



# A Full Nesterov-Todd Step Infeasible Interior-point Method for Symmetric Optimization in the Wider Neighborhood of the Central Path

G. Lesaja<sup>1,\*</sup>, G.Q. Wang<sup>2</sup>, A. Oganian<sup>3</sup>

<sup>1</sup> *Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA 30460, USA.*

<sup>2</sup> *School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science, Shanghai 201620, P.R. China.*

<sup>3</sup> *National Center for Health Statistics, 3311 Toledo Rd, Hyattsville, MD, 20782, USA.*

**Abstract** In this paper, an improved Interior-Point Method (IPM) for solving symmetric optimization problems is presented. Symmetric optimization (SO) problems are linear optimization problems over symmetric cones. In particular, the method can be efficiently applied to an important instance of SO, a Controlled Tabular Adjustment (CTA) problem which is a method used for Statistical Disclosure Limitation (SDL) of tabular data. The presented method is a full Nesterov-Todd step infeasible IPM for SO. The algorithm converges to  $\varepsilon$ -approximate solution from any starting point whether feasible or infeasible. Each iteration consists of the feasibility step and several centering steps, however, the iterates are obtained in the wider neighborhood of the central path in comparison to the similar algorithms of this type which is the main improvement of the method. However, the currently best known iteration bound known for infeasible short-step methods is still achieved.

**Keywords** Interior-point methods; Euclidean Jordan algebras; Linear optimization over symmetric cones; Full Nesterov-Todd step; Polynomial complexity, Control tabular adjustment problem, Statistical Disclosure Limitation, Tabular data

**AMS 2010 subject classifications** 90C05, 90C51

**DOI:**10.19139/soic-2310-5070-1175

## 1. Introduction

Interior-Point Methods (IPMs) are theoretically powerful and numerically efficient iterative methods that are based on Newton's method. However, unlike Newton's method, IPMs guarantee convergence to the  $\varepsilon$ -approximate solution of a problem in a polynomial number of iterations. It would be ambitious to claim such a result for any optimization problem, however, this is the case for quite a large class of optimization problems that include well known and important optimization problems such as linear optimization (LO), quadratic optimization (QO), semidefinite optimization (SDO), conic optimization problems and many others. There is extensive literature on IPMs. The following references [21, 26] and the references therein may serve as a good start.

IPMs have shown to be a good alternative to the classic simplex method and they can solve efficiently LO problems of large size. Moreover, they can be applied to important optimization problems not previously accessible by simplex-type methods such as conic optimization problems. The development of IPM presented in this paper was motivated by the desire to provide a theoretical foundation for the efficient solution of the conic formulation of the Controlled Tabular Adjustment (CTA) problem [13].

---

\*Correspondence to: Goran Lesaja (Email: [goran@georgiasouthern.edu](mailto:goran@georgiasouthern.edu), [lesaja@usna.edu](mailto:lesaja@usna.edu)). Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA 30460, USA. On leave at Department of Mathematics, US Naval Academy, 527C Holloway Road, Chauvenet Hall, Annapolis, MD 21402.

The CTA is a method of Statistical Disclosure Limitation (SDL) that was first introduced in [2, 6]. SDL is an increasingly important area of research and practice for the statistical agencies that collect data from individuals or enterprises and then release it to the public, researchers, and policymakers for statistical analysis and research. Prior to such a release, the collected data have to undergo some SDL procedure in order to guarantee the privacy and confidentiality of data providers. The goals of such procedures are two-fold: minimize the risk of disclosure of confidential information about data providers and, at the same time, maximize the amount of released information, that is, maximize the utility of the data for the legitimate data users. These are conflicting goals and therefore SDL practice as a whole can be thought of as a search for the solution of complex and multifaceted optimization problem: maximize the utility of the released data subject to some upper bound on disclosure risk. The way how utility and risk are formulated depends on the scenario of data release and on the data format.

