcal{K}_A^\circ) \cap \operatorname{hzn} h^*$ satisfies (3.20). Therefore, the nonexistence of a W satisfying (3.20) implies that $(-\mathcal{K}_A^\circ) \cap \operatorname{hzn} h^* = \{0\}$, which proves the result. ∎

Note that for any proper, convex function f we always have $\operatorname{hzn} f \subset (\operatorname{dom} f)^\infty$, which, in view of Lemma 3.13, implies that the condition

$$(-\mathcal{K}_A^\circ) \cap (\operatorname{dom} h^*)^\infty = \{0\} \tag{3.21}$$

is stronger than the CCQ. However, (3.21) is not used in our study.

3.4. Infimal projection II. We return to our analysis of the infimal projection defining p in (3.2).

The following result shows that the two key conditions appearing in Proposition 3.7, $0 \in \operatorname{ri}(\operatorname{dom} q_{\bar{X}})$ and $\bar{X} \in \operatorname{ri}(\operatorname{dom} p)$, correspond nicely to the constraint qualifications studied in section 3.3.

COROLLARY 3.14. *Let p be defined by* (3.2), *let $\bar{X} \in \operatorname{dom} p$, and let $q_{\bar{X}}$ be defined by* (3.15). *Then,*

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

(b) *if the CCQ holds, $\bar{X} \in \operatorname{ri}(\operatorname{dom} p)$.*

Proof. (a) This follows immediately from Lemma 3.8 and the definition of the PCQ.

(b) Under the CCQ we have $\operatorname{dom} p = \mathbb{R}^{n \times m}$ (see the proof of Theorem 3.6(II)), and hence (b) follows. ∎

As a consequence of Corollary 3.14 and Proposition 3.7 we can add to the properties of p proven in Theorem 3.6.

THEOREM 3.15 (properties of p under the PCQ). *Let p, defined in* (3.2), *be such that the PCQ is satisfied and $\operatorname{dom} h \cap \mathcal{K}_A \neq \emptyset$ (i.e., $\operatorname{dom} p \neq \emptyset$). Let $q_{\bar{X}}$ be given by* (3.15). *Then the following hold:*

(a) $p \in \Gamma_0(\mathbb{R}^{n \times m})$,

(b) $\operatorname{argmin}_V \psi(\bar{X}, V) \neq \emptyset$ ($\bar{X} \in \operatorname{dom} p$) *(primal attainment),*

(c) $p(\bar{X}) = -q_{\bar{X}}(0)$ ($\bar{X} \in \operatorname{dom} p$) *(zero duality gap).*

Proof. Let $\bar{X} \in \operatorname{dom} p$. Under the PCQ, by Corollary 3.14, we have

$$0 \in \operatorname{ri}(\operatorname{dom} q_{\bar{X}}).$$

Hence, by Proposition 3.7(a), there is a $\bar{V} \in \mathbb{S}^n$ such that $p(\bar{X}) = \psi(\bar{X}, \bar{V})$, and so, by Proposition 3.7(c), p is l.s.c. at $\bar{X}$ with $p(\bar{x}) \in \mathbb{R}$. The discussion in [1, Page 153] tells us that p is, in fact, closed, proper, convex.

Finally, the equality $p(\bar{X}) = -q_{\bar{X}}(0)$ also follows from Proposition 3.7(a). ∎

Theorem 3.15 can be proven entirely without the shifted duality framework in Proposition 3.7 by using the linear projection $L : (X, V) \to X$ used implicitly throughout our study. It can be seen that $p = L\psi$ is a *linear image* in the sense described in [15, Page 38]. Then [15, Theorem 9.2] gives all statements from Proposition 3.15 after realizing that the constraint qualification in [15, Theorem 9.2], which reads

$$\psi(0, V) > 0 \quad \text{or} \quad \psi^\infty(0, -V) \leq 0 \quad (V \in \mathbb{S}^n) \tag{3.22}$$

since $\ker L = \{0\} \times \mathbb{S}^n$, is equivalent to the PCQ in this setting. However, we choose to derive Theorem 3.15 from the shifted duality scheme since this assists in the subdifferential analysis.

The next result follows readily from the foregoing study.

COROLLARY 3.16. *Let p be given by* (3.2) *and η_0 by* (3.12). *If the PCQ and CCQ are satisfied for p, then the following hold.*

(a) *The SCCQ holds and p is finite valued.*

(b) *Primal attainment: $p \in \Gamma_0(\mathbb{R}^{n\times m})$ is finite valued and for all $\bar{X} \in \mathbb{R}^{n\times m}$ there exists $\bar{V}$ such that $p(\bar{X}) = \psi(\bar{X}, \bar{V})$.*

(c) *Dual attainment: $p^* = \eta_0$ 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. (a) This follows readily from Lemma 3.11(a) and the definition of the SCCQ.

(b) By (a), the SCCQ holds, so the first statement follows from Theorem 3.6(c). The second statement is due to Theorem 3.15(b).

(c) Since the SCCQ holds (see (b)), Theorem 3.15(c) applies. ∎

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

Consequence\hypothesis	PCQ	SPCQ	BPCQ	CCQ	SCCQ	PCQ + CCQ
$p \in \Gamma_0 \vee p \equiv -\infty$	✓	✓	✓	✓	✓	✓
$p \in \Gamma_0$	✓	✓	✓		✓	✓
$p(\bar{X}) = -q_{\bar{X}}(0)$	✓	✓	✓	✓[1]	✓	✓
$\operatorname{argmin} \psi(\bar{X}, \cdot) \neq \emptyset$	✓	✓	✓			✓
$\operatorname{argmin} \psi(\bar{X}, \cdot)$ compact		✓	✓[2]			✓
$\operatorname{dom} p = \mathbb{R}^{n\times m}$				✓	✓	✓
$\underset{(\bar{Y},T)\in\Omega(A,B)}{\operatorname{argmin}} h^*(-T) \neq \emptyset$				✓	✓	✓

In view of Proposition 3.7(b) and Corollary 3.14 one might be inclined to think that using the CCQ instead of the pointwise condition $\bar{X} \in \operatorname{ri}(\operatorname{dom} p)$ is excessively strong. However, computing the relative interior of $\operatorname{dom} p$ without the CCQ is problematic (cf. the derivations in the proof of Theorem 3.6(c) under (CCQ)). Hence, we do not consider constraint qualifications weaker than the CCQ.

