

# A semi-smooth Newton method for general projection equations applied to the nearest correlation matrix problem

Nicolas F. Armijo <sup>\*</sup>      Yunier Bello-Cruz<sup>†</sup>      Gabriel Haeser <sup>‡</sup>

January 7, 2024

## Abstract

In this paper, we extend and investigate the properties of the semi-smooth Newton method when applied to a general projection equation in finite dimensional spaces. We first present results concerning Clarke’s generalized Jacobian of the projection onto a closed and convex cone. We then describe the iterative process for the general cone case and establish two convergence theorems. We apply these results to the constrained quadratic conic programming problem, emphasizing its connection to the projection equation. To illustrate the performance of our method, we conduct numerical experiments focusing on semidefinite least squares, in particular the nearest correlation matrix problem. In the latter scenario, we benchmark our outcomes against previous literature, presenting performance profiles and tabulated results for clarity and comparison.

**Keywords:** Conic programming, nearest correlation matrix, quadratic programming, semi-smooth Newton method.

**2010 AMS Subject Classification:** 90C33, 15A48.

## 1 Introduction

We begin by considering the following special nonlinear system:

$$P_{\mathcal{K}}(x) + Tx = b, \tag{1}$$

where  $\mathcal{K} \subseteq \mathbb{X}$  is a non-empty, closed and convex cone of a finite dimensional vector space  $\mathbb{X}$  with an inner product  $\langle \cdot, \cdot \rangle$ ,  $P_{\mathcal{K}}(x)$  is the projection of  $x \in \mathbb{X}$  onto  $\mathcal{K}$ ,  $b \in \mathbb{X}$ , and  $T: \mathbb{X} \rightarrow \mathbb{X}$  is a linear operator. Some particular cases of equation (1) have been studied, for instance, in [1–5, 8, 10, 14, 16, 19]. Among those problems, particular attention has been given to the cases where  $\mathcal{K}$  is the  $n$ -dimensional non-negative orthant or Lorentz’s cone. For these cases, novel iterative methods have been proposed; see, for instance, [1, 2, 4].

---

<sup>\*</sup>Department of Applied Mathematics, University of São Paulo, Brazil (e-mail: [nfarmijo@ime.usp.br](mailto:nfarmijo@ime.usp.br)). The author was supported by Fapesp grant 2019/13096-2.

<sup>†</sup>Northern Illinois University, USA (e-mail: [yunierbello@niu.edu](mailto:yunierbello@niu.edu)). The author was partially supported by the NSF Grant DMS-2307328 and by an internal grant from NIU.

<sup>‡</sup>Department of Applied Mathematics, University of São Paulo, Brazil (e-mail: [ghaeser@ime.usp.br](mailto:ghaeser@ime.usp.br)). The author was supported by CNPq and Fapesp grant 2018/24293-0.

Equation (1) is closely related to the quadratic cone-constrained programming:

$$\begin{aligned} \min \quad & \frac{1}{2}\langle x, Qx \rangle + \langle q, x \rangle, \\ \text{s.t.} \quad & x \in \mathcal{K}, \end{aligned} \tag{2}$$

for a linear operator  $Q: \mathbb{X} \rightarrow \mathbb{X}$  and a vector  $q \in \mathbb{X}$ . The particularly relevant case occurs when  $\mathbb{X} = \mathbb{R}^n$  and  $\mathcal{K}$  is either the non-negative orthant or Lorentz's cone. The connection of (1) with (2) is established by setting  $T = (Q - \text{Id})^{-1}$  (where  $\text{Id}$  is the  $n \times n$  identity matrix) and  $b = -Tq$ . Moreover, the projection onto  $\mathcal{K}$  of a solution of equation (1) satisfies the first order necessary optimality conditions for problem (2). Here, we also prove that this property holds in the general case, that is, for any closed and convex cone  $\mathcal{K}$  in a finite dimensional vector space  $\mathbb{X}$  and considering any linear operator  $Q: \mathbb{X} \rightarrow \mathbb{X}$ . Additional linear equality constraints are also considered in (2).

We here focus our attention on the *semi-smooth Newton method* for solving equation (1). This method finds a zero of the mapping  $F: \mathbb{X} \rightarrow \mathbb{X}$ ,

$$F(x) = P_{\mathcal{K}}(x) + Tx - b, \quad x \in \mathbb{X}. \tag{3}$$

By starting at a point  $x^0 \in \mathbb{X}$ , the classical semi-smooth Newton method iterates as follows:

$$x^{k+1} = x^k - \left[ F'(x^k) \right]^{-1} F(x^k), \tag{4}$$

where  $F'(x^k)$  is a generalized Jacobian of  $F$  at  $x^k$ . This iteration applying to (3) will take the following simple form:

$$\left( V(x^k) + T \right) x^{k+1} = b, \quad k \in \mathbb{N}, \tag{5}$$

where  $V(x^k) \in \partial_C P_{\mathcal{K}}(x^k)$  is a Clarke's generalized Jacobian of the projection  $P_{\mathcal{K}}(\cdot)$  at  $x^k$ . This is a consequence of the relation  $V(x^k)x^k = P_{\mathcal{K}}(x^k)$  which we will prove later.

Our approach allows us to consider the relevant case when  $\mathcal{K} = \mathbb{S}_+^n$ , the cone of positive semidefinite  $n \times n$  matrices, where a subdifferential  $V(x^k)$  can be computed by the spectral decomposition of  $x^k$ . For this case, we will tackle the quadratic cone-constrained problem (2), where we also include additional linear constraints; see problem (13) below. This is a well-known optimization problem with several applications. The goal is to study further the properties of the semi-smooth Newton method and prove the Q-linear global convergence of the iteration under some standard conditions on the linear operator  $T$  for a general cone  $\mathcal{K}$ , allowing a more detailed study for the cone of positive semidefinite matrices.

The paper is organized as follows: we start by introducing our notation and presenting some preliminary results needed in our analysis. In Section 2, we present general results about the projection operator onto cones and its differentiability. In Section 3, we present the semi-smooth Newton method and prove the previously mentioned global convergence theorems. Section 4 is devoted to exploring the relationship between equation (1) and the quadratic conic programming problem (2), with special emphasis on the case when  $\mathcal{K} = \mathbb{S}_+^n$ , including additional linear constraints. Finally, we present some numerical experiments for the positive semidefinite cone  $\mathbb{S}_+^n$ , focusing particularly on semidefinite least squares problems. Specifically, we include a comparison with the algorithm from [18] in the context of the nearest correlation matrix problem, which is a well studied topic in economics [11].

## 1.1 Notations and preliminaries

In this section, we present some relevant results and definitions that are used in this paper. We denote the nonnegative integers by  $\mathbb{N}$  and by  $\text{Id}$  the identity operator. The bracket notation  $\langle \cdot, \cdot \rangle$  is

referred to the inner product in any finite dimensional space  $\mathbb{X}$ . Given a linear operator  $T: \mathbb{X} \rightarrow \mathbb{X}$ , we use the notation  $\|T\|$  for the operator norm of  $T$ , that is,  $\|T\| := \max\{\|Tx\| \mid x \in \mathbb{X}, \|x\| = 1\}$ , where  $\|x\| := \sqrt{\langle x, x \rangle}$  is the norm associated with the inner product. We also denote by  $T^*$  its adjoint linear operator  $T^*: \mathbb{X} \rightarrow \mathbb{X}$ , that is,  $\langle Tx, y \rangle = \langle x, T^*y \rangle$ , for all  $x, y \in \mathbb{X}$ . When  $T$  is self-adjoint, that is,  $T = T^*$ , we say that  $T$  is positive semidefinite (definite) if  $\langle Tx, x \rangle \geq 0$  ( $> 0$ , respectively),  $\forall x \in \mathbb{X}, x \neq 0$  and we denote by  $\lambda_{\min}(T)$  and  $\lambda_{\max}(T)$  the smallest and largest eigenvalues of  $T$ , respectively. For a cone  $\mathcal{K} \subseteq \mathbb{X}$ , the dual of  $\mathcal{K}$  is denoted by  $\mathcal{K}^* = \{y \in \mathbb{X} \mid \langle y, x \rangle \geq 0, \forall x \in \mathcal{K}\}$ . When  $\mathbb{X} = \mathbb{S}^n$  is the space of symmetric  $n \times n$  matrices with the inner product  $\langle A, B \rangle = \text{trace}(AB)$ ,  $A, B \in \mathbb{S}^n$ , we consider the self-dual cone  $\mathcal{K} = \mathbb{S}_+^n$  of positive semidefinite matrices.

