Exact penalty methods for minimizing a smooth function over the nonnegative orthogonal set

Yitian Qian,† Shaohua Pan ‡ and Lianghai Xiao ‡

Abstract

This paper is concerned with the problem of minimizing a smooth function over the nonnegative orthogonal set. For this class of nonconvex nonsmooth optimization problems, we show that the penalized problem induced by the elementwise $\ell_1$-norm distance to the nonnegative cone is a global exact penalty, and so is the one induced by its Moreau envelop under a growth condition on the objective function. Then, we develop the exact penalty methods by solving the smooth and nonsmooth penalty problems, which are respectively minimizing a smooth function and a composite nonsmooth function over the Stiefel manifold in each step. Two retraction-based proximal gradient methods with a line-search strategy are provided for solving the subproblems. Numerical tests on quadratic assignment problems and graph matching problems demonstrate that the penalty method based on the smooth exact penalty is superior to the one based on the nonsmooth exact penalty.

Keywords: nonnegative orthogonal constraint; global exact penalty; KL property
MSC 2010: 90C26; 49Q12; 49M05

1 Introduction

Let $\mathbb{R}^{n \times r}$ ($n \geq r$) represent the vector space of all $n \times r$ real matrices, equipped with the trace inner product $\langle \cdot, \cdot \rangle$ (i.e., $\langle X, Y \rangle = \text{tr}(X^TY)$ for any $X, Y \in \mathbb{R}^{n \times r}$) and its induced Frobenius norm $\| \cdot \|_F$, and let $\mathbb{R}^+_{n \times r}$ be the cone consisting of all nonnegative matrices in $\mathbb{R}^{n \times r}$. We are interested in the problem with the nonnegative orthogonal constraint

$$\min_{X \in \mathbb{R}^+_{n \times r} \cap \text{St}(n,r)} f(X)$$

where $\text{St}(n,r) := \{ X \in \mathbb{R}^{n \times r} | X^T X = I_r \}$ is the orthogonal set in $\mathbb{R}^{n \times r}$, also called the Stiefel manifold in $\mathbb{R}^{n \times r}$, and $f : \mathbb{R}^{n \times r} \to \mathbb{R}$ is a continuously differentiable ($C^1$) function. Clearly, $f$ is Lipschitz continuous on $\text{St}(n,r)$ owing to the compactness of $\text{St}(n,r)$.

---

*School of Mathematics, South China University of Technology, Guangzhou, China.
†(shhpan@scut.edu.cn), School of Mathematics, South China University of Technology, Guangzhou.
‡Corresponding author, School of Mathematics, South China University of Technology, Guangzhou.
Such an optimization problem frequently arises from machine learning and data sciences such as the nonnegative principal component analysis [23, 36], the nonnegative Laplacian embedding [25], the discriminative nonnegative spectral clustering [31], and the orthogonal nonnegative matrix factorization [12]. For more details on its applications, the interested reader is referred to [17, Section 1.1]. Observe that the nonnegative constraint $X \in \mathbb{R}^{n \times r}_+$ introduces a sparsity structure to problem (1), that is, for every $X \in \mathbb{R}^{n \times r}_+ \cap \text{St}(n, r)$, its each row has at most one nonzero element. Hence, when $n = r$, the column orthogonality makes the feasible set of (1) reduce to a permutation matrix set. Consequently, the problem (1) with $n = r$ also covers some combinatorial optimization problems, and a typical example is the quadratic assignment problem [7]:

$$\min_{X \in \mathbb{R}^{n \times n}} \left\{ \langle A, XBX^T \rangle \mid X^TX = I_n, \ X \geq 0 \right\},$$

where $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times n}$ are the given matrices. In addition, the discrete constraint $X \in \mathbb{R}^{n \times n}_+ \cap \text{St}(n, n)$ is also found to have applications in graph matching (see [18, 34]).

The constraint $X \in \mathbb{R}^{n \times r}_+$ hinders the direct application of the existing algorithms on the Stiefel manifold to problem (1). To overcome this difficulty, we consider its penalty problem induced by $\vartheta(X) := \langle E, \max(0, -X) \rangle$, the elementwise $\ell_1$-norm distance to $\mathbb{R}^{n \times r}_+:

$$\min_{X \in \text{St}(n, r)} \Phi_\rho(X) := f(X) + \rho \vartheta(X),$$

where $\rho > 0$ is the penalty parameter and $E \in \mathbb{R}^{n \times r}$ denotes the matrix of all ones. The main contribution of this work is to show that the problem (3) is a global exact penalty for (1), i.e., there exists a threshold $\rho > 0$ such that the problem (3) associated to every $\rho > \rho_0$ has the same global optimal solution set as problem (1) does; and so is the penalty problem induced by the Moreau envelop of $\vartheta$ under a suitable growth condition of $f$:

$$\min_{X \in \text{St}(n, r)} \Theta_{\rho, \gamma}(X) := f(X) + \rho e_\gamma \vartheta(X),$$

where $e_\gamma \vartheta$ with $\gamma > 0$ denotes the Moreau envelop of $\vartheta$; see Section 2 for its definition. Although the problem (3) is precisely the popular $\ell_1$-penalty for problem (1), a special nonlinear program, the existing global exact penalty result for it in the nonlinear program setting requires an unverifiable condition; see [19, Theorem 4.1]. Our global exact penalty result for the problem (3) does not require any condition when $r = n$, and when $r < n$ requires every global optimal solution to have no zero rows. To the best of our knowledge, this is the first global exact penalty for problem (1) that keeps the structure of $\text{St}(n, r)$ undamaged. The structure of $\text{St}(n, r)$ is significant because it is the source to design effective manifold algorithms for solving the penalty problems (3) and (4), which can guarantee that the iterate sequence satisfies the orthogonal constraint very well. We notice that Jiang et al. [16, 17] recently established two different $\ell_p$-norm exact penalties for problem (1) by splitting the Stiefel manifold $\text{St}(n, r)$. Since the structure of $\text{St}(n, r)$ is destroyed, the orthogonality is not well satisfied for the generated iterates. As a result, their heuristic strategy in [16] involves the projection onto the permutation matrix set.
Due to the nonconvexity of the penalty problems (3) and (4), solving a single penalty problem associated to a certain \( \rho > \rho \) generally can not yield a solution of high quality. In view of this, in Section 4 we develop an exact penalty method for problem (1) by seeking the approximate critical points for a finite number of penalty problem (4) (resp. problem (3)) with increasing \( \rho \). For a fixed \( \rho > 0 \), by the smoothness of \( \Theta_{\rho, \gamma} \), the existing algorithms proposed in [29, 20] for manifold optimization can be directly used to solve problem (4). We notice that the work [29] only achieved the convergence of subsequences, while the work [20] obtained the convergence of whole sequence under the retraction convexity of the objective function. Recently, for the problem of minimizing the sum of a smooth function and a nonsmooth function on a Riemannian manifold, Huang and Wei [21] developed a Riemannian proximal gradient method (PGM) that has a global convergence for the restriction of a semi-algebraic function on \( St(n, r) \). Their algorithm is applicable to the problems (4) and (3) for a fixed \( \rho \), but the efficiency will face a challenge if the Lipschitz constant \( L_{\nabla f} \) of \( \nabla f \) is unavailable because the adopted constant step-size is determined by the upper estimation of \( L_{\nabla f} \). Although an accelerated variant was proposed, the inverse of the retraction mapping in each step requires \( O(r^3) \) flops to solve a Lyapunov equation, so is unsuitable for the problem (1) with a large \( r \).

In Section 4.1 we present a line-search retraction-based PGM for solving problem (4) associated to a given \( \rho > 0 \) and verify that the generated whole sequence is convergent if \( f \) is definable in an o-minimal structure \( \mathcal{O} \) over \( \mathbb{R} \). Unlike the one in [21], our convergence analysis depends on both the Euclidean KL property and the recipe established in [3] for nonconvex and nonsmooth optimization problems in the Euclidean space. In Section 4.2, we follow the same line as in [8] to provide a line-search retraction-based PGM for solving (3) associated to a given \( \rho > 0 \), but the descent direction in the tangent space of \( St(n, r) \) is obtained by using a dual regularized semismooth Newton method. This retraction-based PGM, similar to the one in [8], has only the convergence of subsequences. As we discussed beneath Theorem 4.3, its global convergence is almost impossible because the subdifferential mapping is lack of the Lipschitz continuity. This difficulty was overcome by the Riemannian PGM [21] via the computation of the Riemannian proximal mapping in each iterate. However, the computation cost of the Riemannian proximal mapping is expensive because in each step, besides seeking the descent direction in the tangent space, it requires computing the inverse of the adjoint operator of a vector transport.

For the proposed exact penalty methods, we conduct numerical experiments on the 134 instances in QAPLIB, a quadratic assignment problem (QAP) library, and graph pair matching on CMU house image dataset, and compare their performance with that of the augmented Lagrangian function method (ALM) in [29] and its refined implementation version (ALMre). Numerical results show that the penalty method based on \( \Theta_{\rho, \gamma} \) (SEPPG for short) is superior to the one based on \( \Phi_{\rho} \) (EPPGSN for short) and the ALM in terms of the quality of solutions, and SEPPG is comparable with ALMre in terms of the quality of solutions and the CPU time. In addition, SEPPG also yields the comparable matching accuracy and objective value with that of FGM-D in [32], a path-following algorithm with a heuristic strategy designed by the convex and concave relaxations for (1).
2 Notation and preliminaries

In this paper, $M$ represents the Stiefel manifold $\text{St}(n, r)$, an embedded submanifold of the Euclidean space $\mathbb{R}^{n \times r}$, $(\mathcal{M}, \langle \cdot, \cdot \rangle_\gamma)$ denotes a finite dimensional Riemannian manifold, and $C^1(\mathcal{M})$ represents the collection of all $C^1$ functions on $\mathcal{M}$. Let $e$ denote a vector of all ones with dimension known from the context. For a closed set $\mathcal{O} \subseteq \mathbb{R}^{n \times r}$, $\delta_{\mathcal{O}}$ means the indicator function of $\mathcal{O}$, i.e., $\delta_{\mathcal{O}}(x) = 0$ if $x \in \mathcal{O}$ and otherwise $+\infty$; $\text{Proj}_{\mathcal{O}}(\cdot)$ denotes the projection operator onto $\mathcal{O}$ and $\text{dist}(X, \mathcal{O}):= \inf_{Z \in \mathcal{O}} \|Z-X\|_F$; and for a given $x \in \mathcal{O}$, $N_{\mathcal{O}}(x)$ denotes the limiting normal cone to $\mathcal{O}$ at $x$. The tangent and normal space to $\mathcal{M}$ at $x \in \mathcal{M}$ is denoted by $T_x \mathcal{M}$ and $N_x \mathcal{M}$, respectively. We consider the tangent spaces of $\mathcal{M}$ endowed with the Riemannian metric $\langle \cdot, \cdot \rangle_x$ that is induced from the trace inner product, i.e., $\langle G, H \rangle_x = \text{tr}(G^T H)$ for any $G, H \in T_x \mathcal{M}$. For a function $h: \mathbb{R}^{n \times r} \to \mathbb{R} := \mathbb{R} \cup \{+\infty\}$, $h|_\mathcal{M}: \mathcal{M} \to \mathbb{R}$ represents a restriction of $h$ to $\mathcal{M}$; and if $h$ is differentiable at $x \in \mathcal{M}$, we use $\nabla h(x)$ to denote the Riemannian gradient of $h$ at $x$, the (unique) tangent vector at $x$ such that $\langle \nabla h(x), \eta \rangle = \langle \text{grad} h(x), \eta \rangle$ for any $\eta \in T_x \mathcal{M}$, which is also the gradient of $h|_\mathcal{M}$ at $x$. By our choice of the Riemannian metric, $\text{grad} h(x) = \text{Proj}_{T_x \mathcal{M}}(\nabla h(x))$. For a given $X \in \mathcal{M}$, let $A_X(H) := X^T H + H^T X$ for $H \in \mathbb{R}^{n \times r}$. From [1, Chapter 3], it follows that $T_X \mathcal{M} = \{V \in \mathbb{R}^{n \times r} \mid A_X(V) = 0\}$ and $N_X \mathcal{M} = \{2XS \mid S \in S^r\}$, and moreover,

$$\text{Proj}_{T_X \mathcal{M}}(Z) := Z - \frac{1}{2} X A_X(Z) \quad \forall Z \in \mathbb{R}^{n \times r}.$$ 

For a proper lsc function $h: \mathbb{R}^{n \times r} \to \mathbb{R}$ and a constant $\gamma > 0$, $P_\gamma h$ and $e_\gamma h$ denote the proximal mapping and Moreau envelope of $h$ associated to $\gamma$, respectively, defined by

$$P_\gamma h(x) := \arg \min_{z \in \mathbb{R}^{n \times r}} \left\{ \frac{1}{2\gamma} \|z - x\|^2 + h(z) \right\}, \quad e_\gamma h(x) := \min_{z \in \mathbb{R}^{n \times r}} \left\{ \frac{1}{2\gamma} \|z - x\|^2 + h(z) \right\}.$$ 

When $h$ is convex, $P_\gamma h$ is a Lipschitz mapping of modulus $1$ from $\mathbb{R}^{n \times r}$ to $\mathbb{R}^{n \times r}$, and $e_\gamma h$ is a continuously differentiable convex function with $\nabla e_\gamma h(x) = \gamma^{-1}(x - P_\gamma h(x))$. The following lemma characterizes the Clarke Jacobian of the proximal mapping of $\vartheta$.

**Lemma 2.1** Fix any $\gamma > 0$. Then, $P_\gamma \vartheta(X) = \min(X + \gamma, \max(X, 0))$ for $X \in \mathbb{R}^{n \times r}$. Consequently, $\mathcal{E} \in \partial C P_\gamma \vartheta(X)$ if and only if for any $H \in \mathbb{R}^{n \times r}$, $[\mathcal{E}(H)]_{ij} \in \Gamma_{ij}$ with

$$\Gamma_{ij} = \begin{cases} 
1 & \text{if } X_{ij} > 0 \text{ or } X_{ij} < -\gamma; \\
0 & \text{if } -\gamma < X_{ij} < 0; \\
0,1 & \text{if } X_{ij} = 0 \text{ or } -\gamma 
\end{cases} \quad \text{for } i = 1, \ldots, n, j = 1, \ldots, r.$$ 

2.1 Retractions on manifold

**Definition 2.1** A retraction on a manifold $\mathcal{M}$ is a $C^\infty$-mapping $R$ from the tangent bundle $T \mathcal{M} := \bigcup_{x \in \mathcal{M}} T_x \mathcal{M}$ onto $\mathcal{M}$ and for any $x \in \mathcal{M}$, the restriction of $R$ to $T_x \mathcal{M}$, denoted by $R_x$, satisfies the following two conditions:

(i) $R_x(0_x) = x$, where $0_x$ denotes the zero element of $T_x \mathcal{M}$;
We first recall from [22] the subdifferentials of a nonsmooth function on \( \mathbb{R}^{n \times r} \).

### 2.2 Subdifferentials of nonsmooth functions

We first recall from [22] the subdifferentials of a nonsmooth function on \((\mathcal{M}, \langle \cdot, \cdot \rangle_{g})\).

**Definition 2.2** Let \( h : \mathcal{M} \rightarrow \mathbb{R} \) be a function with \( \text{dom} h := \{ x \in \mathcal{M} \mid h(x) < +\infty \} \neq \emptyset \). We define the Fréchet-subdifferential of \( h \) at a point \( x \in \mathcal{M} \) by

\[
\partial h(x) := \begin{cases} \{ d\theta_x \mid \exists \theta \in C^1(\mathcal{M}) \text{ and } h - \theta \text{ attains a local minimum at } x \} & \text{if } x \in \text{dom} h, \\ \emptyset & \text{if } x \notin \text{dom} h, \end{cases}
\]

where \( d\theta_x \in T_x^*\mathcal{M} \), the dual space of \( T_x\mathcal{M} \), is defined by \( d\theta_x(\xi) = \langle \text{grad} h(x), \xi \rangle_g \) for any \( \xi \in T_x\mathcal{M} \); and define its (limiting) subdifferential and singular subdifferential at \( x \in \mathcal{M} \) by

\[
\partial h(x) = \left\{ \lim_{k \rightarrow \infty} v^k \mid \exists v^k \in \partial h(x^k) \text{ with } (x^k, h(x^k)) \rightarrow (x, h(x)) \right\}, \quad \text{and}
\]

\[
\partial^\infty h(x) = \left\{ \lim_{k \rightarrow \infty} \lambda^k v^k \mid \exists v^k \in \partial h(x^k) \text{ with } (x^k, h(x^k)) \rightarrow (x, h(x)) \text{ and } \lambda^k \downarrow 0 \right\}.
\]

**Remark 2.2** (a) From [22, Remark 3.2], the support function \( \theta \) in the definition of the Fréchet subdifferential need only be \( C^1 \) in a neighborhood of \( x \). In addition, by making a translation, the function \( \theta \) can be required to satisfy \( h - \theta \geq 0 \) in a neighborhood of \( x \). If \( \mathcal{M} \) is a Euclidean space, the above subdifferentials are same as in [28, Definition 8.3].