[1] $p(\bar{X}) \equiv -\infty$ is possible.

[2] The BPCQ also implies that $\operatorname{dom} \psi(\bar{X}, \cdot)$ is bounded.

We now turn our attention to subdifferentiation of p.

Proposition 3.17 (subdifferential of p). *Let p be given by* (3.2). *Then the following hold.*

(a) *Under the SCCQ,* $\mathrm{dom}\, p = \mathrm{dom}\, \partial p = \mathbb{R}^{n\times m}$ *and we have*

$$\partial p(\bar{X}) = \underset{Y}{\mathrm{argmax}} \left\{ \langle \bar{X},\, Y\rangle - \inf_{(Y,T)\in\Omega(A,B)} h^*(-T) \right\}, \tag{3.23}$$

which is nonempty and compact.

(b) *Under the PCQ* (3.23) *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 the PCQ and CCQ,* $\mathrm{dom}\, p = \mathrm{dom}\, \partial p = \mathbb{R}^{n\times m}$ *and we have*

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

which is nonempty and compact.

Proof. (a) Under the SCCQ, p is convex and finite valued (hence closed and proper). Therefore, $\mathrm{dom}\, p = \mathrm{dom}\, \partial p = \mathbb{R}^{n\times m}$ with $\partial p(X)$ compact for all $X \in \mathbb{R}^{n\times m}$. The representation (3.23) follows from [15, Theorem 23.5] and the fact that the closure for p^* can be dropped in the argmax problem.

(b) Under PCQ we also have that $p \in \Gamma_0$. Hence, the same reasoning as in (a) gives (3.23). We now prove the remainder: for the first identity notice that (see, e.g., [10, Chapter D, Corollary 4.5.3])

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

the latter argmin set being nonempty due to the arguments above. The "$\subset$" inclusion is thus clear. For the reverse inclusion, invoke the results in [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 identity follows from Proposition 3.7 in combination with Corollary 3.14 and recalling that $\psi^*(\bar{Y}, 0) = p^*(\bar{Y})$.

(c) Apply Corollary 3.5 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 the BPCQ, as the following example shows.

Example 3.18. 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 $\mathrm{dom}\, h \cap \mathcal{K}_A = \{\left(\begin{smallmatrix} v & 0 \\ 0 & 0 \end{smallmatrix}\right) \mid v \in [0,1]\}$ and $\mathrm{dom}\, h \cap \mathrm{int}\, \mathcal{K}_A = \emptyset$, so that the CCQ is violated but the BPCQ (and hence the (S)PCQ) holds. We find that

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

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

4. Infimal projection with a support function.

We now study the case where h is a support function:

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

where $\mathcal{V}$ is a given closed, convex subset of $\mathbb{S}^n$. Our first task is to interpret the constraint qualifications of section 3.3 when $h = \sigma_{\mathcal{V}}$. Here, and for the remainder of this section, 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 the CCQ for p in (4.1). *Moreover, if the CCQ holds, then the SCCQ holds if and only if*

$$\emptyset \neq \Xi(A,B) = \{Y \mid (Y,0) \in \Omega(A,B) + (\{0\} \times \mathcal{V})\}, \tag{4.5}$$

where $\Xi(A,B)$ is defined in (3.17).

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

$$\operatorname{pos}(\Omega_2(A,B) + \mathcal{V}) = \operatorname{span}(\Omega_2(A,B) + \mathcal{V}), \tag{4.6}$$

where $\Omega_2(A,B)$ is defined as in (3.16).

(c) *(BPCQ) The conditions*

$$\operatorname{bar} \mathcal{V} \cap \mathcal{K}_A \neq \emptyset \quad \textit{and} \quad \operatorname{cl}(\operatorname{bar} \mathcal{V}) \cap \mathcal{K}_A = \{0\}, \tag{4.7}$$

$$\operatorname{bar} \mathcal{V} \cap \mathcal{K}_A \ \textit{is nonempty and bounded}, \tag{4.8}$$

$$\operatorname{bar} \mathcal{V} \cap \mathcal{K}_A \neq \emptyset \quad \textit{and} \quad \mathcal{V}^\infty + \mathcal{K}_A^\circ = \mathbb{S}^n \tag{4.9}$$

are each equivalent to the BPCQ for p, and hence imply (4.6).

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) Equation (4.2) is condition (i) in Lemma 3.13 for $h = \sigma_{\mathcal{V}}$, while (4.3) is condition (iii). Employing the results in [3, section 3.3, Exercise 16] we have that (4.3) holds if and only if $\operatorname{cl}(\operatorname{bar} \mathcal{V} - \mathcal{K}_A) = \mathbb{S}^n$. The final statement follows from (3.17) in the definition of the SCCQ.

(b) This is an application of (2.1) and the definition of the PCQ.

(c) As the horizon cone of any cone is its closure, we see that (4.7) is exactly the BPCQ (for $h = \sigma_{\mathcal{V}}$), while the equivalence to (4.8) follows from Lemma 3.11(b). The equivalence of (4.9) to the former follows from the fact that (4.7) holds if and only if $\operatorname{cl}(\mathcal{V}^\infty + \mathcal{K}_A^\circ) = \mathbb{S}^n$ (see [3, section 3.3, Exercise 16]), where the closure can be dropped by interpreting [15, Theorem 6.3] accordingly. ∎

The additivity of support functions tells us that

$$p(X) = \inf_{V \in \mathbb{S}^n} \sigma_\Sigma(X, V) \quad (X \in \mathbb{R}^{n\times m}), \tag{4.10}$$

where

$$\Sigma := \Omega(A, B) + \{0\} \times \mathcal{V} \subset \mathbb{E}. \tag{4.11}$$

In particular, this implies that $p(\lambda X) = \lambda p(X)$ for all $\lambda > 0$ and $p(X_1 + X_2) \leq p(X_1) + p(X_2)$. Hence, if p is proper, it is a support function. In addition, by (3.17), $\Xi(A, B) = \{Y \mid (Y, 0) \in \Sigma\}$ is the set featured in (3.17), (3.18), and (4.5).