The projection of a point  $x$  onto a closed and convex cone  $\mathcal{K} \neq \emptyset$  is denoted by  $P_{\mathcal{K}}(x)$  and is defined by  $P_{\mathcal{K}}(x) = \text{argmin}\{\|y - x\| \mid y \in \mathcal{K}\}$ . For a given mapping  $F: \mathbb{X} \rightarrow \mathbb{X}$ , we denote the set where it is differentiable by  $D_F$  and the Jacobian at a point  $x \in D_F$  by  $F'(x)$ . The set of Clarke's generalized Jacobians at a point  $x \in \mathbb{X}$  is denoted by  $\partial_C F(x)$  and it is defined by

$$\partial_C F(x) = \text{conv} \left\{ \lim_{k \rightarrow \infty} F'(x_k) \mid x_k \rightarrow x, x_k \in D_F \right\},$$

that is, the convex hull of all limits of Jacobians of  $F$  nearby  $x$ . Throughout this paper, a Clarke's generalized Jacobian of the projection  $P_{\mathcal{K}}(x)$  will be denoted by  $V(x)$ , so there is no confusion in not referring to a particular mapping  $F$ .

We will make use of the well-known results below, which we state in the context of a general finite dimensional inner product space  $\mathbb{X}$  as follows.

**Theorem 1.1** (Mean Value Theorem [6], Proposition 2.6.5, Page 79). *Let  $F: \mathbb{X} \rightarrow \mathbb{X}$  be a Lipschitz mapping. Then, we have*

$$F(y) - F(x) \in \text{conv}(\partial_C F([x, y]))(y - x),$$

that is,  $F(y) - F(x) = U(z)(y - x)$  where  $U(z) \in \partial_C F(z)$  and  $z$  is a convex combination of  $x$  and  $y$ .

**Lemma 1.1** (Banach's Lemma [13], Page 351). *Let  $E: \mathbb{X} \rightarrow \mathbb{X}$  be a mapping onto  $\mathbb{X}$ . If  $\|E\| < 1$ , then  $E - \text{Id}$  is invertible and*

$$\|(E - \text{Id})^{-1}\| \leq \frac{1}{1 - \|E\|}.$$

**Lemma 1.2** (Weyl's inequality [13], Theorem 4.3.1, Page 239). *Let  $A, B: \mathbb{X} \rightarrow \mathbb{X}$  be self-adjoint linear operators. Then, it holds*

$$\lambda_{\min}(A) + \lambda_{\min}(B) \leq \lambda_{\min}(A + B) \leq \lambda_{\max}(A + B) \leq \lambda_{\max}(A) + \lambda_{\max}(B).$$

Finally, an important result to ensure the existence and uniqueness of solutions of equation (1) is the contraction mapping principle.

**Theorem 1.2** (Contraction mapping principle [17], Thm. 8.2.2, page 153). *Let  $\Phi: \mathbb{X} \rightarrow \mathbb{X}$  and suppose that there exists  $\lambda \in [0, 1)$  such that  $\|\Phi(y) - \Phi(x)\| \leq \lambda\|y - x\|$ , for all  $x, y \in \mathbb{X}$ . Then, there exists a unique  $\bar{x} \in \mathbb{X}$  such that  $\Phi(\bar{x}) = \bar{x}$ .*

## 2 On the projection mapping onto a closed and convex cone

In this section, we study some useful results that will be important in the well-definiteness and global convergence of the semi-smooth Newton method for equation (1). We begin by presenting the following result on the properties of generalized Jacobians.

**Theorem 2.1.** *The projection operator  $P_{\mathcal{K}}(\cdot)$  is differentiable almost everywhere. The Jacobian  $P'_{\mathcal{K}}(x)$  (when it exists) and any generalized Jacobian  $V(x) \in \partial_C P_{\mathcal{K}}(x)$  for all  $x \in \mathbb{X}$ , are self-adjoint and positive semidefinite operators. Moreover, the following properties hold:*

- (i)  $\|V(x)\| \leq 1, \forall V(x) \in \partial_C P_{\mathcal{K}}(x)$  with  $x \in \mathbb{X}$ .
- (ii)  $P'_{\mathcal{K}}(x)x = P_{\mathcal{K}}(x), \forall x \in D_{P_{\mathcal{K}}}$ .
- (iii)  $V(x)x = P_{\mathcal{K}}(x), \forall V(x) \in \partial_C P_{\mathcal{K}}(x)$  with  $x \in \mathbb{X}$ .
- (iv) For all  $x \in \mathbb{X}$ ,

$$0 \leq \lambda_{\min}(V(x)) \leq \lambda_{\max}(V(x)) \leq 1, \forall V(x) \in \partial_C P_{\mathcal{K}}(x).$$

*Proof.* The fact that the projection is differentiable almost everywhere is well-known due to its non-expansiveness (that is, the projection is 1-Lipschitz). When  $P'_{\mathcal{K}}(x)$  exists, it is self-adjoint and positive semidefinite due to Proposition 2.2 of [9]. Now, let  $x \in \mathbb{X}$  and  $V(x) \in \partial_C P_{\mathcal{K}}(x)$ . By definition we have that there exist  $V_1(x), \dots, V_m(x)$  and  $\{x_k^j\} \subset D_{P_{\mathcal{K}}}$  such that  $\lim_{k \rightarrow \infty} x_k^j = x$ ,  $P'_{\mathcal{K}}(x_k^j) \rightarrow V_j(x), \forall j = 1, \dots, m$ , and  $V(x) = \sum_{j=1}^m \alpha_j V_j(x)$ , with  $\sum_{j=1}^m \alpha_j = 1$  and  $\alpha_j \in [0, 1]$  for all  $j$ . By the continuity of the inner product, we can deduce that each  $V_j(x)$  is also self-adjoint and positive semidefinite, and therefore, by linearity, the same holds true for  $V(x)$ .

To prove item (i), note that  $\|P'_{\mathcal{K}}(x_k^j)\| \leq 1$  for all  $j$  and  $k$  due to non-expansiveness. Hence

$$\begin{aligned} \|V(x)\| &= \left\| \sum_{j=1}^m \alpha_j V_j(x) \right\| \leq \sum_{j=1}^m \alpha_j \|V_j(x)\| \\ &= \sum_{j=1}^m \alpha_j \lim_{k \rightarrow \infty} \|P'_{\mathcal{K}}(x_k^j)\| \leq 1. \end{aligned}$$

For item (ii), note that since  $\mathcal{K}$  is a cone,  $P_{\mathcal{K}}(\cdot)$  is positive homogeneous, that is,  $P_{\mathcal{K}}(tx) = tP_{\mathcal{K}}(x), \forall t \geq 0$  for any  $x \in \mathbb{X}$ . Let  $x \in D_{P_{\mathcal{K}}}$ . If  $x = 0$ , the equality is evident. Assume that  $x \neq 0$ . By the definition of  $P'_{\mathcal{K}}(x)$ , we have that

$$0 = \lim_{t \rightarrow 0} \frac{\|P_{\mathcal{K}}(x + tx) - P_{\mathcal{K}}(x) - tP'_{\mathcal{K}}(x)x\|}{\|tx\|} = \frac{\|P_{\mathcal{K}}(x) - P'_{\mathcal{K}}(x)x\|}{\|x\|}.$$

Hence  $P'_{\mathcal{K}}(x)x = P_{\mathcal{K}}(x)$ , which proves item (ii).