(b) For a proper function \( h : \mathcal{M} \rightarrow \mathbb{R} \) and a point \( x \in \text{dom} h \), the inclusion \( \partial h(x) \subseteq \partial h(x) \) always holds and the multifunction \( \partial h : \mathcal{M} \rightrightarrows \mathcal{T} \mathcal{M} \) is closed. For a proper lsc function \( h : \mathbb{R}^{n \times r} \rightarrow \mathbb{R} \), if \( x^* \) is a local optimal solution of \( \min_{x \in \mathcal{M}} h(x) \), then \( 0 \in \partial h_{|\mathcal{M}}(x^*) \).

For a proper function \( h : \mathbb{R}^{n \times r} \rightarrow \mathbb{R} \), the following lemma discloses the relations between its subdifferentials and those of its restriction on \( \mathcal{M} \). These relations still hold when \( \mathcal{M} \) is replaced by a general embedded submanifold of a Euclidean space.
Lemma 2.3 Consider a proper $h: \mathbb{R}^{n \times r} \to \mathbb{R}$ and $x \in \text{dom}h \cap M$. Let $\tilde{h} := h + \delta_M$. Then, $\partial^h h\vert_M(x) = \text{Proj}_{T_M}(\partial^\tilde{h}(x)) \subseteq \partial^h h\vert_M(x)$ with $\partial^\tilde{h} = \partial h + \partial \delta_M$. Consequently, when $h$ is locally Lipschitz at $x$, $\partial^h h\vert_M(x)$ can take the Clarke subdifferential operator $\partial h$.

Proof: Pick any $v \in \partial h\vert_M(x)$. There exists $\theta \in C^1(M)$ such that $v = \text{grad}\theta(x)$ and $\partial h\vert_M(x)$ is a local minimizer of $h\vert_M - \theta$ with $h\vert_M(x) = \theta(x)$. Since $\theta \in C^1(M)$, there exists a representative function $\tilde{\theta} : \mathbb{R}^{n \times r} \to \mathbb{R}$ that is $C^1$-smooth around $x$ with $\tilde{\theta}\vert_M = \theta$ locally around $x$. From [1, Section 3.6.1], $\text{Proj}_{T_M}(\nabla \tilde{\theta}(x)) = \text{grad}\theta(x) = v$. Then, on some neighborhood of $x$, the function $\tilde{\theta}$ is smooth and $\tilde{\theta} \leq \tilde{h}\vert_M = h\vert_M + \delta_M = \tilde{h}$ with $\tilde{\theta}(x) = \tilde{h}(x)$. By [28, Proposition 8.5], $\nabla \tilde{\theta}(x) \subseteq \partial \tilde{h}(x)$, which means that $v \in \text{Proj}_{T_M}(\partial \tilde{h}(x))$. So, $\partial h\vert_M(x) \subseteq \text{Proj}_{T_M}(\partial \tilde{h}(x))$. To establish the converse inclusion, pick any $v \in \text{Proj}_{T_M}(\partial \tilde{h}(x))$. Then, there exists $\xi \in \partial \tilde{h}(x)$ such that $v = \text{Proj}_{T_M}(\xi)$. Since $\xi \in \partial \tilde{h}(x)$, by [28, Proposition 8.5], on some neighborhood $U$ of $x$, there is a smooth function $\theta \leq \tilde{h}$ with $\nabla \theta(x) = \xi$. Clearly, $\theta\vert_M$ is a smooth function on $U \cap M$ such that $\partial h\vert_M(x)$ is a local minimizer of $h\vert_M - \theta\vert_M$. By Remark 2.2 (a), $v = \text{grad}\theta\vert_M(x) \in \partial h\vert_M(x)$. The converse inclusion follows. The first equality holds. For the first inclusion, pick any $v \in \text{Proj}_{T_M}(\partial \tilde{h}(x))$. Then, there exists $\xi \in \partial \tilde{h}(x)$ such that $v = \text{Proj}_{T_M}(\xi)$. Thus,

$$
\liminf_{\xi \neq x \in M} \frac{\tilde{h}(x) - \tilde{h}(\xi) - \langle v, x - \xi \rangle}{\|x - \xi\|} = \liminf_{\xi \neq x \in M} \frac{\tilde{h}(x) - \tilde{h}(\xi) - \langle v, x - \xi \rangle}{\|x - \xi\|}
= \frac{\tilde{h}(x) - \tilde{h}(\xi) - \langle \xi, x - \xi \rangle + \langle \text{Proj}_{M^*}(\xi), x - \xi \rangle}{\|x - \xi\|}
\geq \frac{\tilde{h}(x) - \tilde{h}(\xi) - \langle \xi, x - \xi \rangle}{\|x - \xi\|} - \limsup_{\xi \neq x \in M} \frac{\langle -\text{Proj}_{M^*}(\xi), x - \xi \rangle}{\|x - \xi\|}
\geq \frac{\tilde{h}(x) - \tilde{h}(\xi) - \langle \xi, x - \xi \rangle}{\|x - \xi\|} - 0
= 0
$$

where the second inequality holds by $-\text{Proj}_{M^*}(\xi) \in N^*_M$ and the last is by $\xi \in \partial \tilde{h}(x)$. By the definition of the Fréchet subgradient, $v \in \partial \tilde{h}(x)$. The first inclusion holds.

To prove that $\partial h\vert_M(x) = \text{Proj}_{T_M}(\partial \tilde{h}(x))$, pick any $v \in \partial h\vert_M(x)$. Clearly, $v \in T_M$. Also, there exist $(x^k, h\vert_M(x^k)) \to (\xi, h\vert_M(\xi))$ and $v^k \in \partial h\vert_M(x^k)$ for each $k$ such that $v^k \to v$. Notice that the first equality and inclusion hold for all $x \in \text{dom}h \cap M$. So, $v^k \in \text{Proj}_{T_{x^k}}(\partial \tilde{h}(x^k)) \subseteq \partial \tilde{h}(x^k)$ for each $k$. Since $h\vert_M(x^k) \to h\vert_M(\xi)$ implies $(x^k, \tilde{h}(x^k)) \to (\xi, \tilde{h}(\xi))$, we have $v \in \partial \tilde{h}(\xi)$. Together with $v \in T_M$, it follows that $v = \text{Proj}_{T_M}(v) \in \text{Proj}_{T_M}(\partial \tilde{h}(x))$ and $\partial h\vert_M(x) \subseteq \text{Proj}_{T_M}(\partial \tilde{h}(x))$. For the converse inclusion, pick any $v \in \text{Proj}_{T_M}(\partial \tilde{h}(x))$. There exists $\xi \in \partial \tilde{h}(x)$ such that $v = \text{Proj}_{T_M}(\xi)$. Since $\xi \in \partial \tilde{h}(x)$, there exists $(x^k, \tilde{h}(x^k)) \to (\xi, \tilde{h}(\xi))$ with $\xi^k \in \partial \tilde{h}(x^k)$ for every $k$ such that $\xi^k \to \xi$. Since $\tilde{h}(x^k) \to \tilde{h}(\xi)$ implies that $\{x^k\} \subseteq M$ and $h(x^k) \to h(\xi)$, we have $h\vert_M(x^k) \to h\vert_M(\xi)$. Let $v^k := \text{Proj}_{T_{x^k}}(\xi^k)$. From (5), $v^k \to v$. By the first equality, $v^k \in \partial h\vert_M(x^k)$ for each
k. So, \( v \in \partial h|_M(\xi) \) and \( \partial h|_M(\xi) \supset \text{Proj}_{T_xM}(\partial \tilde{h}(\xi)) \). The second equality holds. For the second inclusion, pick any \( v \in \text{Proj}_{T_xM}(\partial \tilde{h}(\xi)) \). There exists \( \xi \in \partial \tilde{h}(\xi) \) such that \( v = \text{Proj}_{T_xM}(\xi) \). Since \( \xi \in \partial \tilde{h}(\xi) \), there exist \( (x^k, \tilde{h}(x^k)) \rightarrow (\xi, h(\xi)) \) and \( \xi^k \in \partial \tilde{h}(x^k) \) for each \( k \) such that \( \xi^k \rightarrow \xi \) as \( k \rightarrow \infty \). By the first inclusion, \( \text{Proj}_{T_xM}(\xi^k) \in \partial \tilde{h}(\xi^k) \). By the outer semicontinuity of \( \partial \tilde{h} \) and (5), \( v = \text{Proj}_{T_xM}(\xi) \in \partial \tilde{h}(\xi) \). The second inclusion holds. Using the similar arguments yields the third equality and inclusion.

By Lemma 2.3, the smoothness of \( f \) and the convexity of \( \partial \), the following result holds.

**Corollary 2.1** Fix any \( \rho > 0, \gamma > 0 \). For any \( x \in \text{St}(n, r) \), the following relations hold
\[
\partial(\Theta_{\rho, \gamma})|_M(X) = \{ \text{grad} \Theta_{\rho, \gamma}(X) \} = \{ \text{Proj}_{T_M}(\nabla \Theta_{\rho, \gamma}(X)) \},
\partial(\Phi_{\rho, \gamma})|_M(X) = \hat{\partial}(\Phi_{\rho, \gamma})|_M(X) = \text{Proj}_{T_M}(\nabla f(X) + \rho \partial \vartheta(X)).
\]

### 2.3 KL property on Riemannian manifold

**Definition 2.3** (see [9, Definition 3.3]) Let \( h : M \rightarrow \mathbb{R} \) be a proper function. The function \( h \) is said to have the KL property at \( x \in \text{dom} \partial h \) if there exist \( \eta \in (0, +\infty) \), a neighborhood \( U \) of \( x \), and a continuous concave function \( \varphi : [0, \eta) \rightarrow \mathbb{R}_+ \) such that
\[
\begin{align*}
(\text{i}) & \quad \varphi(0) = 0, \ \varphi \text{ is } C^1 \text{ on } (0, \eta), \text{ and } \varphi'(s) > 0 \text{ for all } s \in (0, \eta); \\
(\text{ii}) & \quad \text{for all } x \in U \cap \{ x \in M \mid h(x) < h(\xi) < h(x) + \eta \}, \text{ the KL inequality holds:}
\end{align*}
\]
\[
\varphi'(h(x) - h(\xi))d(0, \partial h(x)) \geq 1,
\]
where \( d(0, \partial h(x)) := \inf_{v \in \partial h(x)} \| v \|_g \) and \( \| \cdot \|_g \) is the norm induced by \( \langle \cdot, \cdot \rangle_g \).

If \( h \) has the KL property at every point of \( \text{dom} \partial h \), then it is called a KL function.

Clearly, when \( M \) is a Euclidean space, Definition 2.3 reduces to [2, Definition 3.1]. For a proper \( h : \mathbb{R}^{n \times r} \rightarrow \mathbb{R} \), the following lemma discloses the relation between the KL property of \( h + \delta_M \) and its restriction \( h|_M \) on the manifold \( M \), which still holds if \( M \) is replaced by a general embedded submanifold of \( \mathbb{R}^{n \times r} \). Although this result is not used in this paper, as will be shown in Remark 2.3, it provides a more convenient criterion to identify the KL property of a function defined on \( M \) than the one in [21, Theorem 6].

**Lemma 2.4** Consider a proper function \( h : \mathbb{R}^{n \times r} \rightarrow \mathbb{R} \). If \( h + \delta_M \) is a KL function, then \( h|_M \) is a KL function. Conversely, if \( h|_M \) is a KL function, then so is \( h + \delta_M \).

**Proof:** Write \( \tilde{h} := h + \delta_M \). Suppose \( \tilde{h} \) is a KL function. Fix any \( \xi \in \text{dom} \partial h|_M \). If \( 0 \notin \partial h|_M(\xi) \), then \( h|_M \) has the KL property at \( \xi \) by [9, Lemma 3.1]. So, it suffices to consider the case \( 0 \in \partial h|_M(\xi) \). Since \( \tilde{h} \) has the KL property at \( \xi \), by [2, Definition 3.1], there exist \( \eta \in (0, +\infty) \), a neighborhood \( U \) of \( \xi \), and a continuous concave function \( \varphi : [0, \eta) \rightarrow \mathbb{R}_+ \) satisfying Definition 2.3 (i) such that for all \( x \in U \cap \{ z \in \text{dom} \tilde{h} \mid \tilde{h}(\xi) < \tilde{h}(z) < \tilde{h}(\xi) + \eta \},
\]
\[
1 \leq \varphi'(\tilde{h}(x) - \tilde{h}(\xi))d(0, \partial \tilde{h}(x)) = \varphi'(h(x) - h(\xi))d(0, \partial \tilde{h}(x)).
\]
Now fix any \( x \in U \cap \left\{ z \in M \mid h_{|M}(\overline{x}) < h_{|M}(z) < h_{|M}(\overline{x}) + \eta \right\} \). If \( \partial h_{|M}(x) = \emptyset \), clearly,
\[
\varphi'(h_{|M}(x) - h_{|M}(\overline{x})) \text{dist}(0, \partial h_{|M}(x)) \geq 1. \tag{9}
\]

We next consider the case that \( \partial h_{|M}(x) \neq \emptyset \). By Lemma 2.3, \( \partial \tilde{h}(x) \neq \emptyset \) and moreover,
\[
\text{dist}(0, \partial h_{|M}(x)) = \text{dist}(0, \text{Proj}_{T_{x,M}}(\partial \tilde{h}(x))) = \inf_{v \in \text{Proj}_{T_{x,M}}(\partial \tilde{h}(x))} \|v\|_F \geq \inf_{v \in \partial \tilde{h}(x)} \|v\|_F.
\]
Together with (8), it follows that \( \varphi'(h(x) - h(\overline{x})) \text{dist}(0, \partial h_{|M}(x)) \geq 1 \), i.e., the inequality (9) also holds. This shows that \( h_{|M} \) satisfies the KL property at \( \overline{x} \). By the arbitrariness of \( \overline{x} \in \text{dom} \partial h_{|M} \), we conclude that \( h_{|M} \) is a KL function.

Now suppose that \( h_{|M} \) is a KL function. Fix any \( \overline{x} \in \text{dom} \partial \tilde{h} \). If \( 0 \notin \partial \tilde{h}(\overline{x}) \), then \( \tilde{h} \) has the KL property at \( \overline{x} \) by [2, Lemma 2.1]. So, it suffices to consider the case that \( 0 \in \partial \tilde{h}(\overline{x}) \). Since the function \( h_{|M} \) admits the KL property at \( \overline{x} \), there exist \( \eta \in (0, +\infty) \), a neighborhood \( U \) of \( \overline{x} \), and a continuous concave function \( \varphi : [0, \eta) \to \mathbb{R}_+ \) satisfying Definition 2.3 (i) such that for all \( x \in U \cap \left\{ z \in M \mid h_{|M}(\overline{x}) < h_{|M}(z) < h_{|M}(\overline{x}) + \eta \right\} \),
\[
1 \leq \varphi'(h_{|M}(x) - h_{|M}(\overline{x})) \text{dist}(0, \partial h_{|M}(x)) = \varphi'(\tilde{h}(x) - \tilde{h}(\overline{x})) \text{dist}(0, \text{Proj}_{T_{x,M}}(\partial \tilde{h}(x)))
\]
which the equality is due to Lemma 2.3 and the definition of \( \tilde{h} \). In addition, by Lemma 2.3, \( \text{Proj}_{T_{x,M}}(\partial \tilde{h}(x)) \subseteq \partial \tilde{h}(x) \), which implies that \( 0 \in \text{Proj}_{N_x,M}(\partial \tilde{h}(x)) \). Then,
\[
dist(0, \partial \tilde{h}(x)) = \inf_{v \in \partial \tilde{h}(x)} \|v\|_F \geq \inf_{v \in \text{Proj}_{T_{x,M}}(\partial \tilde{h}(x)) + \text{Proj}_{N_x,M}(\partial \tilde{h}(x))} \|v\|_F = \inf_{v^1 \in \text{Proj}_{T_{x,M}}(\partial \tilde{h}(x))} \|v^1\|_F + \inf_{v^2 \in \text{Proj}_{N_x,M}(\partial \tilde{h}(x))} \|v^2\|_F = \inf_{v^1 \in \text{Proj}_{T_{x,M}}(\partial \tilde{h}(x))} \|v^1\|_F = \text{dist}(0, \text{Proj}_{T_{x,M}}(\partial \tilde{h}(x))).
\]
From the last two inequalities, \( \tilde{h} \) has the KL property at \( \overline{x} \) and \( \tilde{h} \) is a KL function. \( \square \)

Remark 2.3 Lemma 2.4 provides a criterion to identify the KL property of a function \( h : M \to \mathbb{R}_+ \) in terms of that of its extension \( \tilde{h} \) with \( \tilde{h}(x) = \left\{ \begin{array}{ll} h(x) & \text{if } x \in M, \\ +\infty & \text{otherwise,} \end{array} \right. \) since a collection of convenient criteria are included in [2, Section 4] for that of \( \tilde{h} \). From the last part of Lemma 2.3 and the proof of Lemma 2.4, the conclusions of Lemma 2.4 still hold if the KL property is defined by the Clarke subdifferential. Recently, Huang and Wei [21] provided a criterion to identify such a KL property of \( h \), but their criterion involves verifying that the inverse of the chart associated to \( M \) is semialgebraic, which as shown by the proof of [21, Lemma 8 & 9] is not an easy task.