Data can be released in two basic formats: microdata - a collection of individual records, and tabular data - a table of cumulative data that is obtained from cross-tabulations of attributes from microdata. CTA is a perturbative method of protecting tabular data when a specified subset of its cells, called sensitive cells, must be modified to avoid re-identification of an individual respondent. The goal of CTA is to guarantee that the modified value of a sensitive cell is outside of the *disclosure*- an interval that is determined by the data protector (usually a statistical agency). The remaining cells are minimally adjusted to satisfy table equations which usually represent the requirement that the sum of elements in each row and column should be constant and remain unchanged. Hence, the goal of the CTA is to find the closest *safe* table to the original table with respect to the constraints outlined above. The closeness of the original and modified table is measured by the weighted distance between the tables with respect to a certain norm. Most commonly used norms are  $\ell_1$  and  $\ell_2$  norms. Thus, the problem can be formulated as a minimization problem with the objective function being a particular weighted distance function and constraints being table equations and lower and upper bounds on the cell values.

$\ell_2$ -CTA reduces to a QO problem while  $\ell_1$ -CTA is a convex but nonsmooth problem that can be reformulated as LO problem, however, the number of variables and inequality constraints doubles. Alternatively, in [13] a novel second-order cone (SOC) reformulation of  $\ell_1$ -CTA is proposed that does not increase the dimension of the problem as much. As it is shown in [13], conic reformulation of  $\ell_1$ -CTA is a viable alternative to LO reformulation of the problem.

Our motivation was to design an IPM to solve conic  $\ell_1$ -CTA that has good theoretical convergence properties and it is practical to implement which includes the fact that the method can start with any starting point, feasible or infeasible. However, the method is more general, that is, it is designed to solve a more general class of problems, a class of LO problems over symmetric cones of which conic  $\ell_1$ -CTA is just one instance. Nevertheless, the general formulations of CTA and conic reformulation will be listed in the next section as examples of problems to which the proposed method can be applied. These problems are called symmetric optimization (SO) problems. The paper is entirely devoted to the design and analysis of the method, including convergence and complexity analysis because these results are important in their own right. Implementation, numerical testing, and application to conic  $\ell_1$ -CTA and other conic problems will be the subject of a separate paper.

Symmetric cones (SC) is an important class of cones that has been known for quite some time, however, more in the field of algebraic geometry than in the field of optimization. They can be defined in different ways but the one that has shown to be useful in optimization is that symmetric cones are cones consisting of squares of elements of the related Euclidean Jordan Algebras (EJAs). The basic definitions and concepts related to EJAs and corresponding SC that are pertinent to the development of the method in the paper are listed in the next section. Additionally, the classical monograph of Faraut and Korányi [7] provides a wealth of information on Jordan algebras, SCs, and related topics.

Güler [12] was first to realize that symmetric cones, serve as a unifying framework to which the important types of cones used in optimization, such as non-negative orthant, second-order cone, and cone of positive semidefinite matrices belong. That opened a whole new field of research of designing and analyzing optimization algorithms for the SO problems which is still very active today. Faybusovich [8] was the first to generalize IPMs from LO to SO by using EJAs, and associated SC. Subsequently, different versions of IPMs for SO and related optimization problems on SC have been developed (see, e.g., [11, 16, 19, 22, 24]). For an overview of the relevant results, we refer a reader to the monograph on this subject [1] and the references therein.

A full Newton-step infeasible IPM for LO was first analyzed by Roos [20]. The method was generalized by Gu et al. [11] to SO by using the full Nesterov-Todd (NT) direction as a search direction. The obtained iteration bound coincides with the one derived for LO, where  $n$  is replaced by  $r$ , the rank of EJAs, and matches currently best-known iteration bounds for infeasible IPMs for SO.

In this paper, we present an infeasible full NT-step IPM for SO that is a generalization of the feasible IPM discussed in [25]. In particular, Lemma 2.3 and Lemma 2.5 from [25] were used to obtain convergence of the method in the wider neighborhood of the central path while still maintaining the best iteration bound known for these types of methods.

The outline of the paper is as follows. In Section 2 we briefly recall some important definitions and results on EJAs and symmetric cones that are needed in the paper. In addition, we give a brief description and formulation of CTA problem. In Section 3, we briefly recall the framework of the full-NT step feasible IPM for SO with its improved convergence and complexity analysis. The full-NT step infeasible IPM for SO with its convergence and complexity analysis is presented in Section 4. Finally, some concluding remarks follow in Section 5.

## 2. Preliminaries

### 2.1. Euclidean Jordan Algebras and Symmetric Cones

In this section, we recall some important definitions and results on EJAs and associated symmetric cones that will be used in the rest of the paper.