In order to prove item (iii), by noting that  $V(x)x - P_{\mathcal{K}}(x) = \sum_{j=1}^m \alpha_j (V_j(x)x - P_{\mathcal{K}}(x))$ , it is enough to show that for all  $j = 1, \dots, m$ ,  $V_j(x)x = P_{\mathcal{K}}(x)$ . Recalling that  $P'_{\mathcal{K}}(x_k^j) \rightarrow V_j(x)$ , we have that

$$\begin{aligned} \|V_j(x)x - P'_{\mathcal{K}}(x_k^j)x_k^j\| &\leq \|V_j(x)x - V_j(x)x_k^j\| + \|V_j(x)x_k^j - P'_{\mathcal{K}}(x_k^j)x_k^j\| \\ &\leq \|V_j(x)\|\|x - x_k^j\| + \|V_j(x) - P'_{\mathcal{K}}(x_k^j)\|\|x_k^j\| \\ &\rightarrow 0. \end{aligned}$$

Using item (ii) and the continuity of  $P_{\mathcal{K}}(\cdot)$  we conclude that

$$\begin{aligned} \|V_j(x)x - P_{\mathcal{K}}(x)\| &\leq \|V_j(x)x - P_{\mathcal{K}}(x_k^j)\| + \|P_{\mathcal{K}}(x_k^j) - P_{\mathcal{K}}(x)\| \\ &= \|V_j(x)x - P'_{\mathcal{K}}(x_k^j)x_k^j\| + \|P_{\mathcal{K}}(x_k^j) - P_{\mathcal{K}}(x)\| \rightarrow 0. \end{aligned}$$

Finally, for item (iv), it is enough to note that  $0 \leq \lambda_{\min}(V_j(x)) \leq \lambda_{\max}(V_j(x)) \leq 1$  for all  $j = 1, \dots, m$  due to the fact that  $V_j(x)$  is self-adjoint and positive semidefinite with  $\|V_j(x)\| \leq 1$ . The result now follows easily from Lemma 1.2.  $\square$

We conclude the section with the following useful result.

**Lemma 2.1.** *Let  $x, y \in \mathbb{X}$  and  $V(x) \in \partial_C P_{\mathcal{K}}(x)$ . Then  $\|P_{\mathcal{K}}(y) - P_{\mathcal{K}}(x) - V(x)(y - x)\| \leq \|y - x\|$ .*

*Proof.* By Theorem 1.1 we have that

$$P_{\mathcal{K}}(y) - P_{\mathcal{K}}(x) - V(x)(y - x) = (V(z) - V(x))(y - x),$$

with  $V(z) \in \partial_C P_{\mathcal{K}}(z)$ , where  $z$  is a convex combination of  $x$  and  $y$ . The result follows from the fact that  $\|V(z) - V(x)\| \leq 1$  due to Lemma 1.2 and Theorem 2.1 item (iv).  $\square$

Item (ii) from Theorem 2.1 provides a foundation for introducing the semi-smooth Newton method for solving equation (1). Specifically, since the projection can be expressed as

$$P_{\mathcal{K}}(x) = V(x)x,$$

where  $V(x) \in \partial_C P_{\mathcal{K}}(x)$ , we have that  $F(x)$  as defined in (3) can be written as

$$F(x) = (V(x) + T)x - b,$$

with  $V(x) + T \in \partial_C F(x)$ . Then, iteration (4) can be expressed as

$$x^{k+1} = x^k - (V(x^k) + T)^{-1}[(V(x^k) + T)x^k - b] = (V(x^k) + T)^{-1}b.$$

In the next section, we explore the convergence properties of this iteration.

### 3 A semi-smooth Newton method for general projection equations

In this section, we define a semi-smooth Newton method for solving equation (1) and study the convergence along with the sufficient conditions required to achieve it. Our goal is to extend the application of the semi-smooth Newton method, previously studied in [2, 4], for the cases where  $\mathcal{K} \subseteq \mathbb{R}^n$  is either the non-negative orthant or Lorentz's cone. This extension considers any closed and convex cone  $\mathcal{K} \subseteq \mathbb{X}$ . First, we establish a sufficient condition to the existence and uniqueness of the solution to the equation (1).

**Theorem 3.1** (Sufficient condition for existence and uniqueness of a solution). *If  $T$  is invertible and  $\|T^{-1}\| < 1$ , then equation (1) has a unique solution for any  $b \in \mathbb{X}$ .*

*Proof.* Equation (1) has a unique solution if and only if the mapping  $\Phi(x) = -T^{-1}P_{\mathcal{K}}(x) + T^{-1}b$  has a unique fixed point. Hence, it is sufficient to prove that  $\Phi$  is a contraction and use Theorem 1.2 to guarantee the existence and uniqueness of a fixed point. From the definition of  $\Phi$ , we have

$$\Phi(x) - \Phi(y) = -T^{-1}(P_{\mathcal{K}}(x) - P_{\mathcal{K}}(y)).$$

Since  $\|P_{\mathcal{K}}(x) - P_{\mathcal{K}}(y)\| \leq \|x - y\|$  we deduce that  $\|\Phi(x) - \Phi(y)\| \leq \|T^{-1}\|\|x - y\|$  concluding that  $\Phi$  is a contraction since  $\|T^{-1}\| < 1$ .  $\square$

We define the semi-smooth Newton method for the mapping  $F(x) = P_{\mathcal{K}}(x) + Tx - b$  starting on an initial point  $x^0 \in \mathbb{X}$  as the iteration

$$(V(x^k) + T)x^{k+1} = b, \quad k \in \mathbb{N} \tag{6}$$

with  $V(x) \in \partial_C P_{\mathcal{K}}(x)$ .

Notice that if  $x^k \rightarrow \bar{x}$ , then  $\bar{x}$  is a solution of equation (1). To see this, we subtract  $(V(x^k) + T)x^k$  from both sides of (6). Using that  $P_{\mathcal{K}}(x^k) = V(x^k)x^k$ , we arrive at  $(V(x^k) + T)(x^{k+1} - x^k) = b - P_{\mathcal{K}}(x^k) - Tx^k$ . Since  $V(x^k) + T$  is bounded, the left-hand side converges to zero, while from the continuity of the projection the right-hand side converges to  $b - P_{\mathcal{K}}(\bar{x}) - T\bar{x}$ . Therefore,  $\bar{x}$  is a solution of (1).

We start by showing a sufficient condition for stopping the method (6) at a solution.

**Proposition 3.1** (Stopping criterion). *If  $V(x^{k+1}) = V(x^k)$ , then  $x^{k+1}$  is a solution of equation (1).*

*Proof.* From Theorem 2.1 item (iii) and (6), we have that

$$P_{\mathcal{K}}(x^{k+1}) + Tx^{k+1} = (V(x^{k+1}) + T)x^{k+1} = (V(x^k) + T)x^{k+1} = b.$$

$\square$

Now, we show a sufficient condition for the global convergence of iteration (6). Provided certain conditions regarding the norm of the inverse of  $T$  are met, we can guarantee the existence and uniqueness of the solution of equation (1). In addition to that, by imposing an additional norm condition, we obtain linear global convergence.