Lemma 2.5 Let \( h : \mathbb{R}^n \to \mathbb{R}_+ \) be a proper lsc function, and let \( a > 0 \) and \( b > 0 \) be two constants. Consider a sequence \( \{x^k\}_{k=1}^\infty \) satisfying the following three conditions:

\[ \]
H1. For each integer \( k \geq 0 \), \( h(x^{k+1}) + a\|x^{k+1} - x^k\|^2 \leq h(x^k) \).

H2. For each integer \( k \geq 0 \), there exists \( w^k \in \partial h(x^k) \) such that \( \|w^k\| \leq b\|x^{k+1} - x^k\| \).

H3. There is a subsequence \( \{x^j\}_{j=0}^{\infty} \) such that \( x^{k_j} \to \pi \) and \( f(x^{k_j}) \to f(\pi) \) as \( j \to \infty \). Then \( \sum_{k=0}^{j} \|x^{k+1} - x^k\| < \infty \) whenever \( h \) has the KL property at the cluster point \( \pi \).

**Proof:** Notice that condition H2 has a little difference from that of [3]. However, by rechecking the proof of [3, Lemma 2.6], it is not difficult to find that the conclusion of [3, Lemma 2.6] still holds under the above H1-H3 by modifying (7) and (9) there as

\[
\sum_{i=1}^{j} \|x^{i+1} - x^i\| \leq \frac{b}{a}[\varphi(h(x^1) - h(x^*)) - \varphi(h(x^{j+1}) - h(x^*))],
\]

\[
\|x^{k+1} - x^k\| \leq \frac{b}{a}[\varphi(h(x^k) - h(x^*)) - \varphi(h(x^{k+1}) - h(x^*))].
\]

Then, following the same arguments as those in [3, Theorem 2.9] yields the result. \( \square \)

3 Exactness of penalty problems

To establish that problem (3) is a global exact penalty for (1), we need two technical lemmas: Lemma 3.1 and 3.2. Lemma 3.1 states that the distance between a nonnegative matrix and a rectangular unit matrix is locally upper bounded by the distance between their singular value vectors, and Lemma 3.2 states that the distance of a nonnegative matrix and a rectangular unit matrix is locally upper bounded by the distance between their singular value vectors, and that of an identity matrix. For convenience, in the rest of this section, \( I_{n \times r} \) represents the \( n \times r \) rectangular unit matrix; for any \( X \in \mathbb{R}^{n \times r} \), \( X_r \) denotes the submatrix consisting of its first \( r \) rows, \( \sigma(X) \) denotes the singular value vector with entries arranged in a nonincreasing order, and \( \|X\|_1, \|X\|_* \) and \( \|X\| \) denote the elementwise \( \ell_1 \)-norm, the nuclear norm and the spectral norm of \( X \), respectively.

**Lemma 3.1** There exists \( \delta > 0 \) such that for all \( X \in B(I_{n \times r}, \delta) \cap \{X \in \mathbb{R}^{n \times r} | X_r \geq 0 \} \) with \( \|X - I_{n \times r}\|_F \leq c\|X_r - I_r\|_F \) for some constant \( c > 0 \),

\[
\|X_r - I_r\|_F \leq 2.1\sqrt{r}\|\sigma(X) - e\|.
\]

**Proof:** Suppose on the contradiction that the conclusion does not hold. There exists \( X^k \to I_{n \times r} \) with \( X^k_r \geq 0 \) and \( 0 < \|X^k - I_{n \times r}\|_F \leq c\|X^k_r - I_r\|_F \) for each \( k \) such that

\[
\|\sigma(X^k) - e\| < \frac{1}{2.1\sqrt{r}}\|X^k_r - I_r\|_F \quad \text{for each} \quad k.
\]

(10)

For each \( k \), write \( H^k := X^k - I_{n \times r} \). From [11, Proposition 6], for all sufficiently large \( k \),

\[
\sigma_i(X^k) - 1 = \lambda_i \left( \frac{H^k_r + (H^k_r)^T}{2} \right) + O(\|H^k\|^2_F) \quad \text{for} \quad i = 1, 2, \ldots, r,
\]
where \( \lambda_i(Z) \) denotes the \( i \)th largest eigenvalue of a \( r \times r \) symmetric matrix \( Z \). Notice that \( \|\sigma(X^k) - e\| \geq \frac{1}{\sqrt{r}} \sum_{i=1}^r |\sigma_i(X^k) - 1| \). Hence, for all \( k \) large enough,

\[
\|\sigma(X^k) - e\| \geq \frac{1}{\sqrt{r}} \sum_{i=1}^r \left| \lambda_i \left( \frac{H^k_i + (H^k_i)^T}{2} \right) \right| + O(\|H^k\|_F)
\]

\[
= \frac{1}{2\sqrt{r}} \|H^k_i + (H^k_i)^T\|_* + O(\|H^k\|_F^2)
\]

\[
\geq \frac{1}{2\sqrt{r}} \|H^k_i + (H^k_i)^T\|_F + O(\|H^k\|_F^2) > \frac{1}{2\sqrt{r}} \|H^k\|_F,
\]

where the third inequality holds by \( X^k_i \geq 0, H^k_i = X^k_i - I_r \) and \( \|H^k\|_F \leq c\|H^k\|_F \). This yields a contradiction to inequality (10). Consequently, the conclusion follows. \( \square \)

**Lemma 3.2** Fix any \( \overline{X} \in \mathbb{R}^{n \times r}_+ \cap \text{St}(n, r) \) with \( n > r \). Let \( \overline{X}_{i,j} \) be the smallest nonzero entry of \( \overline{X} \). If the matrix \( \overline{X} \) has no zero rows, then there exists \( \delta > 0 \) such that

\[
\text{dist}(X, \mathbb{R}^{n \times r}_+ \cap \text{St}(n, r)) \leq \frac{10.5 \nu^{3/2} (n-r)}{\overline{X}_{i,j}},
\]

**Proof:** For \( k = 1, 2, \ldots, n \), write \( l_k := \arg \max_{1 \leq j \leq r} X_{kj} \). Since \( \overline{X} \in \mathbb{R}^{n \times r}_+ \cap \text{St}(n, r) \) with \( n > r \) and \( \overline{X} \) has no zero rows, each row of \( \overline{X} \) has only one nonzero entry and each column of \( \overline{X} \) has at least one nonzero entry, which implies that \( X_{kl_k} > 0 \) for \( k = 1, \ldots, n \) and \( \{l_1, l_2, \ldots, l_n\} \) contains \( \{1, 2, \ldots, r\} \). Then there exists \( \delta_1 \in (0, 1/3) \) such that for every \( \overline{X} \in \mathbb{B}(\overline{X}, \delta_1) \), \( \arg \max_{1 \leq j \leq r} X_{kj} = l_k \) and \( X_{kl_k} > 0 \) for all \( k = 1, \ldots, n \). In addition, by Lemma 3.1, there exists \( \delta_2 > 0 \) such that for all \( X' \in \mathbb{B}(I_{n \times r}, \delta_2) \cap \{X' \in \mathbb{R}^{n \times r} | X'_{i,j} \geq 0\} \) with \( \|X' - I_{n \times r}\|_F \leq c\|X'_{i,j} - I_{r}\|_F \) for some constant \( c > 0 \),

\[
\|X'_{i,j} - I_{r}\|_F \leq 2.1 \sqrt{r} \|\sigma(X') - e\|.
\]

Set \( \delta = \min \left( \frac{1}{2} \overline{X}_{i,j}, \frac{\delta_2}{\min(2\sqrt{r} + 1, \delta_1)} \right) \). Pick any \( X' \in \mathbb{B}(\overline{X}, \delta) \cap \mathbb{R}^{n \times r}_+ \). Define \( \check{X} \in \mathbb{R}^{n \times r} \) with

\[
\check{X}_{ij} := \begin{cases} X_{ij} & \text{if } j = l_i, \\ 0 & \text{if } j \neq l_i, \end{cases}
\]

for \( i = 1, 2, \ldots, n \) and \( j = 1, 2, \ldots, r \).

Let \( D = \text{Diag}(\|\check{X}_{1}\|, \ldots, \|\check{X}_{r}\|) \). Then, for \( i = 1, 2, \ldots, r \), \( D_{ii} \leq \|X_i\| \leq 1 + \delta \) and \( D_{ii} \geq \|X_i\| - \|\check{X}_i - X_i\| \geq \|X_i\| - \delta \geq 1 - 2\delta \), and hence \( -\frac{2\delta}{1 - 2\delta} \leq 1 - \frac{\delta}{1 + \delta} \leq \frac{\delta}{1 + \delta} \geq \frac{\delta}{1 + \delta} \). Also, \( \check{X}D^{-1} \in \mathbb{R}^{n \times r}_+ \cap \text{St}(n, r) \). Pick \( Z \in \mathbb{R}^{n \times (n-r)} \) such that \( \check{X}D^{-1} Z \) is orthogonal. Then,

\[
\text{dist}(X, \mathbb{R}^{n \times r}_+ \cap \text{St}(n, r)) \leq \|X - \check{X}D^{-1}\|_F = \|\check{X}D^{-1} \|_F (X - \check{X}D^{-1})\|_F
\]

\[
= \|\check{X}D^{-1} Z\|_F (X - I_{n \times r})\|_F.
\]
We next upper bound \( \|Z^TX\|_1 \) via \( \|D^{-1}\tilde{X}^TX - I_r\|_1 \). Write \( H := X - \overline{X} \). Notice that

\[
\|Z^TX\|_1 = \sum_{i=1}^{n} \sum_{j=1}^{r} |\sum_{k=1}^{n} Z_{ki} (X_{kj} - \overline{X}_{kj})| = \sum_{i=1}^{n} \sum_{j=1}^{r} |\sum_{k=1}^{n} (1 - \mathbb{1}_{\{j=l_k\}})Z_{ki}X_{kj}| \\
= \sum_{i=1}^{n} \sum_{j=1}^{r} \sum_{k=1}^{n} |Z_{ki}H_{kj}| \leq (n-r) \sum_{k=1}^{n} \sum_{j \neq l_k} |H_{kj}|, \quad (13)
\]

where the first equality is using \( Z^T\overline{X} = 0 \), the second inequality is due to \( \overline{X}_{kj} = 0 \) for \( k = 1, \ldots, n, l_i \neq j \in \{1, \ldots, r\} \), and the last one is from \( |Z_{ki}| \leq 1 \) for all \( k = 1, \ldots, n \) and \( i = 1, \ldots, n-r \). On the other hand, from \( \tilde{X}, \overline{X} \in \mathbb{R}^{n \times r} \) it follows that

\[
\|D^{-1}\tilde{X}^TX - I_r\|_1 \geq \frac{1}{\|D\|} \|\sum_{i \neq j} (\tilde{X}^TX)_{ij}\| = \frac{1}{\|D\|} \|\sum_{i \neq j} (\tilde{X}^TX)_{ij}\| \\
\geq \frac{1}{\|D\|} \sum_{i \neq j} \sum_{k=1}^{n} \tilde{X}_{ki}X_{kj} = \frac{1}{\|D\|} \|\sum_{i \neq j} \sum_{k=1}^{n} \mathbb{1}_{\{i=l_k\}}(X_{ki} + H_{ki})(X_{kj} + H_{kj})\| \\
= \frac{1}{\|D\|} \sum_{i \neq j} \sum_{k=1}^{n} \mathbb{1}_{\{i=l_k\}}(X_{ki}H_{kj} + H_{ki}H_{kj}) \\
= \frac{1}{\|D\|} \sum_{k=1}^{n} \sum_{j \neq l_k} (X_{kl_k}H_{kj} + H_{kl_k}H_{kj}) \geq \frac{1}{\|D\|} \sum_{k=1}^{n} \sum_{j \neq l_k} \left[ \frac{1}{2} \tilde{X}^T_{i,j} |H_{kj}| \right],
\]

where the second equality is because each row of \( \overline{X} \) has only one nonzero entry, and the last inequality is using \( H_{kj} \geq 0 \) when \( j \neq l_k \), \( |H_{kl_k}| \leq \delta \) and \( \overline{X}_{kl_k} \geq \overline{X}^*_{i,j} \) for \( k = 1, \ldots, n \). Combining the last inequality with (13) yields that \( \frac{\overline{X}^*_{i,j,r}}{n-r} \|Z^TX\|_1 \leq 4\|D^{-1}\tilde{X}^TX - I_r\|_1 \). Then \( \|\tilde{X}D^{-1}Z^TX - I_{n \times r}\|_F \leq c\|D^{-1}\tilde{X}^TX - I_r\|_F \) with \( c = \frac{\overline{X}^*_{i,j,r} + 4(n-r)r}{\overline{X}^*_{i,j,r}} \). Notice that

\[
\|\tilde{X}D^{-1}Z^TX - I_{n \times r}\|_F = \|X - [\tilde{X}D^{-1}Z]I_{n \times r}\|_F = \|X - \tilde{X}D^{-1}\|_F \\
\leq \|X - XD^{-1}\|_F + \|(X - \tilde{X})D^{-1}\|_F \leq \frac{2(\sqrt{\gamma} + \delta)\delta}{1 - 2\delta} + \frac{\delta}{1 - 2\delta} \leq \delta_2,
\]

where the second inequality is due to \( -\frac{2\delta}{1 - 2\delta} \leq 1 - \frac{1}{2\delta} \leq \frac{\delta}{1 + \delta} \). By invoking the inequality (11) and using \( \frac{\overline{X}^*_{i,j,r}}{n-r} \|Z^TX\|_1 \leq 4\|D^{-1}\tilde{X}^TX - I_r\|_1 \), it follows that

\[
2.1\sqrt{\gamma} \|\sigma(X) - e\| = 2.1\sqrt{\gamma} \|\sigma([\tilde{X}D^{-1}Z]^TX) - e\| \geq \|D^{-1}\tilde{X}^TX - I_r\|_F \\
\geq \frac{1}{5} \|D^{-1}\tilde{X}^TX - I_r\|_F + \frac{\overline{X}^*_{i,j,r}}{5r(n-r)} \|Z^TX\|_F \\
\geq \frac{\overline{X}^*_{i,j,r}}{5r(n-r)} \|\tilde{X}D^{-1}Z^TX - I_{n \times r}\|_F,
\]

11
where the first equality is due to the orthogonality of $[\tilde{X}D^{-1} Z]$. Together with (12), we obtain the desired result. The proof is then completed. □

By using Lemma 3.1 and 3.2, we can establish the following metric qualification.

**Proposition 3.1** Fix any $\overline{X} \in \mathbb{R}^{n \times r}_+ \cap \text{St}(n, r)$. The following statements hold.