A comprehensive treatment of EJAs and SCs can be found in the monograph [7] and in [1, 8, 9, 11, 22, 24] as it relates to optimization.

#### Definition 1

Let  $(\mathcal{V}, \langle \cdot, \cdot \rangle)$  be an  $n$ -dimensional inner product space over  $\mathbb{R}$  and  $\circ : (x, y) \mapsto x \circ y$  be a bilinear map from  $\mathcal{V} \times \mathcal{V}$  to  $\mathcal{V}$ . Then  $(\mathcal{V}, \circ, \langle \cdot, \cdot \rangle)$  (denoted by  $\mathcal{V}$ ) is an EJA, if it satisfies the following conditions:

- (i)  $x \circ y = y \circ x$  for all  $x, y \in \mathcal{V}$  (Commutativity).
- (ii)  $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$  for all  $x, y \in \mathcal{V}$ , where  $x^2 = x \circ x$  (Jordan's Axiom).
- (iii)  $\langle x \circ y, z \rangle = \langle x, y \circ z \rangle$  for all  $x, y, z \in \mathcal{V}$ .

The operation,  $x \circ y$  is called the Jordan product of  $x$  and  $y$ . Moreover, we always assume that there exists an identity element  $e \in \mathcal{V}$  such that  $e \circ x = x \circ e = x$  for all  $x \in \mathcal{V}$ .

For any element  $x \in \mathcal{V}$ , the Lyapunov transformation  $L(x) : \mathcal{V} \rightarrow \mathcal{V}$  is given by

$$L(x)y := x \circ y, \quad \forall y \in \mathcal{V}. \quad (1)$$

Furthermore, we define the quadratic representation of  $x$  in  $\mathcal{V}$  as follows

$$P(x) := 2L(x)^2 - L(x^2), \quad (2)$$

where  $L(x)^2 = L(x)L(x)$ .

In what follows we list some basic facts about symmetric cones.

Let  $V$  be a finite dimensional real Euclidean space. A nonempty subset  $\mathcal{K}$  of  $V$  is a cone if  $x \in \mathcal{K}$  and  $\lambda \geq 0$  imply  $\lambda x \in \mathcal{K}$ . Cone  $\mathcal{K}$  is a convex cone iff it is a cone and a convex set. The dual cone of a cone  $\mathcal{K}$  is defined as a set  $\mathcal{K}^* = \{y \in V : \forall x \in \mathcal{K}, \langle x, y \rangle \geq 0\}$ . It is straightforward to see that  $\mathcal{K}^*$  is a closed convex cone. If  $\mathcal{K} = \mathcal{K}^*$ , then  $\mathcal{K}$  is a self-dual cone. Cone  $\mathcal{K}$  is pointed cone if  $\mathcal{K} \cap -\mathcal{K} = \{0\}$ . In what follows we consider, convex, pointed cone  $\mathcal{K}$ . The interior of  $\mathcal{K}$  is denoted as  $\text{int}\mathcal{K}$ .

A cone  $\mathcal{K}$  is a SC if it is homogeneous and self-dual. A cone  $\mathcal{K}$  is homogeneous if an automorphism group  $\text{Aut}(\mathcal{K})$  of  $\mathcal{K}$  acts transitively on interior  $\text{int}\mathcal{K}$  of a cone  $\mathcal{K}$ , that is, for all  $x, y \in \text{int}\mathcal{K}$  there exists  $g \in \text{Aut}(\mathcal{K})$  such that  $g(x) = y$ . Automorphism group is defined as  $\text{Aut}(\mathcal{K}) = \{g \in GL(V) : g(\mathcal{K}) = \mathcal{K}\}$  where  $GL(V)$  is a set of all invertible linear maps  $g : V \rightarrow V$ .

Let's consider EJA  $\mathcal{V}$ , and a corresponding set of squares

$$\mathcal{K}(\mathcal{V}) := \{x^2 : x \in \mathcal{V}\} \quad (3)$$

It can be shown that  $\mathcal{K}(\mathcal{V})$  is a SC (see, e.g., [7, 23]). This is the form of SC that will be used throughout the rest of the paper.

The importance of SC for optimization lays in the fact that common and frequently used cones used in optimization, such as non-negative orthant, SOC (ice cream cone), and semidefinite cone, the definitions of which are listed below, are all instances of SC.