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

(a) $p \in \Gamma_0(\mathbb{R}^{n\times m})$ *(i.e., $p = p^{**}$) under condition* (4.6), *and hence under any of the conditions* (4.7)–(4.9). *Moreover, this is also true under any condition in* (4.2)–(4.4) *if, in addition,* (4.5) *or* (4.6) *holds, in which case p is finite valued.*

(b) $p^* = \delta_{\mathrm{cl}\,\Sigma}(\cdot, 0)$, *where the closure is superfluous (i.e., Σ is closed) under any of the conditions* (4.2)–(4.4), *in which case $p^* = \delta_{\Xi(A,B)}$.*

(c) *If any of* (4.2)–(4.4) *hold, then $p \equiv -\infty$ or $p = \sigma_{\Xi(A,B)}$ is finite valued. The latter is the case if and only if* (4.5) *holds, which is valid under* (4.6).

Proof. (a) The first statement follows from Lemma 4.1 and Theorem 3.15. The second uses Lemma 4.1, Theorem 3.6(c), and Corollary 3.16.

(b) By [16, Exercise 3.12] and [6, Proposition 10], 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). Moreover, $p^* = \sigma_\Sigma^*(\cdot, 0) = \delta_{\mathrm{cl}\,\Sigma}(\cdot, 0)$; see [16, Proposition 11.23(c)].

(c) This follows from (a), (b), and Theorem 3.6(II), as well as Corollary 3.16. ∎

4.1. The case when $B = 0$. We now consider the case when $B = 0$. Recall from [6, Theorem 11] that this implies $\sigma_{\Omega(A,0)}$ 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 4.3 (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*

$$\gamma(x) \leq \gamma(y) \text{ whenever } 0 \preceq_K x \preceq_K y.$$

The following lemma provides a characterization of cone-compatible gauges and provides a very useful tool for determining whether a gauge is compatible with a given cone.

Lemma 4.4 (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 $K \cap (y - K) \subset C$ $(y \in K \cap C)$.*

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 $\gamma_C(x) \leq \gamma_C(y) \leq 1$, and, consequently, $K \cap (y - K) \subset C$.

Next suppose $K \cap (y - K) \subset C$ for all $y \in K \cap C$, 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 4.5 (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 $B = 0$. Under any of the conditions* (4.2)–(4.4) *we have the following.*

(a) $p^* = \delta_{\Xi(A,0)}$*, where*

$$\Xi(A,0) = \left\{ Y \;\middle|\; AY = 0,\ \exists W \in \mathcal{V} : AW = 0,\ \frac{1}{2}YY^T \preceq W \right\}.$$

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

$$p^*(Y) = \delta_{\{Y \mid AY=0,\ \gamma_{\mathcal{V}}(\frac{1}{2}YY^T) \le 1\}}(Y) = \delta_{(-\mathcal{K}_A^\circ)\cap\mathcal{V}}\left(\frac{1}{2}YY^T\right). \tag{4.12}$$

Proof. (a) This follows from Proposition 4.2, (4.5) with $B = 0$, and using the representation of $\mathcal{K}_A$ in Proposition 2.1.

(b) First observe that $-\mathcal{K}_A^\circ = \{W \in \mathbb{S}^n_+ \mid \operatorname{rge} W \subset \ker A\}$; see Proposition 2.1(b). Recall that $\operatorname{rge} Y = \operatorname{rge} YY^T$ $(Y \in \mathbb{R}^{n\times m})$ and, since $0 \in \mathcal{V}$, $V \in \mathcal{V}$ if and only if $\gamma_{\mathcal{V}}(V) \le 1$. Exploiting these facts and the compatibility hypothesis, we see that

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

Conversely, we have $\frac{1}{2}YY^T \in (-\mathcal{K}_A^\circ) \cap \mathcal{V} \iff AY = 0,\ Y \in \mathcal{K}_A$, and $\frac{1}{2}YY^T \in \mathcal{V}$. Taking $W = \frac{1}{2}YY^T$, we see that $Y \in \Xi(A,0)$. Therefore, (b) follows from (a). ∎

When the support function h is taken to be a linear functional, we obtain the following remarkable result. Here $\|\cdot\|_*$ denotes the nuclear norm.[3]

Corollary 4.6 (h linear). *Let $p : \mathbb{R}^{n\times m} \to \overline{\mathbb{R}}$ be defined by*

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

[3]For a matrix T the nuclear norm $\|T\|_*$ is the sum of its singular values.

for some $\bar{U} \in \mathbb{S}^n_+ \cap \mathrm{Ker}_n A$ *and set* $C(\bar{U}) := \{Y \in \mathbb{R}^{n\times m} \mid \frac{1}{2}YY^T \preceq \bar{U}\}$. *Then,*

(a) $p^* = \delta_{C(\bar{U})}$ *is closed, proper, convex;*

(b) $p = \sigma_{C(\bar{U})} = \gamma_{C(\bar{U})^\circ}$ *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* $A = 0$, $p = \sigma_{C(\bar{U})} = \|L^T(\cdot)\|_*$, *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* $\bar{U} \succ 0$, *then* $C(\bar{U})$ *and* $C(\bar{U})^\circ$ *are compact, convex, symmetric*[4] *with* 0 *in their interior, and thus* $\mathrm{pos}\, C(\bar{U}) = \mathrm{pos}\, C(\bar{U})^\circ = \mathbb{S}^n$.

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, the CCQ is trivially satisfied (cf. (4.2)–(4.4)). Note that $0 \in C(\bar{U}) \neq \emptyset$. Given $Y \in C(\bar{U})$, we must have $\mathrm{rge}\, Y \subset \ker A$ since otherwise there is a nonzero $z \in (\ker A)^\perp$ with $Y^T z \neq 0$ yielding $0 < \|Y^T z\|_2^2 \leq 2z^T\bar{U}z = 0$. Consequently, $C(\bar{U}) = \{Y \in \mathbb{R}^{n\times m} \mid AY = 0,\ \frac{1}{2}YY^T - \bar{U} \in \mathcal{K}_A^\circ\} = \Xi(A, 0) \neq \emptyset$, and the result follows from Proposition 4.2(b).