(i) When $r = n$, there exist $\delta > 0$ and $\kappa > 0$ such that for all $X \in \mathbb{B}(\overline{X}, \delta)$,

$$\text{dist}(X, \mathbb{R}^{n \times r}_+ \cap \text{St}(n, r)) \leq \kappa \left[ \text{dist}(X, \mathbb{R}^{n \times r}_+) + \text{dist}(X, \text{St}(n, r)) \right].$$

(ii) When $r < n$, if $\overline{X}$ has no zero rows, then there exist $\delta > 0$ and $\kappa > 0$ such that

$$\text{dist}(X, \mathbb{R}^{n \times r}_+ \cap \text{St}(n, r)) \leq \kappa \left[ \text{dist}(X, \mathbb{R}^{n \times r}_+) + \text{dist}(X, \text{St}(n, r)) \right] \quad \forall X \in \mathbb{B}(\overline{X}, \delta).

**Proof:** (i) By Lemma 3.1, there exists $\delta > 0$ such that for all $Z \in \mathbb{B}(I_n, \delta) \cap \mathbb{R}^{n \times n}_+$,

$$\|Z - I_n\|_F \leq 2.1\sqrt{n}\|\sigma(Z) - e\|.$$

Pick any $X \in \mathbb{B}(\overline{X}, \delta/2)$. Let $X_+ := \text{Proj}_{\mathbb{R}^{n \times n}_+}(X)$. Then, $\overline{X}^TX_+ \in \mathbb{B}(I_n, \delta) \cap \mathbb{R}^{n \times n}_+$ by noting that $\|\overline{X}^TX_+ - I_n\|_F = \|X_+ - \overline{X}\|_F \leq \delta$. From the last inequality, it follows that

$$\text{dist}(X_+, \mathbb{R}^{n \times n}_+ \cap \text{St}(n, n)) \leq \|X_+ - \overline{X}\|_F = \|\overline{X}^T(X_+ - \overline{X})\|_F = \|\overline{X}^T(X_+ - I_n)\|_F \leq 2.1\sqrt{n}\|\sigma(\overline{X}^T X_+) - e\|.$$

Let $X_+$ have the SVD as $X_+ = U\text{Diag}(\sigma(X_+))V^T$ where $U$ and $V$ are $n \times n$ orthogonal matrices. Then $\text{dist}(X_+, \text{St}(n, n)) = \|X_+ - UV^T\|_F = \|\sigma(X_+) - e\| = \|\sigma(\overline{X}^T X_+) - e\|$. Along with the last inequality $\text{dist}(X_+, \mathbb{R}^{n \times n}_+ \cap \text{St}(n, n)) \leq 2.1\sqrt{n}\text{dist}(X_+, \text{St}(n, n))$. Then

$$\text{dist}(X, \mathbb{R}^{n \times r}_+ \cap \text{St}(n, n)) \leq \|X - X_+\|_F + \text{dist}(X_+, \mathbb{R}^{n \times n}_+ \cap \text{St}(n, n))$$

$$\leq \|X - X_+\|_F + 2.1\sqrt{n}\text{dist}(X_+, \text{St}(n, n))$$

$$\leq \|X - X_+\|_F + 2.1\sqrt{n}[\|X - X_+\|_F + \text{dist}(X, \text{St}(n, n))])$$

$$\leq (2.1\sqrt{n} + 1)\left[ \text{dist}(X, \mathbb{R}^{n \times n}_+) + \text{dist}(X, \text{St}(n, n)) \right].$$

This, by the arbitrariness of $X$ in $\mathbb{B}(\overline{X}, \delta/2)$, shows that the result holds for $\kappa = 2.1\sqrt{n} + 1$.

(ii) By Lemma 3.2, there exists $\delta > 0$ such that for all $Z \in \mathbb{B}(\overline{X}, \delta) \cap \mathbb{R}^{n \times r}_+$,

$$\text{dist}(Z, \mathbb{R}^{n \times r}_+ \cap \text{St}(n, r)) \leq \frac{10.5\delta^{3/2}(n-r)}{X_{i^*, j^*}}\|\sigma(Z) - e\|,$$

where $X_{i^*, j^*}$ is the smallest nonzero entry of $\overline{X}$. Pick any $X \in \mathbb{B}(\overline{X}, \delta/2)$. Let $X_+ := \text{Proj}_{\mathbb{R}^{n \times r}_+}(X)$. Clearly, $X_+ \in \mathbb{B}(\overline{X}, \delta) \cap \mathbb{R}^{n \times r}_+$, since $\|X_+ - \overline{X}\|_F \leq 2\|X - \overline{X}\|_F \leq \delta$. Then

$$\text{dist}(X_+, \mathbb{R}^{n \times r}_+ \cap \text{St}(n, r)) \leq \frac{10.5\delta^{3/2}(n-r)}{X_{i^*, j^*}}\|\sigma(X_+) - e\|.$$
Using the same arguments as those for (i), we get the result with \( \kappa = 1 + \frac{10.5r^{3/2}(n-r)}{X_{i,r,j}^*} \).

It is worthwhile to emphasize that the assumption that \( \mathcal{X} \) has no zero rows is necessary for Proposition 3.1 (ii). To see this, consider \( \mathcal{X} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^T \). Let \( X^k = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}^T \) for each \( k \in \mathbb{N} \). Clearly, \( \lim_{k \to \infty} X^k = \mathcal{X} \). Since each row of \( X \in \mathbb{R}^{3 \times 2} \cap \text{St}(3, 2) \) has at most one nonzero entry, \( \text{dist}(X^k, \mathbb{R}_{+}^{3 \times 2} \cap \text{St}(3, 2)) \geq \frac{1}{k} \). However, \( \text{dist}(X^k, \mathbb{R}_{+}^{n \times r}) + \text{dist}(X^k, \text{St}(n, r)) = \|\sigma(X^k) - \epsilon\| = \sqrt{1 + \frac{1}{k^2} - 1} = O\left(\frac{1}{k}\right) \).

By combining Proposition 3.1 and [27, Section 3.1] and using the regularity of \( \text{St}(n, r) \), we have the following characterization on the normal cone to the set \( \mathbb{R}_{+}^{n \times r} \cap \text{St}(n, r) \).

**Corollary 3.1** Consider any \( X \in \mathbb{R}_{+}^{n \times r} \cap \text{St}(n, r) \). If \( X \) has no row when \( n < r \), then
\[
\mathcal{N}_{\mathbb{R}_{+}^{n \times r} \cap \text{St}(n, r)}(X) = \mathcal{N}_{\mathbb{R}_{+}^{n \times r}}(X) + \mathcal{N}_{\text{St}(n, r)}(X).
\]

Now we are in a position to achieve the global exact penalty result for problem (3).

**Theorem 3.1** Suppose that every global (or local) optimizer of problem (1) has no zero rows when \( r < n \). Then, for any global (or local) optimizer \( X^* \) of (1), there exists \( \delta > 0 \) such that for all \( \epsilon \geq 0 \) and all \( X \in \mathcal{B}(X^*, \delta) \cap \{ X \in \text{St}(n, r) \mid \vartheta(X) = \epsilon \} \),
\[
f(X) - f(X^*) + \kappa L \vartheta(X) \geq 0,
\]
where \( \kappa \) is same as in Proposition 3.1 and \( L \) is the Lipschitz constant of \( f \) on \( \text{St}(n, r) \), and consequently, there exists a threshold \( \tilde{p} > 0 \) such that the penalty problem (3) associated to each \( \rho \geq \tilde{p} \) has the same global optimal solution set as problem (1) does.

**Proof:** Consider a local optimal solution \( X^* \) of (1). Then, there exists \( \epsilon' \geq 0 \) such that
\[
f(X) \geq f(X^*) \quad \forall X \in \mathcal{B}(X^*, \epsilon') \cap [\mathbb{R}_{+}^{n \times r} \cap \text{St}(n, r)].
\]
(14)

By Proposition 3.1, there exist \( 0 < \epsilon < 1 \) such that for all \( X \in \mathcal{B}(X^*, \epsilon) \),
\[
\text{dist}(X, \mathbb{R}_{+}^{n \times r} \cap \text{St}(n, r)) \leq \kappa \left[ \text{dist}(X, \mathbb{R}_{+}^{n \times r}) + \text{dist}(X, \text{St}(n, r)) \right],
\]
(15)
where \( \kappa = 2.1 \sqrt{n} + 1 \) if \( r = n \) and otherwise \( \kappa = \frac{10.5r^{3/2}(n-r)}{X_{i,r,j}^*} + 1 \). Set \( \delta = \min(\epsilon', \epsilon)/2 \).

Pick any \( \epsilon \geq 0 \) and let \( \mathcal{F}_\epsilon := \{ X \in \text{St}(n, r) \mid \vartheta(X) = \epsilon \} \). Fix any \( X \in \mathcal{B}(X^*, \delta) \cap \mathcal{F}_\epsilon \). From the last inequality, there exists \( \mathcal{X} \in \mathbb{R}_{+}^{n \times r} \cap \text{St}(n, r) \) with \( \|X - \mathcal{X}\|_F = \text{dist}(X, \mathbb{R}_{+}^{n \times r} \cap \text{St}(n, r)) \) such that
\[
\|X - \mathcal{X}\|_F \leq \kappa \left[ \text{dist}(X, \mathbb{R}_{+}^{n \times r}) + \text{dist}(X, \text{St}(n, r)) \right] \leq \kappa \vartheta(X),
\]
where the last inequality is due to \( X \in \text{St}(n, r) \) implied by \( X \in \mathcal{F}_\epsilon \). By the definition of \( \mathcal{X} \), we have \( \|X - \mathcal{X}\|_F \leq 2\|X - X^*\|_F \leq \epsilon' \). From (14) and the last inequality,
\[
f(X) - f(X^*) \geq f(X) - f(\mathcal{X}) \geq -L\|X - \mathcal{X}\|_F \geq -L\kappa \vartheta(X),
\]
where the second inequality is using the Lipschitz continuity of \( f \) on \( \text{St}(n, r) \). By the arbitrariness of \( \epsilon \geq 0 \), the desired inequality holds. The second part follows by the first part, [24, Proposition 2.1(b)] and the compactness of \( \mathbb{R}_{+}^{n \times r} \cap \text{St}(n, r) \). □
Remark 3.1 (a) In [19, Theorem 4.1], Han and Mangasarian showed that if $\overline{X}$ is globally optimal to all penalty problems associated to $\rho > \overline{\rho}$, then it is globally optimal to the origin problem (1). Clearly, such a condition is very strong and unverifiable. Our global exact penalty result in Theorem 3.1 does not require any condition when $r = n$.

(b) The first part of Theorem 3.1 implies that every local optimizer of (1) is locally optimal to the problem (3) associated to $\rho \geq \kappa L_f$. Conversely, every nonnegative local optimizer of problem (3) associated to $\rho > 0$ is also locally optimal to problem (1).

(c) Observe that the inequalities in Proposition 3.1 also hold if the distance $\text{dist}(X, \mathbb{R}^{n \times r}_+)$ is defined by other norms in $\mathbb{R}^{n \times r}$. Thus, from the proof of Theorem 3.1, the problem (3) is still a global exact penalty of the problem (1) if the penalty term $\vartheta(X)$ is replaced by a distance on other norms to the cone $\mathbb{R}^{n \times r}_+$.

Next we achieve the global exact penalty of (4) under a growth condition of $f$, which is not necessarily stronger than its locally Lipschitz continuity on the solution set.

Theorem 3.2 Suppose that every global (or local) optimal solution of (1) has no zero rows when $r < n$, and that for any global (or local) optimal solution $X^*$ of (1), there exist $\delta' > 0$ and $L' > 0$ such that for all $X \in \mathbb{B}(X^*, \delta') \cap \text{St}(n, r)$ and $\overline{X} \in \text{Proj}_{\text{St}(n, r) \cap \mathbb{R}^{n \times r}}(X)$,

$$ f(X) - f(X^*) \geq -L'\|X - \overline{X}\|_F^2. \tag{16} $$

Then, for every global (or local) optimizer $X^*$ of (1), there exists $\delta > 0$ such that for any $\epsilon \geq 0, \gamma > 0$ and $X \in \mathbb{B}(X^*, \delta) \cap \mathcal{F}_{\gamma, \epsilon}$ where $\mathcal{F}_{\gamma, \epsilon} := \{X \in \text{St}(n, r) \mid e_{\gamma, \epsilon}(X) = \epsilon\}$,

$$ f(X) - f(X^*) + 2\gamma \kappa^2 L' e_{\gamma, \epsilon}(X) \geq 0, $$

and consequently, there exists a threshold $\hat{\rho} > 0$ such that the penalty problem (4) associated to every $\rho \geq \hat{\rho}$ has the same global optimal solution set as problem (1) does.

Proof: Consider a global (or local) optimal solution $X^*$ of (1). Clearly, the inequalities (14) and (15) still hold. Set $\delta := \min(\epsilon', \epsilon, \delta')/2$ where $\epsilon'$ and $\epsilon$ are same as those in (14) and (15). Pick any $\epsilon \geq 0$ and fix any $X \in \mathbb{B}(X^*, \delta) \cap \mathcal{F}_{\gamma, \epsilon}$. From (15) there exists $\overline{X} \in \text{Proj}_{\text{St}(n, r) \cap \mathbb{R}^{n \times r}}(X)$ such that $\|X - \overline{X}\|_F^2 \leq \kappa^2 \text{dist}^2(X, \mathbb{R}^{n \times r}_+)$. Notice that $\|X - X^*\|_F \leq 1/2$ with $X^* \in \mathbb{R}^{n \times r}_+$ and $e_{\gamma, \epsilon}(X) = \frac{1}{2\gamma^2} \|P_{\gamma, \vartheta}(X) - X\|_F^2 + \vartheta(P_{\gamma, \vartheta}(X))$. By using the expression of $P_{\gamma, \vartheta}(X)$ in Lemma 2.1, it is not difficult to check that

$$ e_{\gamma, \epsilon}(X) \geq \frac{1}{2\gamma^2}\|X - \overline{X}\|_F^2. $$

Along with (16), $f(X) - f(X^*) \geq f(X) - f(\overline{X}) \geq -L'\|X - \overline{X}\|_F^2 \geq -2\gamma \kappa^2 L' e_{\gamma, \epsilon}(X). \quad \square$

Remark 3.2 The second assumption of Theorem 3.2 implies that there exist $\delta' > 0$ and $L' > 0$ such that for every $X \in \mathbb{B}(X^*, \delta')$, $[f(X) + \delta_{\text{St}(n, r)}(X)] - [f(X^*) + \delta_{\text{St}(n, r)}(X^*)] \geq -L'\|X - X^*\|_F^2 \geq -L'\|X - X^*\|_F$, which is the calmness of $f + \delta_{\text{St}(n, r)}$ at $X^*$ from below. Thus, the growth condition of $f$ here is stronger than the below calmness of $f + \delta_{\text{St}(n, r)}$.  

14
To close this section, we introduce a kind of stationary points for problem (1). By Corollary 3.1, if $X^*$ is a local optimal solution of (1) and has no zero row when $r < n$, then $0 \in \nabla f(X^*) + N_{\mathbb{R}^n_+^r}(X^*)$. So, we define the following stationary point.

**Definition 3.1** We call $X \in \text{St}(n, r) \cap \mathbb{R}^{n \times r}_+$ a stationary point of problem (1) if

$$0 \in \nabla f(X) + N_{\mathbb{R}^n_+}X + N_{\mathbb{R}^n_+^r}(X);$$

and call $X \in \text{St}(n, r) \cap \mathbb{R}^{n \times r}_+$ a $\tau$-approximate stationary point of problem (1) if

$$\text{dist}(0, \text{Proj}_{TX}(\nabla f(X) + N_{\mathbb{R}^n_+^r}(X))) \leq \tau.$$

By comparing with [17, Theorem 2.4], one can verify that the stationary point defined above is the one defined by [17, Theorem 2.4], but not vice versa. This is because if $X$ is a stationary point defined by [17, Theorem 2.4] and its $i$th row is nonzero, the entries $[\nabla f(X)]_{ij}$ for those $j$ with $X_{ij} = 0$ are free.

### 4 Exact penalty methods for problem (1)

By Theorem 3.1, the problem (3) associated to every $\rho \geq \overline{\rho}$ is equivalent to problem (1) in a global sense; and by Theorem 3.2, under the growth condition (16) of $f$, the problem (4) associated to every $\rho \geq \hat{\rho}$ is also equivalent to (1) in a global sense. However, when solving a single penalty problem (4) (or problem (3)) associated to some $\rho \geq \rho$ with an efficient algorithm, one can not expect a solution of high quality due to its nonconvexity. In view of this, we propose a penalty algorithm by seeking the approximate stationary points for a finite number of penalty problems (4) or (3) with increasing $\rho$.