1. The linear cone or non-negative orthant:

$$\mathcal{K} = \mathbf{R}_+^n := \{x \in \mathbf{R}^n : x_i \geq 0, i = 1, \dots, n\}.$$

2. The positive semidefinite cone:

$$\mathcal{K} = \mathbf{S}_+^n := \{X \in \mathbf{S}^n : X \succeq 0\},$$

where  $\succeq$  means that  $X$  is positive semidefinite matrix and  $\mathbf{S}^n$  is a set of symmetric  $n$ -dimensional matrices.

3. The quadratic cone or SOC:

$$\mathcal{K} = \mathcal{L}^n = \{x \in \mathbf{R}^n : x_i \geq \sqrt{x_1^2 + \dots + x_{i-1}^2 + x_{i+1}^2 + \dots + x_n^2}\}.$$

In what follows we define an important concept of a rank of EJA and describe two important decompositions of EJA, a spectral decomposition of an element in EJA and a Peirce decomposition of an EJA.

For any  $x \in \mathcal{V}$ , let  $r$  be the smallest integer such that the set  $\{e, x, \dots, x^r\}$  is linearly dependent. Then  $r$  is the *degree* of  $x$  which is denoted as  $\text{deg}(x)$ . Clearly, this degree of  $x$  is bounded by the dimension of the vector space  $\mathcal{V}$ . Furthermore, there exist a polynomial  $p \neq 0$  such that  $p(x) = 0$ . If this polynomial has a leading coefficient one (monic polynomial) and the polynomial is of the minimal degree, then it is called *minimal polynomial* of element  $x$ . The *rank* of  $\mathcal{V}$ , denoted by  $\text{rank}(\mathcal{V})$ , is the largest  $\text{deg}(x)$  of any element  $x \in \mathcal{V}$ . An element  $x \in \mathcal{V}$  is called *regular* if its degree equals the rank of  $\mathcal{V}$ . In the sequel,  $\mathcal{V}$  denotes an EJA with  $\text{rank}(\mathcal{V}) = r$ , unless stated otherwise.

For a regular element  $x \in \mathcal{V}$ , since  $\{e, x, x^2, \dots, x^r\}$  is linearly dependent, there are real numbers  $a_1(x), \dots, a_r(x)$  such that the minimal polynomial of  $x$  is given by

$$f(\lambda; x) = \lambda^r - a_1(x)\lambda^{r-1} + \dots + (-1)^r a_r(x). \tag{4}$$

Hence  $f(x; x) = 0$ . The polynomial  $f(\lambda; x)$  is called a *characteristic polynomial* of a regular element  $x$ . Hence, for regular elements, the minimal and characteristic polynomial coincide, however, for elements that are not regular, that may not be the case. Additionally, it can be proved that if regular element  $x$  vary, then  $a_1(x), \dots, a_r(x)$  are polynomials in  $x$  (Proposition II.2.1 in [7]). The coefficient  $a_1(x)$  is called the *trace* of  $x$ , denoted as  $\text{tr}(x)$ . And the coefficient  $a_r(x)$  is called the *determinant* of  $x$ , denoted as  $\text{det}(x)$ .

An element  $c \in \mathcal{V}$  is said to be an idempotent if  $c^2 = c$ . Two idempotents  $c_1$  and  $c_2$  are said to be orthogonal if  $c_1 \circ c_2 = 0$ . Moreover, an idempotent is primitive if it is non-zero and cannot be written as the sum of two (necessarily orthogonal) non-zero idempotents. We say that  $\{c_1, \dots, c_r\}$  is a complete system of orthogonal primitive idempotents, or *Jordan frame*, if each  $c_i$  is a primitive idempotent,  $c_i \circ c_j = 0, i \neq j$ , and  $\sum_{i=1}^r c_i = e$ . The Löwner partial ordering “ $\succeq_{\mathcal{K}}$ ” of  $\mathcal{V}$  defined by a cone  $\mathcal{K}$  is defined by  $x \succeq_{\mathcal{K}} s$  if  $x - s \in \mathcal{K}$ . Likewise,  $x \succ_{\mathcal{K}} s$  if  $x - s \in \text{int}\mathcal{K}$ .

The following theorem describes a spectral decomposition of elements of EJA  $\mathcal{V}$ , which plays an important role in the analysis of the IPMs for SO and other optimization problems.