(b) This follows from [15, Theorem 14.5], part (a) above, and the fact that $0 \in C(\bar{U})$.

(c) Consider the case when $\bar{U} = \frac{1}{2}I$: by part (a), we have $p^* = \delta_{\{Y \mid YY^T \preceq I\}}$. Observe that $\{Y \mid YY^T \preceq I\} = \{Y \mid \|Y\|_2 \leq 1\} =: \mathbb{B}_\Lambda$ 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 $C(\bar{U}) = \{Y \mid L^{-1}Y \in C(\frac{1}{2}I)\}$, and therefore

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

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

(d) This follows from (c) using [15, Theorem 15.2]. ∎

We point out that Corollary 4.6 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 the indicator $h := \delta_\mathcal{V}$ for some nonempty, closed, and convex set $\mathcal{V} \in \mathbb{S}^n$:

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

We begin by interpreting the constraint qualifications from section 3.3. Here, and for the remainder of this section, $h = \delta_\mathcal{V}$ and so $\mathrm{dom}\, h = \mathcal{V}$ and $\mathrm{dom}\, h^* = \mathrm{bar}\, \mathcal{V}$.

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

[4] We say the set $S \subset \mathcal{E}$ is symmetric if $S = -S$.

(a) *CCQ: the conditions*

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

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

are each equivalent to the CCQ for p. Moreover, if the CCQ holds, then the SCCQ holds if and only if

$$\emptyset \neq \Xi(A, B) = \{Y \in \mathbb{R}^{n\times m} \mid (Y, 0) \in \Omega(A, B) + (\{0\} \times \operatorname{bar} \mathcal{V})\}. \tag{5.4}$$

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

$$\operatorname{pos}(\Omega_2(A, B)) + \operatorname{bar} \mathcal{V} = \operatorname{span}(\Omega_2(A, B) + \operatorname{bar} \mathcal{V}). \tag{5.5}$$

(c) *BPCQ: the conditions*

$$\mathcal{V} \cap \mathcal{K}_A \neq \emptyset \quad \textit{and} \quad \mathcal{V}^\infty \cap \mathcal{K}_A = \{0\}, \tag{5.6}$$

$$\mathcal{V} \cap \mathcal{K}_A \neq \emptyset \textit{ is bounded}, \tag{5.7}$$

$$\mathcal{V} \cap \mathcal{K}_A \neq \emptyset \quad \textit{and} \quad \operatorname{bar} \mathcal{V} + \mathcal{K}_A^\circ = \mathbb{S}^n \tag{5.8}$$

are each equivalent to the BPCQ for p, and hence imply (5.5).

Proof. (a) First, observe that, with $h = \delta_\mathcal{V}$, Lemma 3.13(i) is exactly (5.2). By the same lemma, this is equivalent to $\operatorname{hzn} \sigma_\mathcal{V} \cap (-\mathcal{K}_A^\circ) = \{0\}$. Moreover, since $\sigma_\mathcal{V} = \sigma_\mathcal{V}^\infty$, we have $\operatorname{hzn} \sigma_\mathcal{V} = \{V \mid \sigma_\mathcal{V}(V) \leq 0\} = (\text{cone } \mathcal{V})^\circ$. Invoking the results in [3, section 3.3, Exercise 16(a)] implies that $\operatorname{hzn} \sigma_\mathcal{V} \cap (-\mathcal{K}_A^\circ) = \{0\}$ if and only if $\operatorname{cl}(\overline{\text{cone}}\, \mathcal{V} - \mathcal{K}_A) = \mathbb{S}^n$, 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 the PCQ holds for p if and only if

$$\operatorname{pos}(\Omega_2(A, B)) + \operatorname{bar} V = \operatorname{pos}(\Omega_2(A, B) + \operatorname{bar} V) = \operatorname{span}(\Omega_2(A, B) + \operatorname{bar} \mathcal{V}).$$

(c) The equivalence of the BPCQ, (5.6), and (5.7) is clear. Since $\mathcal{V}^\infty$ and $\operatorname{cl}(\operatorname{bar} \mathcal{V})$ are paired in polarity (see (2.3)), [3, section 3.3, Exercise 16(a)] implies that $\mathcal{V}^\infty \cap \mathcal{K}_A = \{0\}$ if and only if $\operatorname{cl}(\operatorname{bar} \mathcal{V} + \mathcal{K}_A^\circ) = \mathbb{S}^n$, where the closure in the latter statement can be dropped as in part (a) above. This establishes all equivalences. □

The following result provides sufficient conditions for p being closed, proper, convex when h is an indicator function.

Corollary 5.2. *Let p be given by* (5.1). *Then* $p \in \Gamma_0(\mathbb{R}^{n\times m})$ *under any of the following conditions:*

(i) (5.4) *holds along with either* (5.2) *or* (5.3)*;*
(ii) (5.5) *holds;*
(iii) *any one of* (5.6)–(5.8) *holds.*

Proof. This follows from Lemma 5.1, Theorem 3.6(c), and Theorem 3.15, respectively. □

The case when $A = 0$ and $B = 0$ is of particular interest in applications to the variational Gram functions in section 5.2.

Corollary 5.3. *Let p be given as in* (5.1) *with* $A = 0$ *and* $B = 0$ *so that* $\mathcal{K}_A = \mathbb{S}^n_+$ *and* $\mathcal{K}_A^\circ = \mathbb{S}^n_-$. *Assume that* $\mathcal{V} \cap \mathbb{S}^n_+ \neq \emptyset$. *Then*

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

Moreover, $p \in \Gamma_0(\mathbb{R}^{n\times m})$ under any of following conditions.