**Theorem 3.2** (Sufficient condition for global Q-linear convergence). *Let  $b \in \mathbb{X}$  and  $T: \mathbb{X} \rightarrow \mathbb{X}$  be an invertible linear operator. Assume that  $\|T^{-1}\| < 1$ . Then, equation (1) has a unique solution  $\bar{x}$  and for any initial point  $x^0$ , the semi-smooth Newton sequence generated by equation (6) is well-defined. Additionally, if  $\|T^{-1}\| < \frac{1}{2}$  then the sequence  $\{x^k\}$  converges Q-linearly to  $\bar{x}$  and satisfies*

$$\|x^{k+1} - \bar{x}\| \leq \frac{\|T^{-1}\|}{1 - \|T^{-1}\|} \|x^k - \bar{x}\|, \quad k \in \mathbb{N}.$$

*Proof.* First we know from Theorem 2.1 that  $\|V(x)\| \leq 1$  for any  $x \in \mathbb{X}$ . Since  $\|T^{-1}\| < 1$  we deduce that  $\|T^{-1}V(x)\| < 1$  for every  $x \in \mathbb{X}$ . Lemma 1.1 implies that  $-T^{-1}V(x) - \text{Id}$  is invertible and therefore  $V(x) + T$  is also invertible. In particular the semi-smooth Newton method (6) is well defined. Let  $\bar{x}$  be the only solution of problem (1) (which exists and is unique due to Theorem 3.1). So, this point satisfies the relation  $(V(\bar{x}) + T)\bar{x} - b = 0$ . Combining with (6) we deduce that

$$(V(x^k) + T)(x^{k+1} - \bar{x}) = (V(\bar{x}) - V(x^k))\bar{x} = V(\bar{x})\bar{x} - V(x^k)x^k - V(x^k)(\bar{x} - x^k).$$

Since  $V(\bar{x})\bar{x} = P_{\mathcal{K}}(\bar{x})$  and  $V(x^k)x^k = P_{\mathcal{K}}(x^k)$ , using Lemma 2.1, we obtain

$$\|x^{k+1} - \bar{x}\| \leq \|(V(x^k) + T)^{-1}\| \|(P_{\mathcal{K}}(\bar{x}) - P_{\mathcal{K}}(x^k) - V(x^k)(\bar{x} - x^k))\| \leq \|(V(x^k) + T)^{-1}\| \|\bar{x} - x^k\|.$$

But  $\|(V(x^k) + T)^{-1}\| = \|(T(T^{-1}V(x^k) + \text{Id}))^{-1}\| \leq \|(T^{-1}V(x^k) - \text{Id})^{-1}\| \|T^{-1}\|$ . Lemma 1.1 and  $\|T^{-1}V(x^k)\| < 1$  implies that

$$\|(T^{-1}V(x^k) - \text{Id})^{-1}\| \leq \frac{1}{1 - \|T^{-1}V(x^k)\|} \leq \frac{1}{1 - \|T^{-1}\|}.$$

Thus, we have that  $\|x^{k+1} - \bar{x}\| \leq \frac{\|T^{-1}\|}{1 - \|T^{-1}\|} \|x^k - \bar{x}\|$  with  $\frac{\|T^{-1}\|}{1 - \|T^{-1}\|} < 1$  due to the assumption that  $\|T\| < \frac{1}{2}$ . Hence,  $x^k$  converges Q-linearly to the unique solution  $\bar{x}$ .  $\square$

The previous result states that with only a norm condition on the operator  $T^{-1}$ , namely  $\|T^{-1}\| < \frac{1}{2}$ , we can achieve Q-linear convergence of the method. We prove next that for the case where  $T$  being a positive definite linear mapping, the weaker norm condition  $\|T^{-1}\| < 1$  is sufficient to ensure Q-linear convergence of the iteration (6) to the unique solution of the problem (1).

**Theorem 3.3.** *Let  $b \in \mathbb{X}$  and  $T: \mathbb{X} \rightarrow \mathbb{X}$  be a positive definite linear operator. Then, equation (1) has a unique solution  $\bar{x} \in \mathbb{X}$  and for any initial point  $x^0 \in \mathbb{X}$ , the semi-smooth Newton sequence generated by (6) is well-defined. Additionally, if  $\|T^{-1}\| < 1$  then the sequence  $\{x^k\}$  converges Q-linearly to  $\bar{x}$  and satisfies*

$$\|x^{k+1} - \bar{x}\| \leq \|T^{-1}\| \|x^k - \bar{x}\|, \quad k \in \mathbb{N}.$$

*Proof.* First notice that from Lemma 1.2 and the positive definiteness of  $T$ , it follows that  $\text{Id} + T$  is invertible with  $\|(\text{Id} + T)^{-1}\| < 1$ . Using Moreau's decomposition [12, Theorem 3.2.5], we can write any  $x \in \mathbb{X}$  as  $x = P_{\mathcal{K}}(x) - P_{\mathcal{K}^*}(-x)$ . Now, it follows directly that  $x$  is a solution of equation (1) if and only if  $x$  is a fixed point of  $\Phi(x) = (\text{Id} + T)^{-1}(b - P_{\mathcal{K}^*}(-x))$ . Since

$$\Phi(x) - \Phi(y) = (\text{Id} + T)^{-1}(-P_{\mathcal{K}^*}(-x) + P_{\mathcal{K}^*}(-y)), \quad x, y \in \mathbb{X},$$

we deduce that  $\Phi$  is a contraction due to the non-expansiveness of the projection. This gives existence and uniqueness of a solution of problem (1).

By Theorem 2.1 item (iv) and Lemma 1.2, it follows that  $V(x) + T$  is positive definite with  $\|(V(x) + T)^{-1}\| \leq \|T^{-1}\|$  for all  $x \in \mathbb{X}$ . Hence, iteration (6) is well-defined for every starting point  $x^0 \in \mathbb{X}$ . The Q-linear convergence when  $\|T^{-1}\| < 1$  now follows from the relation  $\|x^{k+1} - \bar{x}\| \leq \|(V(x^k) + T)^{-1}\| \|x^k - \bar{x}\|$  deduced in the proof of Theorem 3.2.  $\square$

Note that although the assumption of positive definiteness is sufficient for existence and uniqueness of the solution without imposing a condition on the norm of  $T^{-1}$ , it does not guarantee the convergence of Newton's method; see [1, Example 1].

In the following section, we show relevant applications of our results to quadratic conic programming.

## 4 Application to quadratic conic programming

In this section, we connect equation (1) with the important quadratic conic programming problem

$$\begin{aligned} \min \quad & \frac{1}{2}\langle x, Qx \rangle + \langle q, x \rangle, \\ \text{s.t.} \quad & x \in \mathcal{K}. \end{aligned} \tag{7}$$

This problem has been widely studied and has multiple applications such as semidefinite least squares and, in particular, the nearest correlation matrix problem which we will present next.

The Lagrangian of the problem is given by

$$L(x, \mu) := \frac{1}{2}\langle x, Qx \rangle + \langle q, x \rangle - \langle \mu, x \rangle,$$

where  $\mu \in \mathcal{K}^*$ . Then, the well-known complementary KKT conditions are given by

$$\begin{aligned} Q\bar{x} + q - \bar{\mu} &= 0, \\ \langle \bar{\mu}, \bar{x} \rangle &= 0. \end{aligned}$$

Or equivalently,

$$\langle Q\bar{x} + q, \bar{x} \rangle = 0, \quad Q\bar{x} + q \in \mathcal{K}^*, \quad \bar{x} \in \mathcal{K}. \tag{8}$$

In order to find a solution to the KKT system (8), we use the following modified projection equation:

$$(Q - \text{Id})P_{\mathcal{K}}(x) + x = -q. \tag{9}$$

With this in mind, we have the following connection between the solutions of (9) and the ones of the KKT conditions (8) above. The following theorem is a generalization of Proposition 4 in [4].