**Algorithm 4.1** (Penalty method based on $\Theta_{\rho, \gamma}$ (or $\Phi_{\rho}$))

**Initialization:** Input $\gamma > 0, l_{\max} \in \mathbb{N}$ and $\epsilon > 0$. Choose parameters $\rho_{\max} > 0, \sigma > 1$, $\sigma \in (0, 1), \tau \in (0, 1), \rho_0 > 0$ and $\tau_1 > 0$. Select a starting point $X^0 \in \text{St}(n, r)$.

**For** $l = 0, 1, 2, \ldots, l_{\max}$

1. Seek a $\tau_{l+1}$-approximate stationary point $X^{l+1}$ of the following problem

$$\min_{X \in \text{St}(n, r)} \Theta_{\rho_l, \gamma}(X) \quad (\text{resp. } \Phi_{\rho_l}(X)). \quad (17)$$

2. If $\vartheta(X) \leq \epsilon$, stop; else let $\rho_{l+1} = \min\{\sigma \rho_l, \rho_{\max}\}$ and $\tau_{l+1} = \max\{\sigma \tau_l, \tau\}$.

**end**

By Remark 2.2 (b), a local optimal solution $X^*$ of problem (17) with $\Theta_{\rho_l, \gamma}$ (or $\Phi_{\rho_l}$) satisfies $0 \in \nabla \Theta_{\rho_l, \gamma}(X^*)$ (or $0 \in \nabla \Phi_{\rho_l}(X^*)$), which by Corollary 2.1 is equivalent to $\text{Proj}_{TX}(\nabla \Theta_{\rho_l, \gamma}(X^*)) = 0$ (or $0 \in \text{Proj}_{TX}(\nabla f(X^*) + \rho \partial \vartheta(X^*))$). In view of this, we introduce the following stationary point that is same as the limiting critical point in [5].
Definition 4.1 We call $X \in \mathbb{M}$ a stationary point of problem (17) with $\Theta_{\rho, \gamma}$ (resp. $\Phi_{\rho}$) if $\text{Proj}_{X_M}(\nabla \Theta_{\rho, \gamma}(X)) = 0$ (resp. $0 \in \text{Proj}_{X_M}([\nabla f(X) + \rho_1 \partial \vartheta(X)])$), and a $\tau$-approximate critical point of (17) with $\Theta_{\rho, \gamma}$ (resp. $\Phi_{\rho}$) if $\|\text{Proj}_{X_M}(\nabla \Theta_{\rho, \gamma}(X))\|_F \leq \tau$ (resp. there is $H \in \mathbb{R}^{n \times r}$ with $\|H\|_F \leq \tau$ such that $\|\text{dist}(0, \text{grad} f(X) + \text{Proj}_{X_M}(\rho_1 \partial \vartheta(X + H)))\|_F \leq \tau$).

The following result states that once a certain iterate of Algorithm 4.1 is feasible to problem (1), it is necessarily an approximate stationary point of (1).

Proposition 4.1 Let $\{X^l\}_{l=0}^{\text{tmax}}$ be the iterates generated by Algorithm 4.1 with $\Theta_{\rho, \gamma}$ (or $\Phi_{\rho}$ with $\tau_1 \leq |X^l|$, the smallest nonzero entry of $|X^l|$). If some iterate $X^l$ is a feasible point of problem (1), then it is a $\tau_1$-approximate stationary point of (1).

Proof: Let $\{X^l\}_{l=0}^{\text{tmax}}$ be yielded by Algorithm 4.1 with $\Theta_{\rho, \gamma}$. Since $X^l$ is a $\tau_1$-approximate critical point of problem (17) with $\Theta_{\rho, \gamma}$, we have $\|\text{Proj}_{X_M}(\nabla \Theta_{\rho, \gamma}(X^l))\|_F \leq \tau_1$. Notice that $\nabla \Theta_{\rho, \gamma}(X^l) = \nabla f(X^l) + \frac{\varrho}{\gamma}(X^l - \partial \vartheta(X^l))$. By the expression of $\partial \vartheta(X^l)$ in Lemma 2.1, it is easy to check that $X^l - \partial \vartheta(X^l) / \gamma \in \mathcal{N}^{r \times r}(X^l)$. So, $X^l$ is a $\tau_1$-approximate stationary point of (1). Now let $\{X^l\}_{l=0}^{\text{tmax}}$ be generated by Algorithm 4.1 with $\Phi_{\rho}$. Since $X^l$ is a $\tau_1$-approximate critical point of problem (17) with $\Phi_{\rho}$, there exists $H \in \mathbb{R}^{n \times r}$ with $\|H\|_F \leq \tau_1 \leq |X^l|_1$ such that $\|\text{dist}(0, \text{grad} f(X^l) + \text{Proj}_{X_M}(\rho_1 \partial \vartheta(X^l + H)))\|_F \leq \tau_1$. By Lemma 1 in Appendix and its proof, we have $\rho_1 \partial \vartheta(X^l + H) \subseteq \rho_1 \partial \vartheta(X^l) \subseteq \mathcal{N}^{r \times r}(X^l)$. This means that $X^l$ is a $\tau_1$-approximate stationary point of (1).

Motivated by Proposition 4.1, we next develop two PGMs with a line-search strategy to seek an approximate stationary point of subproblem (17) with $\Theta_{\rho, \gamma}$ and $\Phi_{\rho}$. Unless otherwise stated, in the rest of this section, we make the following assumption.

Assumption 4.1 The gradient $\nabla f$ is Lipschitz continuous on $\text{St}(n, r)$ with modulus $L_{\nabla f}$.

4.1 PGM for subproblem (17) with $\Theta_{\rho, \gamma}$

Fix $\rho = \rho_1$. Inspired by the PGM for the problem $\min_{X \in \mathbb{R}^{n \times r}} \Theta_{\rho, \gamma}(X)$, to deal with the manifold constraint in subproblem (17) with $\Theta_{\rho, \gamma}$, we use the following direction

$$V^k := \arg \min_{V \in T_{X^k} \mathcal{M}} \left\{ \langle \nabla \Theta_{\rho, \gamma}(X^k), V \rangle + \frac{1}{2t_k} \|V\|_F^2 \right\} = -t_k \text{grad} \Theta_{\rho, \gamma}(X^k). \quad (18)$$

It is known that the efficiency of the PGM depends much on the choice of the step-size $t_k$, which is determined by the Lipschitz constant $L_{\nabla \Theta_{\rho, \gamma}}$ of $\nabla \Theta_{\rho, \gamma}$. Consider that the latter is usually unknown because the constant $L_{\nabla f}$ is unobtainable for many problems. Inspired by the work [30], we develop a line-search PGM with the direction $V^k$ defined by (18).

Remark 4.1 (a) A good step-size initialization at each outer iteration can greatly reduce the line-search cost. Let $\Delta X^k := X^k - X^{k-1}$ and $\Delta Y^k := \text{grad} \Theta_{\rho, \gamma}(X^k) - \text{grad} \Theta_{\rho, \gamma}(X^{k-1})$. Similar to [30], we initialize the step-size by adopting the Barzilai-Borwein (BB) rule [4]:

$$\max \left\{ \min \left\{ \frac{\|\Delta X^k\|_F^2}{2}, \frac{|\langle \Delta X^k, \Delta Y^k \rangle|}{\|\Delta Y^k\|_F^2} \right\}, t_{\text{min}} \right\}. \quad (19)$$

16
Algorithm 1 (A line-search PG method for $(17)$ with $\Theta_{\rho,\gamma}$)

**Initialization:** Fix $\rho = \rho_l$. Choose $\eta \in (0, 1), 0 < t_{\text{min}} \leq t_{\text{max}}, X^0 \in M$. Set $k := 0$.

**while** some stopping criterion is not satisfied **do**

1. Choose $t_k \in [t_{\text{min}}, t_{\text{max}}]$;
2. **repeat**
   3. Set $V^k := -t_k \text{grad}\Theta_{\rho,\gamma}(X^k)$;
   4. $t_k \leftarrow \eta t_k$;
   5. **until** a certain line-search criterion is satisfied
5. Set $X^{k+1} = R_{X^k}(V^k)$ and let $k \leftarrow k + 1$;
**end while**

**(b)** Choose an integer $m \geq 0$ and a constant $\alpha \in (0, 1)$. We accept the step-size $t_k$ at the outer iteration $k$ if the nonmonotone line-search criterion (see [14, 15]) is satisfied:

$$\Theta_{\rho,\gamma}(X^{k+1}) \leq \max_{j=\max(0,k-m),...,k} \Theta_{\rho,\gamma}(X^j) - \frac{\alpha}{2t_k} \|V^k\|^2_F. \quad (20)$$

From Lemma 4.1 below, it follows that such a line-search criterion is well defined.

**(c)** From Definition 4.1 and the expression of $V^k$, when $\|V^k\|^2_F \leq \tau$, $X^k$ is at worst a $\tau/t_{\text{min}}$-approximate critical point of problem $(17)$ with $\Theta_{\rho,\gamma}$. By this, our numerical tests in Section 5 are conducted with $\|V^k\|^2_F \leq \text{tol}$ as a stopping criterion of Algorithm 1.

**Lemma 4.1** For each integer $k \geq 0$, the line-search criterion $(20)$ with $m = 0$ is satisfied whenever $t_k \leq \frac{1}{2c_{\rho,\gamma} + \tau^2 L_{\rho,\gamma}}$, where $c_{\rho,\gamma} := \max_{Z \in \text{St}(n,r)} \|\nabla\Theta_{\rho,\gamma}(Z)\|_F$.

**Proof:** Notice that the gradient $\nabla\Theta_{\rho,\gamma}$ is Lipschitz continuous on the set $\text{St}(n, r)$ with constant $L_{\rho,\gamma}$ by Assumption 4.1. From the descent lemma, we have

$$\Theta_{\rho,\gamma}(X^{k+1}) - \Theta_{\rho,\gamma}(X^k) \leq \langle \nabla\Theta_{\rho,\gamma}(X^k), X^{k+1} - X^k \rangle + (L_{\rho,\gamma}/2)\|X^{k+1} - X^k\|^2_F.$$

Notice that $X^{k+1} = R_{X^k}(V^k)$. By Lemma 2.2, $\|X^{k+1} - X^k\|_F \leq c_1 \|V^k\|_F$, and moreover,

$$\langle \nabla\Theta_{\rho,\gamma}(X^k), X^{k+1} - X^k \rangle = \langle \nabla\Theta_{\rho,\gamma}(X^k), X^{k+1} - X^k - V^k \rangle + \langle \nabla\Theta_{\rho,\gamma}(X^k), V^k \rangle$$
$$\leq c_2 \|\nabla\Theta_{\rho,\gamma}(X^k)\|_F \|V^k\|^2_F + \|\text{grad}\Theta_{\rho,\gamma}(X^k), V^k \|$$
$$\leq c_2 \|\nabla\Theta_{\rho,\gamma}(X^k)\|_F \|V^k\|^2_F - \frac{1}{2t_k} \|V^k\|^2_F$$
$$\leq c_2 c_{\rho,\gamma} \|V^k\|^2_F - \frac{1}{2t_k} \|V^k\|^2_F.$$
where the second inequality is also using (18). From the last two inequalities, we obtain
\[
\Theta_{\rho,\gamma}(X^{k+1}) - \Theta_{\rho,\gamma}(X^k) \leq -\left(\frac{1}{2l_k} - c_2^\rho c_{\rho,\gamma} - \frac{1}{2} c_1^2 L_{\rho,\gamma}\right) \|V^k\|_F^2,
\]
which implies that the desired result holds. The proof is completed. \qed

**Theorem 4.1** Let \(\{(X^k, V^k)\}_{k=0}^\infty\) be generated by Algorithm 1 with criterion (20). Then,

(i) the sequences \(\{\Theta_{\rho,\gamma}(X^k)\}_{k=0}^\infty\) and \(\{V^k\}_{k=0}^\infty\) are convergent, and \(\lim_{k \to \infty} V^k = 0\);

(ii) \(\{X^k\}_{k=0}^\infty\) is bounded and every cluster point is a stationary point of (17) with \(\Theta_{\rho,\gamma}\).

**Proof:** (i) For each integer \(k \geq 0\), by letting \(\ell(k) := \arg\max_{j=\max(0,m-1),...,k} \Theta_{\rho,\gamma}(X^j)\), the acceptance criterion (20) is rewritten as \(\Theta_{\rho,\gamma}(X^{k+1}) \leq \Theta_{\rho,\gamma}(X^{k+1}) - \frac{\alpha}{2l_k} \|V^k\|_F^2\). This implies that the sequence \(\{\Theta_{\rho,\gamma}(X^{\ell(k)})\}_{k=0}^\infty\) is monotonically decreasing because
\[
\Theta_{\rho,\gamma}(X^{\ell(k+1)}) = \max \left\{ \max_{j=1,...,\min(m,k+1)} \Theta_{\rho,\gamma}(X^{k+1-j}), \Theta_{\rho,\gamma}(X^{k+1}) \right\} \\
\leq \max \left\{ \Theta_{\rho,\gamma}(X^{\ell(k)}), \Theta_{\rho,\gamma}(X^{\ell(k)}) - \frac{\alpha}{2l_k} \|V^k\|_F^2 \right\} \leq \Theta_{\rho,\gamma}(X^{\ell(k)}).
\]
Together with the lower boundedness of \(\Theta_{\rho,\gamma}\) on \(\text{St}(n, r)\), the sequence \(\{\Theta_{\rho,\gamma}(X^{\ell(k)})\}_{k=0}^\infty\) is convergent, say, \(\lim_{k \to \infty} \Theta_{\rho,\gamma}(X^{\ell(k)}) = \varphi_{\rho,\gamma}^\star\). In addition, by following the same arguments as those for [30, Lemma 4], we can obtain \(\lim_{k \to \infty} V^{\ell(k)-j} = 0\) for all \(j \geq 1\). Notice that \(\ell(k)\) is one of the indices \(k-m, \ldots, k\). Therefore, \(k-(m+1) = \ell(k)-j\) for some \(j \in \{1, 2, \ldots, m+1\}\). Consequently, \(0 = \lim_{k \to \infty} V^{k-(m+1)} = \lim_{k \to \infty} V^k\). Notice that \(X^{\ell(k)} = R_{X^{\ell(k)-1}}(V^{\ell(k)-1})\) for each integer \(k \geq 0\). From \(\lim_{k \to \infty} V^{\ell(k)-j} = 0\) and Remark 2.1, for all \(k\) large enough, \(X^{\ell(k)} = X^{\ell(k)-1} + V^{\ell(k)-1} + o(\|V^{\ell(k)-1}\|_F)\), and hence
\[
X^{\ell(k)} = X^{k-(m+1)} + \sum_{j=1}^{\ell(k)-(m+1)} \left[ V^{\ell(k)-j} + o(\|V^{\ell(k)-j}\|_F) \right].
\]
Thus, \(\lim_{k \to \infty} (X^{\ell(k)} - X^{k-(m+1)} = 0\). Along with the continuity of \(\Theta_{\rho,\gamma}\), it follows that
\(\lim_{k \to \infty} \Theta_{\rho,\gamma}(X^k) = \lim_{k \to \infty} \Theta_{\rho,\gamma}(X^{k-m-1}) = \lim_{k \to \infty} \Theta_{\rho,\gamma}(X^{\ell(k)}) = \varphi_{\rho,\gamma}^\star\).