(i) *SCCQ: $\{Y \in \mathbb{R}^{n\times m} \mid \exists\, T \in \mathrm{bar}\,\mathcal{V} : \frac{1}{2}YY^T \preceq T\} \neq \emptyset$ and $\mathcal{V} \cap \mathbb{S}^n_{++} \neq \emptyset$.*

(ii) *PCQ: $\mathbb{S}^n_- + \mathrm{bar}\,\mathcal{V} = \mathbb{S}^n$.*

(iii) *(B/S)PCQ: $\emptyset \neq \mathcal{V} \cap \mathbb{S}^n_+$ is bounded.*

Proof. First note that

$$\Xi(0,0) = \left\{ Y \in \mathbb{R}^{n\times m} \;\middle|\; \exists\, T \in \mathrm{bar}\,\mathcal{V} : \frac{1}{2}YY^T \preceq T \right\} \quad \text{and} \quad \Omega_2(0,0) = \mathbb{S}^n_- = \mathcal{K}_A^\circ.$$

The first statement now follows from Lemma 5.1 and the definitions of the PCQ and SPCQ, respectively, since the span of a set with interior is the whole space. The remaining implications follow from Corollary 5.2 and Lemma 5.1. ∎

We compute the conjugate p^* directly using techniques from [5, Theorem 3.2].

THEOREM 5.4 (infimal projection with an indicator function). *Let p be given by (5.1). Assume that*

$$\emptyset \neq \mathrm{dom}\,(\varphi + \delta_{\mathcal{V}}) = \left\{ (X,V) \in \mathbb{E} \;\middle|\; V \in \mathcal{V} \cap \mathcal{K}_A \text{ and } \mathrm{rge}\begin{pmatrix} X \\ B \end{pmatrix} \subset \mathrm{rge}\, M(V) \right\}. \tag{5.9}$$

Then $p^* : \mathbb{R}^{n\times m} \to \overline{\mathbb{R}}$ is given by

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

In particular, for $A = 0$ and $B = 0$ we obtain $p^*(Y) = \frac{1}{2}\sigma_{\mathcal{V}\cap\mathbb{S}^n_+}(YY^T)$.

Proof. By (2.4) and our assumption that $\emptyset \neq \mathrm{dom}\,(\varphi + \delta_{\mathcal{V}})$, we have

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

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

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

where the final equality follows since

$$\delta_{\{y\,|\,Ay=b_i\}}(y_i) = \sup_{w_i} \langle w_i,\, Ay_i - b_i\rangle \quad (i = 1, \ldots, m).$$

By hypothesis, $\operatorname{rge} B \subset \operatorname{rge} A$, and so, by [5, Theorem 3.2],

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

Therefore, when $AY = B$, we have

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

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

COROLLARY 5.5. *Let p be given by* (5.1). *If the SCCQ holds, i.e.,*

$$\mathcal{V} \cap \operatorname{int} \mathcal{K}_A \neq \emptyset \quad \textit{and} \quad \{Y \in \mathbb{R}^{n\times m} \mid (Y, 0) \in \Omega(A, B) + (\{0\} + \operatorname{bar} \mathcal{V})\} \neq \emptyset,$$

then

$$\partial p(\bar{X}) = \operatorname*{argmax}_Y \left\{\langle \bar{X},\, Y\rangle - \inf_{(Y,T)\in\Omega(A,B)} \sigma_{\mathcal{V}}(-T)\right\}$$

is nonempty and compact for all $\bar{X} \in \mathbb{R}^{n\times m}$. Alternatively, if $\mathcal{V} \cap \operatorname{int} \mathcal{K}_A \neq \emptyset$ (the CCQ) and $\operatorname{pos} \Omega_2(A, B) + \operatorname{bar} \mathcal{V} = \operatorname{span}(\Omega_2(A, B) + \operatorname{bar} \mathcal{V})$ (the PCQ) hold, then

$$\partial p(\bar{X}) = \{\bar{Y} \mid \exists \bar{V}, \bar{T} : -\bar{T} \in N_{\mathcal{V}}(\bar{V}),\ (\bar{Y}, \bar{T}) \in \partial\varphi(\bar{X}, \bar{V})\}$$

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

Proof. This follows from Proposition 3.17 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 5.8. We start with a technical lemma.

LEMMA 5.6. *Let $C, K \subset \mathbb{E}$ be nonempty and convex with K 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 result in this section is concerned with a representation of the conjugate p^* under the standing assumptions.

Corollary 5.7 (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*

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

Assume that

$$\emptyset \neq \left\{ (X,V) \in \mathbb{E} \,\middle|\, V \in \mathcal{V} \cap \mathcal{K}_A \text{ and } \operatorname{rge} \begin{pmatrix} X \\ 0 \end{pmatrix} \subset \operatorname{rge} M(V) \right\}.$$

Then the following hold.

(a) *We have*

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

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 $\mathcal{K}_A^\circ \cap -(\operatorname{cone} \mathcal{V})^\circ = \{0\}$*), *then*

$$p^*(Y) = \frac{1}{2}\gamma_{(\mathcal{V}^\circ\cap\mathcal{S})+\mathcal{K}_A^\circ}(YY^T), \tag{5.10}$$

where $\operatorname{dom} p^* = \{Y \mid YY^T \in \operatorname{cone}(\mathcal{V}^\circ \cap \mathcal{S}) + \mathcal{K}_A^\circ\}$.

Proof. (a) By Theorem 5.4, we have

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

Here the first equality uses Theorem 5.4, the second equality follows from the fact that $\operatorname{rge} Y = \operatorname{rge} YY^T$, the third can be seen from [16, Example 7.4], 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), which gives (5.10). ∎

Our final goal is to show that, under the standing assumption in this section, p is a squared gauge. Here we denote by $\mathbb{B}_F$ the (closed) unit ball in the Frobenius norm.