**Theorem 4.1** (KKT points and solutions of a generalized projection equation). *If  $x$  is solution of equation (9), then  $\bar{x} = P_{\mathcal{K}}(x)$  is a solution of the KKT system (8). On the other hand, if  $\bar{x}$  is a solution of system (8), then  $x = \bar{x} - (Q\bar{x} + q)$  is a solution of (9).*

*Proof.* For the first part, let  $x$  be a solution of (9). Using Moreau's decomposition [12, Theorem 3.2.5] for  $x$ , we have

$$x = P_{\mathcal{K}}(x) - P_{\mathcal{K}^*}(-x), \tag{10}$$

and

$$\langle P_{\mathcal{K}}(x), P_{\mathcal{K}^*}(-x) \rangle = 0. \tag{11}$$

By hypothesis we get that

$$QP_{\mathcal{K}}(x) + q = P_{\mathcal{K}}(x) - x = P_{\mathcal{K}^*}(-x) \in \mathcal{K}^*,$$

where we used (10) in the last equality. Now using the previous equation and (11), we have

$$\langle QP_{\mathcal{K}}(x) + q, P_{\mathcal{K}}(x) \rangle = \langle P_{\mathcal{K}^*}(-x), P_{\mathcal{K}}(x) \rangle = 0,$$

implying that  $P_{\mathcal{K}}(x) \in \mathcal{K}$  solves (8).

For the second part, let  $x$  be a solution of (8) and  $\bar{x} := x - (Qx + q)$ . Since  $x \in \mathcal{K}$ ,  $Qx + q \in \mathcal{K}^*$  with  $\langle Qx + q, x \rangle = 0$ , it follows by Moreau's decomposition that  $x = P_{\mathcal{K}}(\bar{x})$ . Thus, replacing  $\bar{x}$  in (9) we obtain

$$(Q - \text{Id})P_{\mathcal{K}}(\bar{x}) + \bar{x} = (Q - \text{Id})x + x - (Qx + q) = -q.$$

Therefore,  $\bar{x}$  is a solution of (9), which concludes the proof.  $\square$



Theorem 4.1 establishes a significant connection between system (8) and equation (9). In particular, it asserts that if (9) does not have a solution, then the quadratic conic programming problem (7) lacks points satisfying the complementary optimality conditions.

We now extend our previous results to the case of equation (9). The proofs of the next three results use Theorem 3.1 and closely follow the ideas presented in [4], and therefore, we omit some details for brevity. We begin by presenting two propositions regarding the existence and uniqueness of solutions as follows:

**Proposition 4.1.** *If  $\|Q - \text{Id}\| < 1$ , then equation (9) has a unique solution for any  $q \in \mathbb{X}$ .*

*Proof.* Similar to the proof of Theorem 3.1 but replacing  $T$  by  $(Q - \text{Id})^{-1}$ . □

**Proposition 4.2.** *If  $Q$  is invertible and  $\|Q^{-1} - \text{Id}\| < 1$ , then equation (9) has a unique solution for any  $q \in \mathbb{X}$ .*

*Proof.* Similar to the proof of Theorem 3.1 but defining  $\Phi(x) = (Q^{-1} - \text{Id})P_{\mathcal{K}^*}(-x) - Q^{-1}q$  and using that  $x = P_{\mathcal{K}}(x) - P_{\mathcal{K}^*}(-x)$  and the fact that the projection  $P_{\mathcal{K}^*}(\cdot)$  is non-expansive. □

Next, we specialize our  $Q$ -linear convergence results to the case of equation (9).

**Theorem 4.2.** *Let  $q \in \mathbb{X}$  and  $Q: \mathbb{X} \rightarrow \mathbb{X}$  a linear operator. Assume that  $Q - \text{Id}$  is invertible and  $\|Q - \text{Id}\| < 1$ . Then, (9) has a unique solution  $\bar{x}$ , and for any initial point  $x^0$  the semi-smooth Newton sequence generated by (6) is well-defined. Additionally, if  $\|Q - \text{Id}\| < \frac{1}{2}$  then the sequence  $\{x^k\}$  converges  $Q$ -linearly to  $\bar{x}$  and satisfies*

$$\|x^{k+1} - \bar{x}\| \leq \frac{\|Q - \text{Id}\|}{1 - \|Q - \text{Id}\|} \|x^k - \bar{x}\|, \quad k \in \mathbb{N}.$$

*Proof.* Similar to the proof of Theorem 3.2. □

Before stating a result analogous to Theorem 3.3, we need the following lemma.

**Lemma 4.1.** *Let  $x \in \mathbb{X}$  and  $V(x) \in \partial_C P_{\mathcal{K}}(x)$ . If  $Q$  is a positive definite linear mapping, then  $(Q - \text{Id})V(x) + \text{Id}$  is invertible.*

*Proof.* By contradiction let us suppose that there exists  $u \neq 0$  with  $((Q - \text{Id})V(x) + \text{Id})u = 0$ , or equivalently that

$$QV(x)u = (V(x) - \text{Id})u. \tag{12}$$

Since  $Q$  is positive definite and  $V(x)$  is self-adjoint, we have that

$$0 \leq \langle QV(x)u, V(x)u \rangle = \langle V(x)^*QV(x)u, u \rangle = \langle V(x)QV(x)u, u \rangle = \langle (V(x)^2 - V(x))u, u \rangle \leq 0,$$

where the last inequality is due to Theorem 2.1, item (iv). Then,  $\langle QV(x)u, V(x)u \rangle = 0$ , which implies that  $V(x)u = 0$ . But by (12), we conclude that  $u = 0$ , which is a contradiction. Hence, the operator  $(Q - \text{Id})V(x) + \text{Id}$  is invertible. □

Finally, we present a sufficient condition for the convergence of the semi-smooth Newton method which is analogous to Theorem 3.3 applied to the quadratic conic programming problem when  $Q$  is positive definite.

**Theorem 4.3.** *Let  $q \in \mathbb{X}$  and  $Q: \mathbb{X} \rightarrow \mathbb{X}$  a positive definite linear operator. Then, for any initial point  $x^0 \in \mathbb{X}$ , the semi-smooth Newton sequence generated by (6) is well-defined. Additionally, if  $Q - \text{Id}$  is invertible, then equation (9) has a unique solution  $\bar{x} \in \mathbb{X}$  and, if  $\|Q - \text{Id}\| < 1$  then  $\{x^k\}$  converges  $Q$ -linearly to  $\bar{x}$  satisfying*

$$\|x^{k+1} - \bar{x}\| \leq \|Q - \text{Id}\| \|x^k - \bar{x}\|, \quad k \in \mathbb{N}.$$

*Proof.* Using Lemma 4.1, the proof follows the lines of the proof of Theorem 3.3.  $\square$

We next consider an extension of the quadratic conic programming problem (7) by including additional linear constraints. That is, given an additional linear mapping  $\mathcal{A}: \mathbb{X} \rightarrow \mathbb{Y}$ , where  $\mathbb{Y}$  is also a finite dimensional inner product vector space, and given  $b \in \mathbb{Y}$ , we consider the problem

$$\begin{aligned} \min \quad & \frac{1}{2} \langle x, Qx \rangle + \langle q, x \rangle, \\ \text{s.t.} \quad & \mathcal{A}x = b, \\ & x \in \mathcal{K}. \end{aligned} \tag{13}$$