(ii) The boundedness of \(\{X^k\}_{k=0}^\infty\) is immediate, since \(\{X^k\}_{k=0}^\infty \subset \text{St}(n, r)\). Let \(\overline{X}\) be an arbitrary cluster point of \(\{X^k\}\). Then, there exists an index set \(K \subset \mathbb{N}\) such that \(\lim_{k \to \infty} X^k = \overline{X}\). Together with part (i) and \(t_k \geq t_{\min} > 0\), from equality (5) we have
\[
0 = \lim_{k \to \infty} \grad \Theta_{\rho,\gamma}(X^k) = \grad \Theta_{\rho,\gamma}(\overline{X}).
\]
Recall that \(\rho = \rho_t\). This shows that \(\overline{X}\) is a stationary point of (17) with \(\Theta_{\rho,\gamma}\). \qed

**Theorem 4.2** Let \(\overline{X}\) be a cluster point of the sequence \(\{X^k\}_{k=0}^\infty\) yielded by Algorithm 1 from \(X^0\) with the acceptance criterion (20) for \(m = 0\). If \(f\) is definable in an o-minimal structure \(\Theta\) over \(\mathbb{R}\), then \(\{X^k\}_{k=0}^\infty\) converges to \(\overline{X}\), and moreover, there exists \(k_0 \in \mathbb{N}\) such that \(\sum_{k=k_0}^\infty \|X^{k+1} - X^k\|_F < \infty\).
Proof: Since $\vartheta(X) = \min_{Z \in \mathbb{R}^{n \times r}} \|Z - X\|_1$, it is a semi-algebraic function on $\mathbb{R}^{n \times r}$, which means that $e_\gamma \vartheta$ is a semi-algebraic function on $\mathbb{R}^{n \times r}$. Since $\delta_M$ is a semi-algebraic function, $e_\gamma \vartheta + \delta_M$ is a semi-algebraic function on $\mathbb{R}^{n \times r}$ and hence is definable in the o-minimal structure $O$. Since $f$ is definable in the o-minimal structure $O$, $\hat{\Theta}_{\rho, \gamma} = f + \rho e_\gamma \vartheta + \delta_M$ is definable in the o-minimal structure $O$ by [2, Section 4.3]. By [2, Theorem 4.1], $\hat{\Theta}_{\rho, \gamma}$ has the KL property at $\overline{\mathbb{X}}$. Next we argue that $\{X^k\}_{k=0}^\infty$ satisfies condition H1-H3 of Lemma 2.5 for $h = \hat{\Theta}_{\rho, \gamma}$. First, condition H1 is satisfied because by Lemma 4.1 and the expression of $\hat{\Theta}_{\rho, \gamma}$, for each $k \in \mathbb{N}$, $\hat{\Theta}_{\rho, \gamma}(X^{k+1}) + \frac{\alpha}{1 + \max_{\mu t}} \|V^k\|^2_F \leq \hat{\Theta}_{\rho, \gamma}(X^k)$.

To argue that condition H2 is satisfied, let $W^k = \text{grad} \Theta_{\rho, \gamma}(X^k)$ for each $k$. Noting that $\partial \hat{\Theta}_{\rho, \gamma}(X^k) = \nabla \Theta_{\rho, \gamma}(X^k) + N_{X^k}M$, we have $W^k \in \text{Proj}_{T_{X^k}M}(\partial \hat{\Theta}_{\rho, \gamma}(X^k)) \subseteq \partial \hat{\Theta}_{\rho, \gamma}(X^k)$, where the inclusion is by Lemma 2.3. In addition, from the iterate steps of Algorithm 1,

$$\|X^{k+1} - X^k\|_F = \|R_{X^k}(V^k) - (X^k + V^k) + V^k\|_F.$$  

Since $\lim_{k \to \infty} V^k = 0$, by Remark 2.1 we have $\|R_{X^k}(V^k) - (X^k + V^k)\|_F = o(\|V^k\|_F)$. From the last equality, there exist $k_0 \in \mathbb{N}$ and a constant $\mu \in (0, 1)$ such that

$$\|X^{k+1} - X^k\|_F \geq \mu \|V^k\|_F \quad \text{for all } k \geq k_0.$$  

Together with $V^k = -t_k \text{grad} \Theta_{\rho, \gamma}(X^k) = -t_k W^k$, we have $\|W^k\|_F \leq \frac{1}{\mu} \|X^{k+1} - X^k\|_F$. Thus, condition H2 is satisfied, since for each $k \geq k_0$ there is $W^k \in \partial \hat{\Theta}_{\rho, \gamma}(X^k)$ such that $\|W^k\|_F \leq \frac{1}{\mu} \|X^{k+1} - X^k\|_F$. In addition, condition H3 is automatically satisfied because $\hat{\Theta}_{\rho, \gamma}$ is continuous relative to $M$ and $\{X^k\} \subseteq M$. Since condition H1-H3 of Lemma 2.5 are satisfied for every $k \geq k_0$, the conclusion follows by this lemma.  

4.2 PGM for subproblem (17) with $\Phi_p$

Fix $\rho = \rho_l$. To deal with the manifold constraint in subproblem (17) with $\Phi_p$, it is natural to require $V^k \in T_{X^k}M$. Inspired by the work [8], we seek the descent direction $V^k$ via

$$V^k := \arg \min_{V \in T_{X^k}M} \left\{ \langle \nabla f(X^k), V \rangle + \frac{1}{2t_k} \|V\|^2_F + \rho \vartheta(X^k + V) \right\}. \quad (21)$$

Such a direction is also used by [21, Algorithm 2] to solve the Riemannian proximal map. Our line-search PG method for solving subproblem (17) with $\Phi_p$ is described as follows.
Algorithm 2 (A line-search PGM for (17) with $\Phi_p$)  

**Initialization:** Fix $\rho = \rho_l$. Choose $\eta \in (0, 1), 0 < t_{\min} \leq t_{\max}, X^0 \in M$. Set $k := 0$.

**while** some stopping criterion is not satisfied **do**

1. Choose $t_k \in [t_{\min}, t_{\max}]$ by the BB rule;

2. **repeat**

3. to achieve $V^k$ with Algorithm 3 in Appendix;

4. $t_k \leftarrow \eta t_k$;

5. **until** a certain line-search criterion is satisfied

6. Set $X^{k+1} = R_{X_k}(V^k)$ and let $k \leftarrow k + 1$;

**end while**

Remark 4.2 (a) Algorithm 2 is different from [8, Algorithm 2] in two aspects: (1) the strongly convex program (21) is solved with a dual regularized semismooth Newton method; see Algorithm 3 in Appendix, which is equivalent to the regularized semismooth Newton method applied to the reduced KKT system of (21) without primal variables; (2) the step-size $t_k$ is initialized at the outer iteration $k$ by the BB rule (19) with $\Delta X^k := X^k - X^{k-1}$ and $\Delta Y^k := \nabla f(X^k) - \nabla f(X^{k-1})$, and the step-size $t_k$ is accepted whenever

$$
\Phi_p(X^{k+1}) \leq \max_{j = \max(0, k - m), \ldots, k} \Phi_p(X^j) - \frac{\alpha}{2t_k} \|V^k\|_F^2,
$$

(22)

is satisfied for a given integer $m \geq 0$, where $\alpha \in (0, 1)$ is a constant, and Lemma 4.2 below implies that such a nonmonotone line-search rule is well defined.

(b) Notice that $V^k$ is the unique optimal solution of the strongly convex program (21) iff

$$
0 \in \nabla f(X^k) + t_k^{-1} V^k + \text{Proj}_{X_k^lM}(\rho \partial \partial(X^k + V^k)).
$$

So, when $\|V^k\|_F \leq \min(\tau t_k, \tau)$ for some $\tau > 0$, $X^k$ is a $\tau$-approximate stationary point of (17) with $\Phi_p$. By this, we suggest $\|V^k\|_F \leq \text{tol}$ as a stopping rule of Algorithm 2.

Lemma 4.2 For each integer $k \geq 0$, the line-search criterion (22) with $m = 0$ is satisfied whenever $t_k \leq \frac{2 - \alpha}{2\eta c\eta + c_l^2 L_{\eta f} + 2\eta \rho \sqrt{m}}$, where $c\eta = \max_{Z \in \text{St}(n, r)} \|\nabla f(Z)\|_F$.

**Proof:** Fix an integer $k \geq 0$. Let $\Xi_k$ denote the objective function of (21). From the optimality condition of (21), it follows that $0 \in \partial \Xi_k + \delta_{X_k^lM}(V^k) \subseteq \partial \Xi_k(V^k) + N_{X_k^lM}$. So, $0 \in \text{Proj}_{X_k^lM}(\partial \Xi_k(V^k))$ and there exists $Z^k \in \partial \Xi_k(V^k)$ such that $\text{Proj}_{X_k^lM}(Z^k) = 0$. Consequently, $\langle Z^k, V^k \rangle = 0$. Together with the strong convexity of $\Xi_k$, we have

$$
\Xi_k(0) \geq \Xi_k(V^k) - \langle Z^k, V^k \rangle + \frac{1}{2t_k} \|V^k\|_F^2 = \Xi_k(V^k) + \frac{1}{2t_k} \|V^k\|_F^2.
$$

20
In addition, from Assumption 4.1 and the descent lemma, it follows that
\[
f(X^{k+1}) - f(X^k) \leq \langle \nabla f(X^k), R_{X_k}(V^k) - X^k \rangle + \frac{L \sqrt{f}}{2} \|R_{X_k}(V^k) - X^k\|^2_F
\]
\[
= \langle \nabla f(X^k), R_{X_k}(V^k) - X^k - V^k \rangle + \langle \nabla f(X^k), V^k \rangle + \frac{L \sqrt{f}}{2} \|R_{X_k}(V^k) - X^k\|^2_F
\]
\[
\leq c_2 \|\nabla f(X^k)\|_F \|V^k\|^2_F + \langle \nabla f(X^k), V^k \rangle + \frac{c_2^2 L \sqrt{f}}{2} \|V^k\|^2_F,
\]
where the second inequality is by Lemma 2.2. Now, from the last two inequalities, the expression of \(\Phi_p\), and the Lipschitz continuity of \(\vartheta\), it follows that
\[
\Phi_p(X^{k+1}) - \Phi_p(X^k) = f(X^{k+1}) - f(X^k) - \langle \nabla f(X^k), V^k \rangle - \frac{1}{2t_k} \|V^k\|^2_F
\]
\[
+ \Xi_k(V^k) - \Xi_k(0) + \rho [\vartheta(X^{k+1}) - \vartheta(X^k + V^k)]
\]
\[
\leq [c_2 \|\nabla f(X^k)\|_F + 0.5c_2^2 L \sqrt{f} - t_k^{-1}] \|V^k\|^2_F + \rho \sqrt{\alpha_r} \|X^k + V^k\|^2_F
\]
\[
\leq [c_2 c \sqrt{f} + 0.5c_2^2 L \sqrt{f} - t_k^{-1}] \|V^k\|^2_F + c_2 \rho \sqrt{\alpha_r} \|V^k\|^2_F
\]
\[
= -(t_k^{-1} - c_2 c \sqrt{f} - 0.5c_2^2 L \sqrt{f} - c_2 \rho \sqrt{\alpha_r}) \|V^k\|^2_F,
\]
which shows that the desired result holds. The proof is then completed. \(\square\)

**Theorem 4.3** Let \(\{X^k, V^k\}_{k=0}^\infty\) be generated by Algorithm 2 with criterion (22). Then,

(i) the sequences \(\{\Phi_p(X^k)\}_{k=0}^\infty\) and \(\{V^k\}_{k=0}^\infty\) are convergent, and \(\lim_{k \to \infty} V^k = 0\);

(ii) \(\{X^k\}_{k=0}^\infty\) is bounded and every cluster point is a stationary point of (17) with \(\Phi_{p_t}\).

**Proof:** The proof of part (i) is similar to that of [30, Lemma 4]. So, it suffices to argue that every cluster point of \(\{X^k\}_{k=0}^\infty\) is a critical point of (17) with \(\Phi_p\). Pick any cluster point \(\overline{X}\) of \(\{X^k\}\). Then, there exists an index set \(K \subset \mathbb{N}\) such that \(\lim_{k \to \infty} X^k = \overline{X}\). From the optimality of \(V^k\) to the problem (21), it follows that
\[
0 \in \mathrm{Proj}_{X \times M}(\partial \Xi_k(V^k)) = \mathrm{Proj}_{X \times M}(\nabla f(X^k) + t_k^{-1} V^k + \rho \nabla \vartheta(X^k + V^k)). \tag{23}
\]
Since the mapping \(\partial \vartheta : \mathbb{R}^{n \times r} \to \mathbb{R}^{n \times r}\) is locally bounded by [28, Theorem 9.13], from [28, Proposition 5.52] the mapping \(F(X) := \mathrm{Proj}_{X \times M}(\rho \nabla \vartheta(X))\) for \(X \in \mathbb{R}^{n \times r}\) is outer semicontinuous. By using equality (5), \(\lim_{k \to \infty} V^k = 0\) and \(t_k^{-1} \leq t_{\text{min}}^{-1}\), we obtain from the inclusion (23) that \(-\mathrm{Proj}_{X \times M}(\nabla f(\overline{X})) \in \mathrm{Proj}_{X \times M}(\rho \nabla \vartheta(\overline{X}))\). By Definition 4.1 and \(\rho = p_t\), we conclude that \(\overline{X}\) is a critical point of problem (17) with \(\Phi_{p_t}\). \(\square\)

For the sequence \(\{X^k\}_{k=0}^\infty\) generated by Algorithm 2, we cannot achieve its convergence because it does not satisfy condition H2 of Lemma 2.5. The reason is that by the optimal condition of (21), \(W^k := t_k^{-1} V^k \in \mathrm{Proj}_{X \times M}(\nabla f(X^k) + \rho \nabla \vartheta(X^k + V^k))\), but we cannot have \(W^k \in \mathrm{Proj}_{X \times M}(\nabla f(X^k) + \rho \nabla \vartheta(X^k))\), since there is lack of a constant \(\delta > 0\).
such that for all $X \in \text{St}(n, r)$ and all $H \in \mathbb{R}^{n \times r}$ with $\|H\|_F \leq \delta$, $\partial \vartheta(X + H) \subseteq \partial \vartheta(X)$. To overcome this difficulty, Huang and Wei [21] achieved the direction $V^k$ by solving

$$\min_{V \in \Gamma_{X^k}M} \left\{ (\nabla f(X^k), V) + \frac{1}{2t_k} \|V\|_F^2 + \rho \vartheta(R_{X^k}(V)) \right\}.$$ 

As we mentioned in the introduction, their Algorithm 2 yields such a direction by solving a Lyapunov equation as well as a subproblem similar to (21) in each iterate, and its computation cost in each iterate is at least twice that of our Algorithm 3 to solve (21).

5 Numerical experiments

From Section 4.1 and 4.2, Algorithm 1 has a better convergence result than Algorithm 2 does. Then, it is natural to ask whether Algorithm 4.1 armed with Algorithm 1 (SEPPG for short) is more efficient than Algorithm 4.1 armed with Algorithm 2 (EPPGSN for short). In this section, we shall provide an affirmative answer by testing QAP and graph matching problems with them, and comparing their performance with that of the ALM used in [29]. All the tests are performed in MATLAB on a workstation running on 64-bit Windows System with an Intel(R) Xeon(R) W-2245 CPU 3.90GHz and 128 GB RAM.

5.1 Augmented Lagrangian method

For a given $\mu > 0$, let $L_\mu(X, \Lambda)$ be the augmented Lagrangian function of (1) defined as

$$L_\mu(X, \Lambda) := f(X) + \frac{\mu}{2} \|\min(0, X - \mu^{-1} \Lambda)\|_F^2 - \frac{1}{2\mu} \|\Lambda\|_F^2.$$ 

The ALM proposed in [29] for solving problem (1) consists of the following iterates:

$$\begin{align*}
X^{k+1} &\in \arg\min_{X \in \text{St}(n, r)} L_{\mu_k}(X, \Lambda^k); \\
\Lambda^{k+1} &\leftarrow \max(\Lambda^k - \mu_k X^{k+1}, 0); \\
\mu_{k+1} &= 1.2 \mu_k.
\end{align*}$$

The iterates (24a)-(24c) are actually the ALM for the equivalent reformulation of (1):

$$\min_{X, Y \in \mathbb{R}^{n \times r}} \left\{ f(X) + \delta_{\mathbb{R}^{n \times r}_+}(Y) \right\} \text{ s.t. } X - Y = 0, X \in \text{St}(n, r).$$

Indeed, for a given $\mu > 0$, the augmented Lagrangian function for (25) takes the form of

$$L_\mu(X, Y; \Lambda) = f(X) + \delta_{\text{St}(n, r)}(X) + \delta_{\mathbb{R}^{n \times r}_+}(Y) - \langle \Lambda, X - Y \rangle + \frac{\mu}{2} \|X - Y\|_F^2,$$

and the ALM for problem (25) consists of the following iterates:

$$\begin{align*}
(X^{k+1}, Y^{k+1}) &\in \arg\min_{X, Y \in \mathbb{R}^{n \times r}} L_{\mu_k}(X, Y; \Lambda^k); \\
\Lambda^{k+1} &\leftarrow \Lambda^k - \mu_k (X^{k+1} - Y^{k+1}); \\
\mu_{k+1} &= \rho \mu_k \text{ for some } \rho > 1.
\end{align*}$$
Notice that \( \min_{X,Y \in \mathbb{R}^{n \times r}} L_{\mu_k}(X,Y; \Lambda^k) = \min_{X \in \mathbb{R}^{n \times r}} [\min_{Y \in \mathbb{R}^{n \times r}} L_{\mu_k}(X,Y; \Lambda^k)] \). After an elementary calculation, we have \( \min_{Y \in \mathbb{R}^{n \times r}} L_{\mu_k}(X,Y; \Lambda^k) = L_{\mu_k}(X, \Lambda^k) + \delta_{S\{n,r\}}(X) \) with the unique optimal solution \( Y^*_k(X) = \max(0, X - \mu_{k-1}^{-1} \Lambda^k) \) and \( Y^{k+1} = Y^*_k(X^{k+1}) \). Thus, the iterates (24a)-(24c) are precisely those (26a)-(26c) with \( g = 1.2 \). For the implementation of the above ALM, the interested reader is referred to the toolbox OptM; see https://github.com/optsuite/OptM.