Corollary 5.8 (gauge case II). *Let p be as in Theorem 5.4 with $0 \in \mathcal{V}$ and $B = 0$, and assume that* (5.9) *holds. Let $P \in \mathbb{R}^{n\times n}$ be the orthogonal projector on* $\ker A$ *and define the (closed, convex) sets*

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

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

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

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

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

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

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

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

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

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 second term can be written as

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

Using the definition of $\mathcal{F}$ and the convention $(+\infty)^2 = +\infty$, we can rewrite this equivalence as $\frac{1}{2}\sigma_{\mathcal{F}}(Y)^2 + \delta_{\mathcal{U}}(Y) = \frac{1}{2}[\sigma_{\mathcal{F}}(Y) + \delta_{\mathcal{U}}(Y)]^2$. All in all, using the latter, [16, Examples 11.4, 11.19], and the polar cone calculus from, e.g., [3, Page 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 gives the representation for p^*; the representation for p follows from [15, Corollary 15.3.1]. ∎

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

$$\Phi_{\mathcal{V}} : \mathbb{R}^{n\times m} \to \overline{\mathbb{R}}, \quad \Phi_{\mathcal{V}}(Y) := \frac{1}{2}\sigma_{\mathcal{V}\cap\mathbb{S}^n_+}(YY^T). \tag{5.11}$$

These functions are called *variational Gram functions (VGFs)* and were introduced by Jalali, Fazel, and Xiao [14]. They have received attention from the machine learning community due to their orthogonality promoting properties when used as penalty functions (cf. [14]).

Note that the definition (5.11) explicitly intersects $\mathcal{V}$ with the positive semidefinite cone $\mathbb{S}^n_+$, while Jalali, Fazel, and Xiao [14] employ the standing assumption that $\Phi_{\mathcal{V}} = \Phi_{\mathcal{V}\cap\mathbb{S}^n_+}$. These (equivalent) conventions guarantee that $\Phi_{\mathcal{V}}$ is convex. We also scale by $1/2$ since $\Phi_{\mathcal{V}}$ is positively homogeneous of degree 2.

As an immediate consequence of Theorem 5.4, $\Phi_{\mathcal{V}} = p^*$, where p is defined in (5.1) with $A = 0$, $B = 0$, and $\mathcal{V}\cap\mathbb{S}^n_+ \neq \emptyset$. In addition, the constraint qualifications dramatically simplify in this case. We have already seen in Corollary 5.3 that the PCQ, SPCQ, and BPCQ are all equivalent for VGFs. We now observe that the CCQ and SCCQ are also equivalent.

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

LEMMA 5.9 (CCQ = SCCQ for VGFs). *Let $\Phi_{\mathcal{V}}$ be given by* (5.11) *with $\mathcal{V} \subset \mathbb{S}^n$. Then the condition $\mathcal{V} \cap \mathbb{S}^n_+ \neq \emptyset$ is equivalent to* (5.9), *and* (5.4) *is satisfied with $\Phi_{\mathcal{V}} = p^*$, where $A = 0$, $B = 0$, and p is as defined in* (5.1). *In particular, the CCQ and SCCQ are equivalent, where the CCQ is given by $\mathcal{V} \cap \mathbb{S}^n_{++} \neq \emptyset$.*

Proof. First, note that $0 \in \Xi(0,0) = \{Y \mid \exists W \in \operatorname{bar} \mathcal{V} : \frac{1}{2}YY^T \preceq W\}$ since $0 \in \operatorname{bar} \mathcal{V}$. The relationship between $\Phi_{\mathcal{V}}$ and p is given in Theorem 5.4. ☐

Lemma 5.9 and the results of the previous section allow us to refine [14, Proposition 4].

PROPOSITION 5.10 (conjugate of VGFs and VGFs as squared gauges). *Let $\Phi_{\mathcal{V}}$ be given by* (5.11). *Under either of the assumptions*

(i) *$\mathcal{V} \cap \mathbb{S}^n_{++} \neq \emptyset$ (for the CCQ),*

(ii) *$\mathcal{V} \cap \mathbb{S}^n_+ \neq \emptyset$ is bounded (or, equivalently, $\mathbb{S}^n_- + \operatorname{bar} \mathcal{V} = \mathbb{S}^n$) (for the PCQ),*

we have

$$\Phi^*_{\mathcal{V}}(X) = \inf_V \sigma_{\Omega(0,0)}(X,V) + \delta_{\mathcal{V}}(V) = \frac{1}{2} \inf_{\substack{V \in \mathcal{V} \cap \mathbb{S}^n_+ : \\ \operatorname{rge} X \subset \operatorname{rge} V}} \operatorname{tr}\left(X^T V^\dagger X\right) \quad (X \in \mathbb{R}^{n \times m}).$$

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

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

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

Proof. Lemma 5.9 tells us that assumption (i) is equivalent to the SCCQ, and Corollary 5.3 tells us that assumption (ii) is equivalent to the BPCQ. Hence, by Theorem 5.4, either assumption (i) or (ii) implies that $\Phi^*_{\mathcal{V}} = p^{**} = p$. The remainder of the proof now follows from the definition of p, (2.5), and Corollary 5.8. ☐

Next consider the subdifferential of a VGF when defined by (5.11). Although a VGF is always convex, we take the *convex-composite* perspective (see, e.g., [7]), since a VGF is simply the composition of a closed, proper, convex function $\sigma_{\mathcal{V} \cap \mathbb{S}^n_+}$ and a nonlinear map $H : Y \mapsto YY^T$. The *basic constraint qualification* for the composition $\Phi_{\mathcal{V}} = \frac{1}{2}\sigma_{\mathcal{V} \cap \mathbb{S}^n_+} \circ H$ at a point $\bar{Y} \in \operatorname{dom} \Phi_{\mathcal{V}}$ is given by

$$\text{(BCQ)} \qquad N_{\operatorname{dom} \sigma_{\mathcal{V} \cap \mathbb{S}^n_+}}(\bar{Y}\bar{Y}^T) \cap (\operatorname{Ker}_n \bar{Y}^T) = \{0\}.$$

It is well known that this condition is essential for a full subdifferential calculus of convex-composite functions [16]. We now show that this condition is intimately linked to condition (ii) in Corollary 5.3.