The Lagrangian function associated with (13) is given by:

$$L(x, \lambda, \mu) = \frac{1}{2} \langle x, Qx \rangle + \langle q, x \rangle + \langle \lambda, \mathcal{A}x - b \rangle - \langle \mu, x \rangle. \tag{14}$$

where  $\lambda \in \mathbb{Y}$  and  $\mu \in \mathbb{X}$ . The complementary KKT conditions for problem (13) are given by:

$$\begin{aligned} Q\bar{x} + q + \mathcal{A}^*\bar{\lambda} - \bar{\mu} &= 0, \\ \mathcal{A}\bar{x} &= b, \\ \langle \bar{\mu}, \bar{x} \rangle &= 0, \end{aligned}$$

where the Lagrange multipliers are  $\bar{\mu} \in \mathcal{K}^*$  and  $\bar{\lambda} \in \mathbb{Y}$ . This system can be rewritten as the following complementary system:

$$\left\langle \left( \begin{array}{c} Q\bar{x} + \mathcal{A}^*\bar{\lambda} + q \\ \mathcal{A}\bar{x} - b \end{array} \right), \left( \begin{array}{c} \bar{x} \\ \bar{\lambda} \end{array} \right) \right\rangle = 0, \quad \left( \begin{array}{c} Q\bar{x} + \mathcal{A}^*\bar{\lambda} + q \\ \mathcal{A}\bar{x} - b \end{array} \right) \in K^*, \quad (\bar{x}, \bar{\lambda}) \in K, \tag{15}$$

where  $K = \mathcal{K} \times \mathbb{Y}$  and its dual is given by  $K^* = \mathcal{K}^* \times \{0\}$ . Thus, by Theorem 4.1, this system may be solved by means of the following projection equation

$$\left( \left( \begin{array}{cc} Q & \mathcal{A}^* \\ \mathcal{A} & 0 \end{array} \right) - \text{Id} \right) P_K(x, \lambda) + \left( \begin{array}{c} x \\ \lambda \end{array} \right) = \left( \begin{array}{c} -q \\ b \end{array} \right), \tag{16}$$

where a solution  $(\bar{x}, \bar{\lambda})$  of (16) is such that  $P_K(\bar{x}, \bar{\lambda}) = (P_{\mathcal{K}}(\bar{x}), \bar{\lambda})$  solves the complementary system (15). Notice that equation (16) can be rewritten as

$$\left( \begin{array}{c} (Q - \text{Id})P_{\mathcal{K}}(x) + \mathcal{A}^*\bar{\lambda} + x \\ \mathcal{A}P_{\mathcal{K}}(x) \end{array} \right) = \left( \begin{array}{c} -q \\ b \end{array} \right), \tag{17}$$

and for any starting point  $x^0 \in \mathbb{X}$ , the correspondent semi-smooth Newton iteration can be rewritten as follows:

$$\left( \begin{array}{c} (Q - \text{Id})V(x^k)x^{k+1} + x^{k+1} + \mathcal{A}^*\lambda^{k+1} \\ \mathcal{A}V(x^k)x^{k+1} \end{array} \right) = \left( \begin{array}{c} -q \\ b \end{array} \right), \tag{18}$$

for  $k \in \mathbb{N}$ . The convergence of the sequence  $\{(x^k, \lambda^k)\}$  generated by (18) is guaranteed by applying Theorems 4.2 and 4.3 with respect to equation (16).

## 5 The nearest correlation matrix problem

In this section, we describe the application of the semi-smooth Newton method to the nearest correlation matrix problem. This specific problem represents a special case of the broader positive semidefinite least squares problem (referenced as (19) below), a well-studied area notable for its significant applications and established algorithms; see [11, 15].

Let  $\mathbb{X} = \mathbb{S}^n$  be the set of symmetric  $n \times n$  matrices with real entries and  $\mathcal{K} = \mathbb{S}_+^n \subset \mathbb{S}^n$  be the cone of positive semidefinite matrices. Given any finite dimensional inner product space  $\mathbb{Y}$  and a linear mapping  $\mathcal{A}: \mathbb{S}^n \rightarrow \mathbb{Y}$ , we consider the problem

$$\begin{aligned} \min \quad & \frac{1}{2} \|X - G\|^2, \\ \text{s.t.} \quad & \mathcal{A}(X) = b, \\ & X \in \mathbb{S}_+^n, \end{aligned} \tag{19}$$

where  $G \in \mathbb{S}^n$  and  $b \in \mathbb{Y}$  are given and the Frobenius norm is defined as  $\|A\| = \sqrt{\langle A, A \rangle}$ , where  $\langle A, B \rangle = \text{trace}(AB)$  for  $A, B \in \mathbb{S}^n$ .

Since  $\frac{1}{2} \|X - G\|^2 = \frac{1}{2} \langle X, X \rangle - \langle X, G \rangle$ , problem (19) is an instance of (13) with  $Q = \text{Id}$  and  $q = -G$ . Thus, substituting in (18) we arrive at the following iteration for the semi-smooth Newton method for finding a solution  $(X, \Lambda) \in \mathbb{S}^n \times \mathbb{Y}$  of (16):

$$\begin{pmatrix} X^{k+1} + \mathcal{A}^*(\Lambda^{k+1}) \\ \mathcal{A}V(X^k)X^{k+1} \end{pmatrix} = \begin{pmatrix} G \\ b \end{pmatrix}. \tag{20}$$

The nearest correlation matrix problem is the particular case where  $\mathbb{Y} = \mathbb{R}^n$  and  $\mathcal{A} = \text{diag}: \mathbb{S}^n \rightarrow \mathbb{R}^n$  which maps a symmetric matrix to its diagonal vector. Its dual  $\mathcal{A}^* = \text{Diag}: \mathbb{R}^n \rightarrow \mathbb{S}^n$  maps a vector to a diagonal matrix with the given vector in its diagonal. The right-hand side vector  $b \in \mathbb{R}^n$  will be fixed at the vector  $e$  of all ones. Namely, let us consider problem

$$\begin{aligned} \min \quad & \frac{1}{2} \|X - G\|^2, \\ \text{s.t.} \quad & \text{diag}(X) = e, \\ & X \in \mathbb{S}_+^n. \end{aligned} \tag{21}$$

Thus, substituting in (17) we aim at solving the following projection equation for  $(\bar{X}, \bar{\lambda}) \in \mathbb{S}^n \times \mathbb{R}^n$ :

$$\bar{X} + \text{Diag}(\bar{\lambda}) = G, \tag{22}$$

$$\text{diag}(P_{\mathbb{S}_+^n}(\bar{X})) = e, \tag{23}$$

by means of the semi-smooth Newton iteration

$$X^{k+1} + \text{Diag}(\lambda^{k+1}) = G, \tag{24}$$

$$\text{diag}(V(X^k)X^{k+1}) = e, \tag{25}$$

which is obtained from (20).