### 5.2 Implementation of exact penalty methods

First we focus on the implementation of Algorithm 4.1. The parameter \( \rho_0 \) is chosen by \( \| f(X^0) \| / \vartheta(X^0) \), where \( X^0 \) is given in the experiments, and other parameters are chosen as follows:

\( \gamma = 0.05, \ t_{\text{max}} = 2000, \ \epsilon = 10^{-6}, \ \tau = 10^{-3}, \ \rho_{\text{max}} = 10^{10}, \ \sigma = 1.1, \ \vartheta = 0.99, \ \tau_1 = 0.005 \).

When solving the subproblem (17) associated to \( \rho_1 \) and 2, we use myQR(\( |X^{l-1}| + \varepsilon_{l-1}\text{randn}(n), n \)) with the perturbation \( \varepsilon_{l-1} := 10^{-2} \min(1, \vartheta(X^{l-1}) \) as the starting point. In addition, we terminate Algorithm 4.1 whenever \( \vartheta(X^k) \leq \epsilon \) or \( \frac{|f(X^k)-f(X^{k-9})|}{1+|f(X^k)|} \leq 10^{-8} \).

Next we provide the implementation details of Algorithm 1 and 2. We choose the retraction of the QR decomposition and compute it with myQR in the toolbox OptM. The nonmonotone line-search rules (20) and (22) with \( m = 5 \) and \( \alpha = 10^{-4} \) are used to search the step-size. The parameters \( \eta = 0.3 \) and 0.1 are respectively used for Algorithm 1 and 2.

In addition, we choose an appropriate \( t_{\text{max}} \) in terms of the ratio \( \frac{|f(X^0)|}{\vartheta(X^0)} \) for Algorithm 1 and 2, and set \( t_{\text{min}} = 10^{-12} \). By Remark 4.1 (c), when applying Algorithm 1 and 2 to solving the subproblem (17) with \( \rho_1 \), we terminate them at \( X^k \) once \( k \geq 100 \) or \( \|V^k\|_F \leq \tau_1 \) or \( \|V^k\|_F \leq 5\tau_1 \) but \( \frac{|F(X^k)-F(X^{k-9})|}{1+|F(X^k)|} \leq 10^{-8} \), where \( F = \Theta_{\rho,\gamma} \) or \( \Phi_{\rho} \).

### 5.3 Numerical results on QAPs

Since every feasible point of a QAP has all binary entries, problem (2) is equivalent to

\[
\min_{X \in \mathbb{R}^{n \times n}} \left\{ (A \odot X)B(X \odot X)^T \right\} \quad \text{s.t.} \quad X^T X = I_n, \ X \geq 0 \tag{27}
\]

where “\( \odot \)” denotes the Hadamard product. The lifted reformulation (27) is superior to the origin problem (1) when \( A \) and \( B \) are nonnegative matrices, for example, those instances from QAPLIB, because the objective function of (27) is nonnegative at any \( X \in \mathbb{R}^{n \times r} \) and will account for a larger proportion in the penalty function \( \Theta_{\rho,\gamma} \) or \( \Phi_{\rho} \), compared with the objective function of (1) in the two penalty functions. The lifted problem (27) was first adopted by Wen and Yin [29] when applying the ALM to the QAPs. In this part, we apply SEPPG and EPPGSN to solving (27), and compare their performance with that of ALM proposed in [29]. The parameter \( \mu_0 = 10 \) is used for ALM, which is more stable than using \( \mu_0 = 1 \) and \( \mu_0 = 0.1 \). We measure the quality of a solution yielded by a method in terms of the nonnegative infeasibility and the relative gap, defined by

\[
\text{Ninf} := \vartheta(X^f) \quad \text{and} \quad \text{relgap} := \left[ \frac{f(X^f) - \text{Best}}{\text{Best}} \times 100 \right] \% ,
\]

23
where $X^f$ is the output of a solver and Best is the best upper bound of an instance. During the tests, we find that for all instances the orthogonal infeasibilities yielded by three methods are all less than $10^{-10}$, so we will not report them. Since our purpose is to verify if SEPPG is more efficient than EPPGSN and ALM, rather than to provide an exact method for QAPs like [16], no rounding technique is imposed on the solutions.

We test three solvers on the 134 instances in QAPLIB except esc16f and tai10b because the matrix $A$ for esc16f is zero and the best bound for tai10b is not provided. Since the quality of the returned solution for three solvers depends on the starting point, we run each QAPLIB instance from 100 different starting points generated randomly by the MATLAB command orth(randn($n$, $r$)). Table 1 and 2 report the levels of the average minimum and median gaps of three solvers for the 134 instances. We see that SEPPG yields better minimum and median gaps than EPPGSN and ALM do, and for half of instances its average minimum and median gaps are respectively less than 0.5% and 5%. We also implemented ALM by solving each subproblem (24a) with a starting point myQR($|X^k|+\varepsilon_k$randn($n$), $n$) (ALMre for short), where the perturbation $\varepsilon_k$ is defined as in Subsection 5.2. The last line of Table 1 and 2 show that ALMre yields better results than ALM do, and the performance of SEPPG is comparable with that of ALMre.

Table 1: Levels of the average min gaps of four solvers on 134 QAPLIB instances

| Min Gap| 0.0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 | 2.0 | 3.0 | 4.0 |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| SEPPG  | 40  | 47  | 54  | 60  | 63  | 67  | 71  | 79  | 84  | 93  | 110 | 119 | 121 |     |
| EPPGSN | 29  | 35  | 39  | 41  | 45  | 53  | 60  | 60  | 65  | 69  | 71  | 99  | 113 | 115 |
| ALM    | 29  | 37  | 38  | 42  | 45  | 48  | 51  | 57  | 60  | 68  | 74  | 99  | 109 | 115 |
| ALMre  | 42  | 51  | 55  | 57  | 62  | 64  | 68  | 76  | 80  | 86  | 91  | 102 | 114 | 117 |

Table 2: Levels of the average med gaps of four solvers on 134 QAPLIB instances

| Med Gap| 0.3 | 0.5 | 0.7 | 1.0 | 3.0 | 5.0 | 7.0 | 10.0 | 15.0 | 20.0 | 25.0 | 30.0 | 40.0 |
|--------|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|
| SEPPG  | 3   | 6   | 11  | 22  | 53  | 84  | 94  | 100  | 111  | 115  | 120  | 122  | 123  |
| EPPGSN | 1   | 3   | 8   | 13  | 45  | 76  | 87  | 97   | 107  | 114  | 120  | 122  | 123  |
| ALM    | 1   | 4   | 10  | 19  | 43  | 68  | 81  | 90   | 101  | 107  | 111  | 117  | 121  |
| ALMre  | 1   | 5   | 12  | 20  | 51  | 79  | 88  | 94   | 103  | 109  | 115  | 118  | 122  |

We also make a pairwise comparison for four solvers as in [10, 16]. Given a QAPLIB instance, solver $i$ is called the winner if the average minimum gap (or median gap) obtained by the solver is the smallest one among them. The second line of Table 3 reports the winners of each solver for the 134 QAPLIB instances. From the third line, we perform pairwise comparisons of the solvers. The last line indicates that ALMre is superior to SEPPG, EPPGSN and ALM whether in terms of min or median gaps.

Table 4 reports the average results of four solvers on the 21 instances with $n \geq 80$. We list the average median gaps, the average CPU time in second, and the average nonnegative infeasibility for solving each instance. One can see that EPPGSN yields the
Table 3: Pairwise comparison of four solvers

| Min Gap | SEPPG | EPPGSN | ALM | ALMre |
|---------|-------|--------|-----|-------|
| winners | 70    | 46     | 37  | **81** |
| SEPPG   | –     | 103:63 | 108:57 | 79:92 |
| EPPGSN  | 63:103| –      | 81:76 | **53:112** |
| ALM     | 57:108| 76:81  | –   | 48:121|
| ALMre   | **92:79** | **112:53** | **121:48** | – |

Table 4: Numerical results of three solvers on 21 QAPLIB instances with \(n \geq 80\)

| Name   | SEPPG | SNCG    | ALM     | ALMre |
|--------|-------|---------|---------|-------|
| med gap| Ninf  | med time| Ninf  | med time| Ninf |
| esc128 | 5.719 | 7.47    | 2.4e-6 | 10.500 | 33.69 | 2.4e-13 | 26.312 | 0.47 | 2.3e-6 | 25.938 | 0.68 | 3.0e-5 |
| lipa80a| 0.836 | 2.90    | 2.2e-6 | 1.201  | 14.86  | 6.2e-9 | 0.791  | 1.22 | 6.4e-6 |
| lipa80b| 19.911| 0.92    | 2.2e-6 | 21.184 | 4.56   | 5.8e-17 | 20.565 | 1.36 | 7.0e-7 | 20.226 | 1.55 | 2.7e-6 |
| lipa90a| 1.714 | 0.96    | 3.6e-6 | 2.270  | 2.46   | 4.2e-17 | 1.884  | 0.83 | 9.0e-7 | 1.747  | 0.82 | 1.8e-5 |
| lipa90b| 19.911| 0.92    | 2.2e-6 | 21.184 | 4.56   | 5.8e-17 | 20.565 | 1.36 | 7.0e-7 | 20.226 | 1.55 | 2.7e-6 |
| sko81  | 5.117 | 1.23    | 2.9e-6 | 2.240  | 2.89   | 3.8e-17 | 1.729  | 0.96 | 1.0e-6 | 1.634  | 0.99 | 1.5e-5 |
| sko90  | 1.490 | 1.61    | 2.7e-6 | 2.043  | 4.11   | 1.7e-16 | 1.683  | 1.32 | 1.2e-5 | 1.620  | 1.37 | 2.0e-5 |
| sko100a| 1.508 | 1.57    | 3.1e-6 | 2.023  | 4.06   | 4.8e-15 | 1.615  | 1.31 | 1.2e-5 | 1.547  | 1.35 | 2.0e-5 |
| sko100c| 1.692 | 1.62    | 2.9e-6 | 2.248  | 4.18   | 8.1e-17 | 1.797  | 1.30 | 1.2e-5 | 1.731  | 1.37 | 2.0e-5 |
| sko100d| 1.548 | 1.58    | 3.0e-6 | 2.092  | 4.08   | 4.8e-10 | 1.723  | 1.31 | 1.1e-5 | 1.649  | 1.36 | 2.1e-5 |
| sko100e| 1.713 | 1.54    | 2.8e-6 | 2.301  | 4.14   | 1.2e-16 | 1.861  | 1.33 | 9.8e-7 | 1.790  | 1.35 | 2.2e-5 |
| sko100f| 1.514 | 1.59    | 2.8e-6 | 1.994  | 4.11   | 6.1e-17 | 1.623  | 1.31 | 1.1e-5 | 1.542  | 1.35 | 1.8e-5 |
| tai80a | 3.030 | 0.87    | 2.5e-6 | 3.484  | 3.23   | 5.3e-17 | 3.293  | 1.32 | 5.4e-7 | 2.876  | 1.50 | 2.1e-5 |
| tai80b | 4.639 | 2.75    | 3.4e-6 | 4.738  | 8.07   | 4.7e-13 | 5.152  | 2.36 | 1.5e-6 | 4.975  | 3.94 | 1.2e-5 |
| tai100a| 2.801 | 1.43    | 3.1e-6 | 3.251  | 4.96   | 5.6e-17 | 3.032  | 2.20 | 4.6e-7 | **2.743** | 2.43 | 2.2e-5 |
| tai100b| 4.138 | 4.89    | 3.3e-6 | 4.297  | 13.26  | 4.3e-17 | 4.397  | 3.93 | 1.5e-6 | 4.208  | 6.55 | 8.9e-6 |
| tai150b| 2.970 | 8.56    | 2.8e-6 | **2.807** | 23.30 | 3.2e-16 | 2.928  | 7.30 | 1.2e-6 | 2.849  | 10.59 | 1.8e-5 |
| tai250c| 2.003 | 19.00   | 3.0e-6 | 1.465  | 49.72  | 2.9e-16 | **1.349** | 10.29 | 1.3e-5 | **1.272** | 16.44 | 7.9e-5 |
| thol50 | 1.877 | 4.46    | 3.4e-6 | 2.093  | 20.67  | 3.3e-17 | 2.024  | 5.34 | 1.4e-6 | 1.933  | 6.04 | 5.6e-5 |
| wis100 | 0.717 | 2.26    | 2.4e-6 | 0.893  | 12.75  | 5.5e-14 | 0.803  | 1.30 | 1.3e-6 | 0.768  | 1.41 | 2.2e-5 |

5.4 Numerical results on graph matching

The CMU house image sequence (see [http://www.cs.cmu.edu/~cil/v-images.html](http://www.cs.cmu.edu/~cil/v-images.html)) is commonly used to test the performance of graph matching algorithms [33, 32]. This dataset consists of 111 frames of a house, each of which has been manually labeled with 30 landmarks. We connected the landmarks via Delaunay triangulation. Given an image pair \((G_1, G_2)\), the edge-affinity matrix \(K^q\) is computed by \(K^q_{ci,cj} = \exp\left(-\frac{(q^1_{ci} - q^2_{cj})^2}{2500}\right)\) and the node-affinity \(K^p\) was set to zero, where \(K^q_{ci,cj}\) measures the similarity between the \(c_i\)th edge of \(G_1\) and the \(c_j\)th edge of \(G_2\), and the edge feature \(q_{ic}\) was computed as the pairwise lowest nonnegative infeasibility, while ALMre yields the highest nonnegative infeasibility; SEPPG yields the best average median gap for 15 instances, and ALMre yields the best average median gap only for 4 instances; and the CPU time required by SEPPG is comparable with the one required by ALMre.
distance between the connected nodes. The graph matching problem is modelled as
\[
\max_{X \in \mathbb{R}_+^{n \times n} \cap \text{St}(n,n)} \text{vec}(X)^T K \text{vec}(X). \tag{28}
\]

Since the data matrix $K$ is nonnegative, we solve the problem (1), instead of its lifted reformulation as in (27), with the four solvers and compare their performance with that of FGM-D, a graph matching algorithm [32]. The FGM-D is a path-following algorithm with a heuristic strategy, designed by the convex and concave relaxations for (1). Considering that the matrix $K$ usually has a small spectral norm, for the subsequent experiments, we solve the subproblem (17) with $\rho_t$ from the starting point $\text{myQR}(|X^{l-1}|, n)$, and run ALMre with the subproblem (24a) solved from the starting point $\text{myQR}(|X^k|, n)$. In addition, we choose $\mu_0 = 1$ for ALM, which is found to be much better than $\mu_0 = 10$.

The sequence gap of a pair of images is defined as the number of images lying between them, and the graph matching problems become more difficult in the presence of a larger sequence gap. In the first experiment, we use image $\#1$ to match image $\#30$, $\#60$, $\#90$, and $\#100$, respectively, and remove 5 out of the 30 landmarks for the purpose of noise simulation. The illustration (via SEPPG) is shown in Figure 1, where the correct matches of the landmarks are drawn in yellow line, the incorrect ones are colored in blue, and the green lines label the missing landmarks. The accuracy is defined by $\frac{n_c}{30-n_m}$ where $n_c$ denotes the number of landmarks that are correctly matched, and $n_m$ denotes the number of missing landmarks. Figure 1 displays the matching effect of SEPPG in terms of the accuracy corresponding to the best objective value yielded by running 10 times from different starting points. We see that SEPPG can provide a satisfactory matching by running 10 times randomly (with accuracy attaining 92% ~ 96%).