LEMMA 5.11 (BPCQ = PCQ = BCQ for VGFs). *Let $\Phi_{\mathcal{V}}$ be as in* (5.11) *and assume that $\mathbb{S}^n_+ \cap \mathcal{V} \neq \emptyset$. Then the following are equivalent.*

(i) *There exists $\bar{Y} \in \operatorname{dom} \Phi_{\mathcal{V}}$ such that the BCQ holds.*

(ii) *(B)PCQ: $\mathcal{V}^\infty \cap \mathbb{S}^n_+ = \{0\}$ (or, equivalently, $\mathcal{V} \cap \mathbb{S}^n_+$ is bounded).*

(iii) *The BCQ holds at every $\bar{Y} \in \operatorname{dom} \Phi_{\mathcal{V}}$.*

Proof. (i) $\Longrightarrow$ (ii): let $\bar{V} \in \mathbb{S}^n_+ \cap \mathcal{V}$ and assume (ii) is violated, i.e., there exists $0 \neq W \in (\mathcal{V} \cap \mathbb{S}^n_+)^\infty = \mathcal{V}^\infty \cap \mathbb{S}^n_+$. By (2.2), we have

$$V_t := \bar{V} + tW \in \mathcal{V} \cap \mathbb{S}^n_+ \quad (t > 0). \tag{5.12}$$

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

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

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

$$N_{\operatorname{dom} \sigma_{\mathcal{V} \cap \mathbb{S}^n_+}}(Z) = (\mathcal{V} \cap \mathbb{S}^n_+)^\infty$$

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

(ii) $\implies$ (iii): if $\mathcal{V} \cap \mathbb{S}^n_+$ is bounded, then $\operatorname{dom} \sigma_{\mathcal{V} \cap \mathbb{S}^n_+} = \mathbb{S}^n$, and so, for every $\bar{Y} \in \operatorname{dom} \Phi_{\mathcal{V}}$, $N_{\operatorname{dom} \sigma_{\mathcal{V} \cap \mathbb{S}^n_+}}(\bar{Y}\bar{Y}^T) = \{0\}$, giving the desired implication.

(iii) $\implies$ (i): this is obvious. ∎

We now derive the formula for the subdifferential of the VGF from (5.11).

PROPOSITION 5.12. *Let* $\Phi_{\mathcal{V}}$ *be given by* (5.11). *Then*

$$\partial \Phi_{\mathcal{V}}(\bar{Y}) \supset \{\bar{V}\bar{Y} \mid \bar{V} \in \mathcal{V} \cap \mathbb{S}^n_+ : \langle \bar{V}, \bar{Y}\bar{Y}^T \rangle = \Phi_{\mathcal{V}}(\bar{Y})\} \quad (\bar{Y} \in \operatorname{dom} \Phi_{\mathcal{V}}).$$

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

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 (subdifferential calculus failure for the 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) fails. We have

$$\sigma_{\mathcal{V} \cap \mathbb{S}^n_+}(W) = \sup_{\alpha \geq 0} \alpha \operatorname{tr}(W) = \delta_{\{U \in \mathbb{S}^n \mid \operatorname{tr}(U) \leq 0\}}(W) \quad (W \in \mathbb{S}^n). \tag{5.13}$$

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

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

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

Jalali, Fazel, and Xiao employ great effort to compute the subdifferential of the conjugate of a (convex) VGF (see the proof of [14, Proposition 3.8]). However, a slightly refined version of [14, Proposition 3.8] follows immediately from our analysis.

PROPOSITION 5.14 (subdifferential of $\Phi_{\mathcal{V}}^*$). *Let $\Phi_{\mathcal{V}}$ be given by* (5.11).

(a) *((S)CCQ) If $\mathcal{V} \cap \mathbb{S}^n_{++} \neq \emptyset$, then* $\operatorname{dom} \partial \Phi_{\mathcal{V}}^* = \operatorname{dom} \Phi_{\mathcal{V}}^*$ *and*

$$\partial \Phi_{\mathcal{V}}^*(\bar{X}) = \operatorname*{argmax}_{Y} \left\{ \langle \bar{X}, Y \rangle - \inf_{\frac{1}{2}YY^T \preceq T} \sigma_{\mathcal{V} \cap \mathbb{S}^n_+}(T) \right\}.$$

(b) *((B)PCQ) If the set $\mathcal{V} \cap \mathbb{S}^n_+$ is nonempty and bounded, then* $\operatorname{dom} \partial \Phi_{\mathcal{V}}^* = \operatorname{dom} \Phi_{\mathcal{V}}^*$ *and we have*

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

for all $\bar{X} \in \operatorname{dom} \Phi_{\mathcal{V}}^*$.

Proof. (a) By Lemma 5.9, PCQ = SCCQ and $\Phi_{\mathcal{V}} = p^*$. The subdifferential formula follows from Proposition 3.17(a).

(b) The fact that $\operatorname{dom} \partial \Phi_{\mathcal{V}}^* = \operatorname{dom} \Phi_{\mathcal{V}}^*$ is due to the fact that the latter is a subspace, and hence relatively open; cf. Lemma 3.1(c). The remainder of the proof follows from Lemma 5.9 and Proposition 3.17(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, Example 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 denoting the columns of X, and the final from (5.11). For the Schatten norms, where $k = \min\{n, m\}$, we have $\mathbb{B}^\circ_{s,k} = \mathbb{B}_{\hat{s},k}$, with $\hat{s}$ satisfying $\frac{1}{s} + \frac{1}{\hat{s}} = 1$; see [11]. For other values of k, the representation of $\mathbb{B}^\circ_{s,k}$ can be significantly more complicated; see, e.g., [8].

6. Final remarks. 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 were given for closedness and properness as well as representations of both the conjugate and the subdifferential of the infimal projections under the associated essential constraint qualifications. The general results were applied to the cases when h is a support or an indicator function of a closed, convex set in $\mathbb{S}^n$. These results revealed close connections to a range of important convex functions on $\mathbb{R}^{n\times m}$. In particular, the infimal projection with linear functionals yielded smoothing variational representations for the family of scaled nuclear norms, while the infimal projection with an indicator is often a squared gauge. As a special case, we showed that the conjugate of the infimal projection coincides with a variational Gram function of the underlying set. Hence, the variational calculus for VGFs follows easily as a consequence of our general study. In all cases, the infimal projection offers new smoothing approaches to a range of nonsmooth optimization problems on $\mathbb{R}^{n\times m}$ using the representation (1.4).