From (22) and (24) we deduce that the off-diagonal entries of  $\bar{X}$  and  $X^{k+1}$  are equal to the correspondent off-diagonal entries of  $G$ . Thus, by defining  $D^{k+1} = \text{Diag}(\text{diag}(X^{k+1}))$  and  $\hat{G} = G - \text{Diag}(\text{diag}(G))$  we obtain

$$X^{k+1} = D^{k+1} + \hat{G}, \tag{26}$$

$$\lambda^{k+1} = \text{diag}(G) - \text{diag}(D^{k+1}). \tag{27}$$

Now, a simple calculation using (26) gives

$$\begin{aligned}\text{diag}(V(X^k)X^{k+1}) &= \text{diag}(V(X^k)(D^{k+1} + \hat{G})), \\ &= \text{diag}(V(X^k)D^{k+1}) + \text{diag}(V(X^k)\hat{G}), \\ &= \text{Diag}(\text{diag}(V(X^k)) \text{diag}(D^{k+1}) + \text{diag}(V(X^k)\hat{G})).\end{aligned}$$

Thus, substituting in (25) we arrive at the following expression for the iteration  $D^{k+1}$ :

$$\text{diag}(D^{k+1}) = (\text{Diag}(\text{diag}(V(X^k)))^{-1}[e - \text{diag}(V(X^k)\hat{G})], \quad (28)$$

which gives a fully computable iteration for our Newton method applied to the nearest correlation matrix problem. In particular, we compute  $V(X) = UDU^T$  from the spectral decomposition  $X = U\Lambda U^T$ , where  $D$  is a diagonal matrix with  $D_{ii} = 1$  if  $\Lambda_{ii} > 0$  and  $D_{ii} = 0$  if  $\Lambda_{ii} \leq 0$ . When iteration (28) is not defined, we use in our implementation the pseudoinverse, however this never occurred in the tests we run. We prove next that the iteration is well defined when  $\text{diag}(X^k) > 0$ :

**Proposition 5.1.** *Let  $X \in \mathbb{S}^n$ . If  $\text{diag}(X) > 0$ , then  $\text{diag}(V(X)) > 0$ .*

*Proof.* Let  $X = U\Lambda U^T$  and  $i \in \{1, \dots, n\}$ . Denote by  $u_j$  the  $j$ -th row of  $U$ ,  $j = 1, \dots, n$ . We have

$$X_{ii} = \sum_{j=1}^n \Lambda_{jj} (u_j)_i^2 > 0.$$

Then there exists  $k$  such that  $\Lambda_{kk} (u_k)_i^2 > 0$ . In particular,  $\Lambda_{kk} > 0$  and  $(u_k)_i^2 > 0$ . Then,  $V(X) = UDU^T$  with  $D_{kk} = 1$ . Therefore,

$$V(X)_{ii} = \sum_{j=1}^n D_{jj} (u_j)_i^2 = D_{kk} (u_k)_i^2 + \sum_{j \neq k} D_{jj} (u_j)_i^2 \geq (u_k)_i^2 > 0,$$

proving the result. □

It turns out that in [18], a method closely resembling ours was proposed for the nearest correlation matrix problem (29). In their approach, a semi-smooth Newton method is applied to the following function:

$$\tilde{F}(y) := \mathcal{A}P_{\mathbb{S}_+^n}(G + \mathcal{A}^*y) - e,$$

where  $\mathcal{A} = \text{diag}$ . A significant contribution of their work is the provision of a formula for computing the matrix-vector operation  $h \mapsto V_y h$ , where  $V_y$  belongs to the (B-)subdifferential of  $\tilde{F}$  at  $y$ , without the need to compute  $V_y$  explicitly. The drawback to this approach is that they must then resort to an iterative procedure for computing each Newton iteration, without exploiting the diagonal structure of the operator  $\mathcal{A}$ . In our approach, we compute the subdifferential explicitly, which allows us to solve explicitly the diagonal Newtonian linear system. We note that both methods require computing the full spectral decomposition of an  $n \times n$  matrix at each iteration. The algorithm in [18] also adds a line search procedure in order to ensure global convergence. For a fairer comparison in our study, we turned off this procedure, although its impact on the method's performance was minimal.

## 6 Numerical experiments

In order to observe the behavior of the method, we conducted experiments 5.5 through 5.8 as described in [18]. In these four experiments the *Nearest Correlation Problem* is solved using randomly generated data. We remind the definition of the problem

$$\begin{aligned} \min \quad & \frac{1}{2} \|X - G\|^2, \\ \text{s.t.} \quad & \text{diag}(X) = e, \\ & X \in \mathbb{S}_+^n. \end{aligned} \tag{29}$$

Experiment 5.5 involves generating the matrix  $G$  as  $G := C + \alpha R$ , where  $C$  is a random correlation matrix generated by the `randcorr` Matlab command,  $R_{ij} \in [-1, 1]$  is random, and  $\alpha \geq 0$  is chosen. In Experiment 5.6,  $G$  is defined as a matrix of random numbers  $G_{ij} \in [-1, 1]$  with  $G_{ii}$  fixed at 1 for  $i = 1, \dots, n$ . Experiment 5.7 is analogous to 5.6, but in this case  $G_{ij} \in [0, 2]$ . Finally, Experiment 5.8 defines  $G$  as

$$G := \begin{pmatrix} \frac{\ell}{1-\ell}(E_\ell - \text{Id}_\ell) & 0 \\ 0 & 0 \end{pmatrix} + D + \alpha R,$$

where  $1 \leq \ell \leq n$ ,  $\alpha \geq 0$ ,  $\text{Id}_\ell$  is the identity matrix of dimension  $\ell$ ,  $E_\ell$  is a matrix of 1's of dimension  $\ell$ ,  $D$  is a random diagonal matrix with  $D_{ii} \in [-20000, 20000]$ , and  $R$  is a random matrix such that  $R_{ij} \in [-1, 1]$ . The comparison is made using performance profiles [7]. We built 10 random problems for each choice of parameters. In Experiment 5.5,  $n = 3000$  is fixed and we test all values of  $\alpha \in \{0.01, 0.1, 1, 10\}$ . For Experiments 5.6 and 5.7 we test  $n = 1000, 2000, 3000$ , and for Experiment 5.8 we test  $n = 5000, 8000, 10000$  for  $\alpha = 0.001$  and  $\ell = n/2$ .

The experiments were ran in a 2.30 GHz Intel(R) Core(TM) i5-8300H processor, 16 GB of RAM and operating system Windows 10 using Matlab 9.5.0.944444 (R2018b). We stopped the execution of the methods when the Euclidean residual error is smaller than  $10^{-5}$  and we report the CPU time required by both methods.

In Experiment 5.5, our method was in general slower than that of Qi and Sun [18]. For  $\alpha = 0.01$ , our method took on average 65% more CPU time. This percentage was 137% and 499% for  $\alpha = 1$  and 10, respectively. However, for  $\alpha = 0.1$ , their method took on average 13% more CPU time than ours. In Figure 1 we present a performance profile for Experiment 5.5 with  $\alpha = 0.1$ .

In Experiments 5.6 and 5.7, we observed that our method was slower compared to [18]. For Experiment 5.6 our method took on average 1.6 times the CPU time required by the method from [18] for  $n = 1000$ , 2 times for  $n = 2000$  and 2.5 times for  $n = 3000$ . The situation was worse regarding Experiment 5.7 where we observe that our method took around 4 times the total CPU time required by the method from [18] for  $n = 1000$ . It was 7 times slower for  $n = 2000$  and 9 times slower for  $n = 3000$ . The situation is much more favorable in Experiment 5.8. In Figures 2a, 2b, and 2c, we present the performance profiles concerning  $n = 5000$ ,  $n = 8000$ , and  $n = 10000$ , respectively. In Table 1 we show the average total time and the average number of iterations for both methods and all dimensions tested.