Next we undertake the experiments in two scenarios. In the first scenario, we match a sequence of original image pairs with different sequence gap varying from 0 to 90. Figure 2 (a) displays the performance of FGM-D and four solvers in terms of the accuracy
Figure 2: Performance of solvers on (a) graph pairs with perfect 30 nodes (b) graph pairs with 5 nodes removed randomly

and the objective values of (28), where the accuracy and the objective values of the four solvers take the average for 10 times running from the starting points generated randomly. Here, the accuracy is defined by the same formula as above, and the objective values are normalized to 1 for ease of comparison. We see that SEPPG, ALM and ALMre yield the comparable accuracy and objective values, which are close to those yielded by FGM-D, while EPPGSN yields a little worse accuracy and objective value. In the second scenario, we simulate noisy images by randomly removing 5 landmarks from each image and test the performance on matching the sequence of noisy image pairs. Figure 2 (b) displays the performance of FGM-D and four solvers in terms of the matching accuracy and the objective value of (28), where the accuracy and the objective values of the four solvers take the average for 10 times running from the starting points generated randomly. We see that the performance of SEPPG, ALM and ALMre is still comparable in terms of accuracy and objective values. When the sequence gap is over 30, their performance is worse than that of EPPGSN in terms of the accuracy, but is better than the latter in terms of the objective value. When the the sequence gap is over 60, their objective values are very close to the one yielded by FGM-D, though their accuracies are a little less favorable than the one yielded by FGM-D.

To close this section, we conclude that SEPPG has a better performance than EPPGSN does, which coincides with the convergence results in Section 4; and SEPPG has a comparable performance with ALMre does. The performance of ALMre is superior to that
of ALM for the instances from QAPLIB, and is comparable with that of ALM for the graph matching problems, which means that the ALM with the subproblems solved from a random starting point will have a better performance for those difficult problems.

6 Conclusions

For the problem (1) with the nonnegative orthogonal constraint, we first removed the restrictive assumption in [19] to establish that the penalty problem (3) induced by the nonnegative constraint is a global exact penalty and so is the problem (4) when $f$ satisfies the growth condition in (16). The two exact penalty problems, different from those from [16, 17], keep the structure of $\text{St}(n,r)$ intact. Based on this, we develop two exact penalty methods by seeking the approximate critical points of a finite number of penalty problems (4) (respectively, (3)) with increasing $\rho$. Two PGMs with a line-search strategy were also provided for seeking such approximate critical points. Numerical comparisons with the ALM in [29] on the 134 instances in QAPLIB and with FGM-D on the graph pair matching problems show that the penalty method based on $\Theta_{\rho, \gamma}$ is superior to the one based on $\Phi_{\rho}$, which coincides with the convergence results for them; and the penalty method based on $\Theta_{\rho, \gamma}$ is comparable with the refined implementation of the ALM in [29] in terms of the quality of solutions and the CPU time.

Acknowledgements. The authors are deeply indebted to Professor Wen zaiwen from Peking University for sharing their codes with us to make numerical comparisons.

References

[1] P. A. Absil, R. Mathony and R. Sepulchre, Optimization Algorithms on Matrix Manifolds, Princeton University Press, Princeton, 2008.

[2] H. Attouch, J. Bolte, P. Redont and A. Soubeyran, Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka-Łojasiewicz inequality, Mathematics of Operations Research, 35(2010): 438-457.

[3] H. Attouch, J. Bolte and B. F. Svaiter, Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods, Mathematical Programming, 137(2013): 91-129.

[4] J. Barzilai and J. M. Borwein, Two-point step size gradient methods, IMA Journal of Numerical Analysis, 8(1988): 141-148.

[5] G. C. Bento, J. X. Cruz Neto and P. R. Oliveira, A new approach to the proximal point method: convergence on general Riemannian manifold, Journal of Optimization Theory and Applications, 168(2016): 743-755.
[6] N. Boumal, P. A. Absil and C. Cartis, Global rates of convergence for nonconvex optimization on manifolds, IMA Journal of Numerical Analysis, 1(2018): 1-33.

[7] R. E. Burkard, S. E. Karisch and F. Rendl, QAPLIB: a quadratic assignment problem library, Journal of Global Optimization, 10(1997): 391-403.

[8] S. X. Chen, S. Q. Ma, A. M. So and T. Zhang, Proximal gradient method for nonsmooth optimization over the Stiefel manifold, SIAM Journal on Optimization, 30(2020): 210-239.

[9] J. X. Cruz Neto, P. R. Oliveira, P. A. Soares Júnior and A. Soubeyran, Learning how to play Nash and alternating minimization method for structured non-convex problems on Riemannian manifolds, Journal of Convex Analysis, 20(2013): 395-438.

[10] Y. H. Dai and R. Fletcher, Projected Barzilai-Borwein methods for large-scale box-constrained quadratic programming, Numerische Mathematik, 100(2005): 21-47.

[11] C. Ding, D. F. Sun and K. C. Toh, An introduction to a class of matrix cone programming, Mathematical Programming, 144(2014): 141-179.

[12] C. Ding, T. Li, W. Peng and H. Park, Orthogonal nonnegative matrix t-factorizations for clustering, in Proceedings of the 12th ACM SIGKDD, ACM, 2006: 126-135.

[13] F. Facchinei and J. S. Pang, Finite-dimensional Variational Inequalities and Complementarity Problems. Springer, New York, 2003.

[14] L. Grippo, F. Lampariello and S. Lucidi, A nonmonotone line search technique for newton’s method, SIAM Journal on Numerical Analysis, 23(1986): 707-716.

[15] L. Grippo and M. Sciandrone, Nonmonotone globalization techniques for the barzilai-borwein gradient method, Computational Optimization and Applications, 23(2002): 143-169.

[16] B. Jiang, Y. F. Liu and Z. W. Wen, An exact penalty approach for optimization with nonnegative orthogonality constraints, SIAM Journal on Optimization, 26(2016): 2284-2313.

[17] B. Jiang, X. Meng, Z. W. Wen and X. J. Chen, An exact penalty approach for optimization with nonnegative orthogonality constraints, arXiv:1907.12424v2.

[18] B. Jiang, J. Tang, C. Ding and B. Luo, Nonnegative orthogonal graph matching, In Thirty-First AAAI Conference on Artificial Intelligence, 2017.

[19] S. P. Han and O. L. Mangasarian, Exact penalty functions in nonlinear programming, Mathematical Programming, 17(1979): 251-269.
[20] W. Huang, K. A. Gallivan and P. A. Absil, A Broyden class of quasi-Newton methods for Riemannian optimization, SIAM Journal on Optimization, 25(2015): 1660-1685.

[21] W. Huang and K. Wei, Riemannian proximal gradient methods, Mathematical Programming, 25(2021): 1660-1685.

[22] Y. S. Ledyaev and Q. J. Zhu, Nonsmooth analysis on smooth manifolds, Transactions of the American Mathematical Society, 359(2007): 3687-3732.

[23] A. Montanari and E. Richard, Non-negative principal component analysis: Message passing algorithms and sharp asymptotics, IEEE Transactions on Information Theory, 62(2016): 1458-1484.

[24] Y. L. Liu, S. J. Bi and S. H. Pan, Equivalent Lipschitz surrogates for zero-norm and rank optimization problems, Journal of Global Optimization, 72(2018): 679-704.

[25] D. Luo, C. Ding, H. Huang and T. Li, Non-negative Laplacian embedding, in 2009 Ninth ICDM, IEEE, 2009: 337-346.

[26] J. B. Hiriart-Urruty, J. J. Strodiot and V. H. Nguyen, Generalized Hessian matrix and second-order optimality conditions for problems with $C^{1,1}$ data, Applied Mathematics and Optimization, 11(1984): 43-56.

[27] A. D. Ioffe and J. V. Outrata, On metric and calmness qualification conditions in subdifferential calculus, Set-Valued Analysis, 16(2008): 199-227.

[28] R. T. Rockafellar and R. J-B. Wets, Variational Analysis, Springer, 1998.

[29] Z. W. Wen and W. T. Yin, A feasible method for optimization with orthogonality constraints, Mathematical Programming, 142(2013): 397-434.

[30] S. J. Wright, R. Nowak and M. Figueiredo, Sparse reconstruction by separable approximation, IEEE Transactions on Signal Processing, 57(2009): 2479-2493.

[31] Y. Yang, Y. Yang, H. T. Shen, Y. Zhang, X. Du and X. Zhou, Discriminative nonnegative spectral clustering with out-of-sample extension, IEEE Transaction on Knowledge and Data Engineering, 25(2012): 1760-1771.

[32] F. Zhou and F. De la Torre, Factorized graph matching, IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(2015): 1774-1789

[33] Anita and A. Yadav, Discrete artificial electric field algorithm for high-order graph matching, Applied Soft Computing, 92(2020): 106260.

[34] T. Yu, J. Yan, Y. Wang, W. Liu and B. Li, Generalizing graph matching beyond quadratic assignment model, In Proceedings of the 32nd International Conference on Neural Information Processing Systems, 2018: 861-871.
This part provides a dual regularized semismooth Newton method for solving (21). Fix any $k \in \mathbb{N}$. Let $Z^k = X^k - t_k \nabla f(X^k)$. Then, $V^k = Y^k - X^k$ with $Y^k$ defined by

$$
Y^k := \arg \min \frac{1}{2t_k} \|Y - Z^k\|_F^2 + \rho \vartheta(Y) - \frac{t_k}{2} \|\nabla f(X^k)\|_F^2.
$$

s.t. $A_k(Y) = A_k(X^k)$ \hspace{1cm} (29)

where $A_k$ is exactly the mapping $A_{X^k}$. The dual of problem (29) has the following form

$$
\min_{U \in \mathbb{S}^r} \Psi_k(U) := \frac{t_k}{2} \|A_k^*(U) - \nabla f(X^k)\|_F^2 - \rho_k e_{t_k \rho_k} \vartheta(Z^k + t_k A_k^*(U)),
$$

where $A_k^*$ denotes the adjoint of $A_k$. Since $A_k A_k^* = 4I$, the dual can be simplified as

$$
\min_{U \in \mathbb{S}^r} \Psi_k(U) := 2t_k \|U\|_F^2 - \rho_k e_{t_k \rho_k} \vartheta(Z^k + t_k A_k^*(U)) + \frac{t_k}{2} \|\nabla f(X^k)\|_F^2, \hspace{1cm} (30)
$$

Since the linear operator $A_k$ is surjective, the strong duality holds for problems (29) and (30). By the smoothness and convexity of $\Psi_k$, solving problem (30) is equivalent to seeking a root to the system

$$
0 = \nabla \Psi_k(U) = -A_k(Z^k) + A_k P_{t_k \rho_k} \vartheta(Z^k + t_k A_k^*(U)), \hspace{1cm} (31)
$$

which is strongly semismooth by the piecewise linearity of $\vartheta$ and [13, Proposition 7.4.7]. Since the mapping $\nabla \Psi_k : \mathbb{S}^r \to \mathbb{S}^r$ is globally Lipschitz continuous, we define the generalized Hessian of $\Psi_k$ at $U$ by $\partial^2 \Psi_k(U) \vartheta = \partial_{C, F}^2 \Psi_k(U)$. By [26, Theorem 2.2],

$$
\partial^2 \Psi_k(U)H = \hat{\partial}^2 \Psi_k(U)H \quad \text{for any} \quad H \in \mathbb{S}^r,
$$

with $\hat{\partial}^2 \Psi_k(U) = t_k A_k \partial_{C, F} \vartheta(Z^k + t_k A_k^*(U)) A_k^*$. By Lemma 2.1, it follows that for every $W \in \partial_{C, F} \vartheta(Z^k + t_k A_k^*(U))$, the linear operator $A_k W A_k^* : \mathbb{S}^r \to \mathbb{S}^r$ is self-adjoint and positive semidefinite. Since the element of $\hat{\partial}^2 \Psi_k(U)$ may be singular, a direct application of the semismooth Newton method to (31) will fail. Inspired by the work [35], we employ the following regularized semismooth Newton conjugate gradient (CG) to solve (31), and its global and local convergence analysis, the reader may refer to [35, Theorem 3.3-3.4].

[35] X. Y. Zhao, D. F. Sun and K. C. Toh, A Newton-CG augmented Lagrangian method for semidefinite programming, SIAM Journal on Optimization, 20(2010): 1737-1765.

[36] R. Zass and A. Shashua, Nonnegative sparse PCA, in NeurIPS, 2007: 1561-1568.
**Algorithm 3 (Dual regularized semismooth Newton CG)**

**Initialization:** Fix any \( k \in \mathbb{N} \cup \{0\} \). Choose a small constant \( \tau > 0 \). Choose constants \( \eta, \delta \in (0, 1), \varsigma \in (0, 1], 0 < \varphi_1 < \varphi_2 < \frac{1}{2} \), and a starting point \( U^0 \in S^r \). Set \( j = 0 \).

while the stopping conditions are not satisfied do

1. Choose \( \mathcal{W}^j \in \partial_C \mathcal{P}_{t_k \rho_k} \partial( Z^k + t_k \mathcal{A}_k^* (U^j) ) \). Solve the following linear system
   \[
   \mathcal{B}^j \mathcal{H} = - \nabla \Psi_k(U^j) \text{ with } \mathcal{B}^j := \tau_j I + \gamma \mathcal{A}_k \mathcal{W}^j \mathcal{A}_k^* \\
   \text{for } \tau_j = \min(\tau, \| \nabla \Psi_k(U^j) \|_F) \text{ with the CG method to find } \mathcal{H}^j \text{ such that}
   \]
   \[
   \| \mathcal{B}^j \mathcal{H} + \nabla \Psi_k(U^j) \|_F \leq \min(\eta, \| \nabla \Psi_k(U^j) \|_F)^{1+\varsigma}).
   \]

2. Set \( \alpha_j = \delta^m_j \), where \( m_j \) is the smallest nonnegative integer \( m \) satisfying
   \[
   \left\{ \\
   \Psi_k(U^j + \delta^m \mathcal{H}^j) \leq \Psi_k(U^j) + \varrho_1 \delta^m \langle \nabla \Psi_k(U^j), \mathcal{H}^j \rangle, \\
   |\langle \nabla \Psi_k(U^j + \delta^m \mathcal{H}^j), \mathcal{H}^j \rangle| \leq \varrho_2 |\langle \nabla \Psi_k(U^j), \mathcal{H}^j \rangle|.
   \right. \]
   (32a)
   (32b)

3. Set \( U^{j+1} = U^j + \alpha_j \mathcal{H}^j \). Let \( j \leftarrow j + 1 \), and then go to Step 1.

end while

**Lemma 1** Fix any \( X \in \mathbb{R}^{n \times r} \). There is \( \delta > 0 \) such that for all \( H \in \mathbb{R}^{n \times r} \) with \( \| H \|_F \leq \delta \),
   \[
   \partial \partial(X + H) \subseteq \partial \partial(X).
   \]

**Proof:** Let \( h(t) := \max(0, -t) \) for \( t \in \mathbb{R} \). For any given \( t \in \mathbb{R} \), it is immediate to have \( \partial h(t) = \{0\} \) if \( t > 0 \); \( \partial h(t) = [-1, 0] \) if \( t = 0 \); and \( \partial h(t) = \{-1\} \) if \( t < 0 \). Clearly, when \( X = 0 \), any \( \delta > 0 \) satisfies the requirement. So, we only need to consider the case that \( X \neq 0 \). For convenience, for any \( Z \in \mathbb{R}^{n \times r} \), let \( \Omega_+(Z) = \{(i, j) \mid Z_{ij} > 0\} \), \( \Omega_-(Z) = \{(i, j) \mid Z_{ij} < 0\} \) and \( \Omega_0(Z) = \{(i, j) \mid Z_{ij} = 0\} \). Take \( \delta = |X_{i^*, j^*}| \) where \( |X_{i^*, j^*}| \) is the smallest nonzero entry of the matrix \( |X| \). Fix any \( H \in \mathbb{R}^{n \times r} \) with \( \| H \|_F \leq \delta \). Clearly, \( \Omega_+(X + H) = \Omega_+(X) \) and \( \Omega_-(X + H) = \Omega_-(X) \), so \( \partial h(X_{ij} + H_{ij}) = \partial h(X_{ij}) \) whenever \( (i, j) \in \Omega_+(X) \) or \( \Omega_-(X) \). In addition, it is clear that \( \partial h(X_{ij} + H_{ij}) \subseteq \partial h(X_{ij}) \) whenever \( (i, j) \in \Omega_0(X) \). Together with \( \partial \partial(Z) = \bigotimes_{i=1}^n \bigotimes_{j=1}^r \partial h(Z_{ij}) \) for any \( Z \in \mathbb{R}^{n \times r} \), it follows that \( \partial \partial(X + H) \subseteq \partial \partial(X) \). By the arbitrariness of \( H \), the conclusion holds. \( \square \)