Appendix A. In what follows we use the *direct sum* of functions $f_i : \mathcal{E} \to \mathbb{R}\cup\{+\infty\}$ $(i = 1,\dots,m)$, which is defined by

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

THEOREM A.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) *The conjugate of f is given by $f^* = \mathrm{cl}\,(f_1^*\Box f_2^*\Box\cdots\Box f_m^*)$. Under the condition*

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

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* (A.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*

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

(d) *Under* (A.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 [15, Theorem 23.5], $\bar x \in$

$\partial f_i^*(z^i)$ $(i = 1,\dots,m)$. Hence, by [15, Theorems 23.8, 23.9] and [2, Proposition 16.8] we obtain

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

Therefore, $(z^1,\dots,z^m) \in \mathcal{T}(\bar z)$, and we have

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

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

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

so that $0 = \sum_{i=1}^m (f_i^*(z^i) + f_i(\bar x) - \langle z^i, \bar x\rangle)$. By the Fenchel–Young inequality, $f_i^*(z^i) + f_i(\bar x) - \langle z^i, \bar x\rangle \geq 0$ $(i = 1,\dots,m)$, and hence the equality must hold for each $i = 1,\dots,m$, or, equivalently, $z^i \in \partial f_i(\bar x)$ $(i = 1,\dots,m)$. 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 $(z^1,\dots,z^m) \in \mathcal{T}(\bar z)$, we have $z^i \in \partial f_i(\bar x)$ $(i = 1,\dots,m)$. Hence,

$$\partial f(\bar x) \subset \{\bar z \mid (z^1,\dots,z^m) \in \mathcal{T}(\bar z),\ z^i \in \partial f_i(\bar x)\ (i = 1,\dots,m)\}.$$

The reverse inclusion follows from the first equivalence.

(d) By (a), $f^* = f_1^* \,\square\, f_2^* \,\square \cdots \square\, f_m^* \in \Gamma_0(\mathcal{E})$ and $\mathcal{T}(z) \neq \emptyset$ for all $z \in \operatorname{dom} f^*$.

Let us first suppose that $\bar z \in \operatorname{dom} \partial f^* \subset \operatorname{dom} f^*$, and 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 $\bar x \in \{\bigcap_{i=1}^m \partial f_i^*(z^i) \mid \bar z = \sum_{i=1}^m z^i\}$. 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 $\bar x \in \{\bigcap_{i=1}^m \partial f_i^*(z^i) \mid \bar z = \sum_{i=1}^m z^i\}$. 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 $(z^1,\dots,z^m) \in \mathcal{T}(\bar z) \neq \emptyset$. 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. ∎

Proposition A.2 (partial conjugates). *Let $f \in \Gamma(\mathcal{E}_1 \times \mathcal{E}_2)$ and $\bar x \in \mathcal{E}_1$ be such that $\bar g := f(\bar x, \cdot)$ is proper and $\bar x \in \operatorname{ri} L(\operatorname{dom} f)$, where $L : (x, v) \mapsto x$. Then*

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

Proof. By [15, Theorem 6.6], $\operatorname{ri} L(\operatorname{dom} f) = L(\operatorname{ri}\operatorname{dom} f)$, so the hypothesis implies the existence of a $\bar{w} \in \mathcal{E}_2$ such that $(\bar{x}, \bar{w}) \in \operatorname{ri}\operatorname{dom} f$. By [15, Theorem 16.4],

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

where the closure can be dropped if $\operatorname{ri}(\operatorname{dom} f) \cap \operatorname{ri}(\operatorname{dom}\delta_{\{\bar{x}\}\times\mathcal{E}_2}) \neq \emptyset$. But this intersection is nonempty by hypothesis since $(\bar{x}, \bar{w}) \in \{\bar{x}\}\times\mathcal{E}_2 = \operatorname{ri}(\operatorname{dom}\delta_{\{\bar{x}\}\times\mathcal{E}_2})$. Hence,

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

REFERENCES

[1] A. Auslender and M. Teboulle, *Asymptotic Cones and Functions in Optimization and Variational Inequalities,* Monogr. Math., Springer, New York, 2003.

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

[3] J. M. Borwein and A. S. Lewis, *Convex Analysis and Nonlinear Optimization. Theory and Examples*, CMS Books Math., Springer, New York, 2000.

[4] S. Boyd and L. Vandenbergh, *Convex Optimization,* Cambridge University Press, Cambridge, 2004.

[5] J. V. Burke and T. Hoheisel, *Matrix support functionals for inverse problems, regularization, and learning*, SIAM J. Optim., 25 (2015), pp. 1135–1159.

[6] J. V. Burke, Y. Gao, and T. Hoheisel, *Convex geometry of the generalized matrix-fractional function*, SIAM J. Optim., 28 (2018), pp. 2189–2200.

[7] J. V. Burke and R. A. Poliquin, *Optimality conditions for non-finite valued convex composite functions*, Math. Program., 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*, preprint, https://arxiv.org/pdf/1403.5901v2, 2015.

[9] J. Dattorro, *Convex Optimization & Euclidean Distance Geometry*, Version 2014.04.08, M$\varepsilon\beta oo$ Publishing, Palo Alto, CA, 2005.

[10] J.-B. Hiriart-Urrruty and C. Lemaréchal, *Fundamentals of Convex Analysis,* Grundlehren Text Editions, Springer, Berlin, 2001.

[11] R. A. Horn and C. R. Johnson, *Matrix Analysis,* Cambridge University Press, New York, 1985.

[12] R.A. Horn and C. R. Johnson, *Topics in Matrix Analysis,* Cambridge University Press, New York, 1991.

[13] C.-J. Hsieh and P. Olsen, *Nuclear norm minimization via active subspace selection,* Proceedings of the International Conference on Machine Learning 2014, Proc. Mach. Learn. Res., 32, 2014, pp. 575–583; available at http://proceedings.mlr.press/v32/.

[14] A. Jalali, M. Fazel, and L. Xiao, *Variational Gram functions: Convex analysis and optimization,* SIAM J. Optim., 27 (2017), pp. 2634–2661.

[15] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[16] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Grundlehren Math. Wiss. 317, Springer, Cham, 1998.