We observed that the method of Qi and Sun was faster in all problem with dimension 5000 and in 75% of the problems with dimension 8000, while our method was faster in all problems of dimension 10000. In Table 1 we can see that our method in general requires more iterations to converge, yet the computation of the iteration is cheaper. This pattern was consistent across other experiments we conducted. Despite this, the difference in the number of iterations was not significantly enough in order for the better iteration cost to yield a better overall performance. In Experiment 5.8, we noted that as the dimension increases, our method becomes more efficient compared to the method

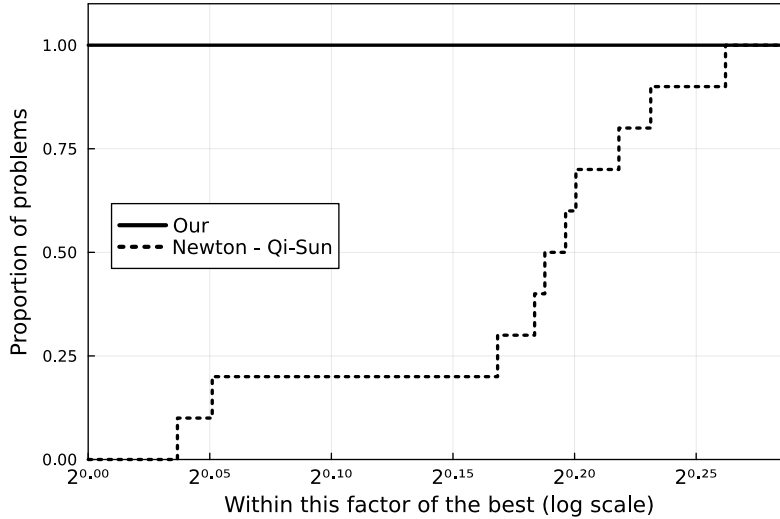


Figure 1: Performance profile for Experiment 5.5 with  $\alpha = 0.1$ .

of Qi and Sun as the number of iterations becomes similar. We could not, however, replicate this phenomenon for different values of  $\alpha$ . This suggests that although we can exploit the diagonal structure of the linear system in order to compute Newton’s direction, while Qi and Sun’s method resorts to a conjugate gradient method, the subgradient they compute is somewhat more efficient towards solving the problem than the one we obtain in our approach. Specifically, in Experiments 5.6 and 5.7 where the matrix  $G$  is far from being positive semidefinite, the better subgradient of Qi and Sun pays off considerably. Nonetheless, our method shows potential superiority for the low noise level regime ( $\alpha \approx 0$ ) and high values of  $n$  in Experiments 5.5 and 5.8.

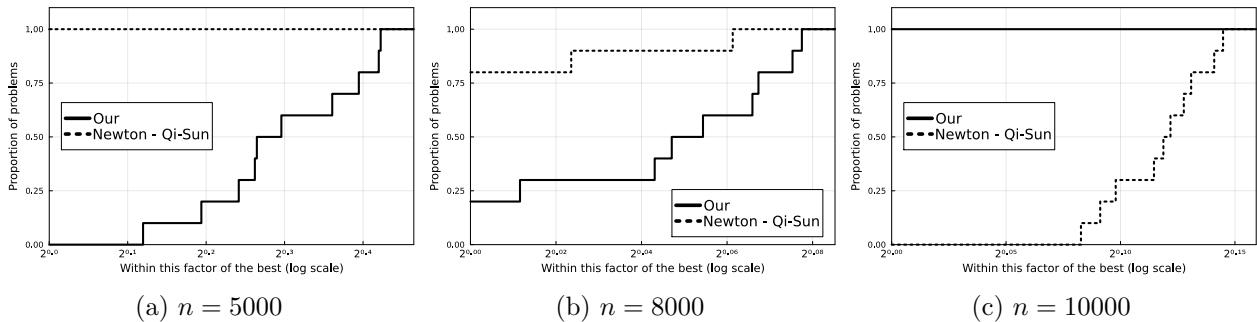


Figure 2: Performance profiles for Experiment 5.8.

	$n = 5000$		$n = 8000$		$n = 10000$	
Method	time (s)	it	time (s)	it	time (s)	it
Our	972,99	13	3677,27	10	6880,66	10
Qi-Sun	791,12	8	3588,47	8	7463,56	9

Table 1: Average total time and number of iterations for  $n = 5000$ ,  $n = 8000$ , and  $n = 10000$ .

## 7 Concluding remarks

In this paper, we investigated the global convergence properties of the semi-smooth Newton method when applied to a general projection equation within finite-dimensional spaces. We have further highlighted the intrinsic connection between the solutions of this projection equation and the constrained quadratic conic programming problem. Comparative experiments on semidefinite least squares problems, particularly the nearest correlation matrix problem, benchmarked against existing literature, underscore the efficacy of the proposed method. The methodology introduced herein paves the way for future research into the versatility and performance of the semi-smooth Newton method in addressing a broader conic constrained problems via generalized projection equations.

## References

- [1] F. Armijo, N., Bello Cruz, Y.J. and Haeser, G.: On the convergence of iterative schemes for solving a piecewise linear system of equations. *Linear Algebra and its Applications*, 665:291–314, (2023).
- [2] Barrios, J.G., Bello Cruz, Y.J., Ferreira, O.P., and Németh, S.Z.: A semi-smooth Newton method for a special piecewise linear system with application to positively constrained convex quadratic programming. *J. Comput. Appl. Math.*, 301:91–100, (2016).
- [3] Barrios, J.G., Ferreira, O.P., and Németh, S.Z.: Projection onto simplicial cones by Picard’s method. *Linear Algebra Appl.*, 480:27–43, (2015).
- [4] Bello Cruz, Y.J., Ferreira, O.P., Németh, S.Z. and Prudente, L.F.: A semi-smooth Newton method for projection equations and linear complementarity problems with respect to the second order cone. *Linear Algebra Its Appl.*, 513:160–181, (2017).
- [5] Chen, J. and Agarwal, R.P.: On Newton-type approach for piecewise linear systems. *Linear Algebra and its Applications*, 433(7):1463–1471, (2010).
- [6] Clarke, F.H.: *Optimization and Nonsmooth Analysis*. Society for Industrial and Applied Mathematics, second edition edition, (1990).
- [7] Dolan, E.D. and Moré, J.J.: Benchmarking optimization software with performance profiles. *Mathematical programming*, 91:201–213, (2002).
- [8] Ferreira, O.P. and Németh, S.Z.: Projection onto simplicial cones by a semi-smooth Newton method. *Optim. Lett.*, 9(4):731–741, (2015).
- [9] Fitzpatrick, S. and Phelps, R.R.: Differentiability of the metric projection in hilbert space. *Transactions of the American Mathematical Society*, 270(2):483–501, (1982).
- [10] Griewank, A., Bernt, J.U., Radons, M. and Streubel, T.: Solving piecewise linear systems in abs-normal form. *Linear Algebra and its Applications*, 471:500–530, (2015).
- [11] Higham, N.J.: Computing the nearest correlation Matrix—a problem from finance. *IMA Journal of Numerical Analysis*, 22(3):329–343, (2002).

- [12] Hiriart-Urruty, J.B. and Lemaréchal, C.: *Convex Analysis and Minimization Algorithms II*, volume 306 of *Grundlehren Der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, (1993).
- [13] Horn, R.A. and Johnson, C.R.: *Matrix Analysis*. Cambridge University Press, 2 edition, (2012).
- [14] Bello Cruz, Y.J., Prudente, L.F. and Ferreira, O.P.: On the global convergence of the inexact semi-smooth newton method for absolute value equation. *Computational Optimization and Applications*, 65(1):93–108, (2016).
- [15] Malick, J.: A Dual Approach to Semidefinite Least-Squares Problems. *SIAM J. Matrix Anal. Appl.*, 26(1):272–284, (2004).
- [16] Mangasarian, O.L.: A generalized Newton method for absolute value equations. *Optim Lett*, 3(1):101–108, (2009).
- [17] Ortega, J.: *Numerical Analysis: A Second Course*. Society for Industrial and Applied Mathematics, Philadelphia, (1987).
- [18] Qi, H. and Sun, D.: A quadratically convergent Newton method for computing the nearest correlation matrix. *SIAM J. Matrix Anal. Appl.*, 28(2):360–385, (2006).
- [19] Sun, Z., Wu, L. and Liu, Z.: A damped semismooth Newton method for the Brugnano–Casulli piecewise linear system. *Bit Numer Math*, 55(2):569–589, (2015).