*Theorem 1*

(Theorem III.1.2 in [7]) Let  $x \in \mathcal{V}$ . Then there exist a Jordan frame  $\{c_1, \dots, c_r\}$  and real numbers  $\lambda_1(x), \dots, \lambda_r(x)$  such that

$$x = \sum_{i=1}^r \lambda_i(x) c_i. \tag{5}$$

The numbers  $\lambda_i(x)$  (with their multiplicities) are the eigenvalues of  $x$ . Furthermore, the trace and the determinant of  $x$  are given by

$$\text{tr}(x) = \sum_{i=1}^r \lambda_i(x) \quad \text{and} \quad \text{det}(x) = \prod_{i=1}^r \lambda_i(x), \tag{6}$$

respectively.

For a fixed Jordan frame  $\{c_1, c_2, \dots, c_r\}$  in a EJA  $\mathcal{V}$  and for  $i, j \in \{1, 2, \dots, r\}$ , we define the following eigenspaces

$$\mathcal{V}_{ii} := \{x \in \mathcal{V} : x \circ c_i = x\} = \mathbb{R}c_i, \mathcal{V}_{ij} := \{x \in \mathcal{V} : x \circ c_i = \frac{1}{2}x = x \circ c_j\}, i \neq j.$$

The theorem below provides another important decomposition, the Peirce decomposition, of the space  $\mathcal{V}$ .

*Theorem 2*

(Theorem IV.2.1 in [7]) The space  $\mathcal{V}$  is the orthogonal direct sum of the spaces  $\mathcal{V}_{ij}$  ( $i \leq j$ ), i.e.,

$$\mathcal{V} = \bigoplus_{i \leq j} \mathcal{V}_{ij}.$$

Furthermore,

$$\begin{aligned} \mathcal{V}_{ij} \circ \mathcal{V}_{ij} &\subset \mathcal{V}_{ii} + \mathcal{V}_{jj}, \mathcal{V}_{ij} \circ \mathcal{V}_{jk} \subset \mathcal{V}_{ik}, \text{ if } i \neq k, \\ \mathcal{V}_{ij} \circ \mathcal{V}_{kl} &= \{0\}, \text{ if } \{i, j\} \cap \{k, l\} = \emptyset. \end{aligned}$$

Thus, the Peirce decomposition of  $x \in \mathcal{V}$  with respect to the Jordan frame  $\{c_1, \dots, c_r\}$  is given by

$$x = \sum_{i=1}^r x_i c_i + \sum_{i < j} x_{ij} \tag{7}$$

with  $x_i \in \mathbb{R}, i = 1, \dots, r$  and  $x_{ij} \in \mathcal{V}_{ij}, 1 \leq i < j \leq r$ .

As a consequence of Theorem 2, we have the following corollary.

*Corollary 1*

(Lemma 12 in [22]) Let  $x \in \mathcal{V}$  and its spectral decomposition with respect to the Jordan frame  $\{c_1, \dots, c_r\}$  is given by (5). Then the following statements hold.

- (i) The matrices,  $L(x)$  and  $P(x)$  commute and thus share a common system of eigenvectors; in fact the  $c_i, 1 \leq i \leq r$  are among their common eigenvectors.
- (ii) The eigenvalues of  $L(x)$  have the form  $\frac{\lambda_i + \lambda_j}{2}, 1 \leq i \leq j \leq r$ .
- (iii) The eigenvalues of  $P(x)$  have the form  $\lambda_i \lambda_j, 1 \leq i \leq j \leq r$ .

As already indicated, for any  $x, s \in \mathcal{V}$ , the trace inner product is given by

$$\langle x, s \rangle := \text{tr}(x \circ s). \tag{8}$$

Thus,  $\text{tr}(x) = \langle x, e \rangle$ . Hence, it is easy to verify that

$$\text{tr}(x + s) = \text{tr}(x) + \text{tr}(s) \text{ and } x \preceq_{\mathcal{K}} s \Rightarrow \text{tr}(x) \leq \text{tr}(s). \tag{9}$$

The Frobenius norm induced by this trace inner product is then defined by

$$\|x\|_F := \sqrt{\langle x, x \rangle}. \tag{10}$$

It follows from Theorem 1 that

$$\|x\|_F = \sqrt{\text{tr}(x^2)} = \sqrt{\sum_{i=1}^r \lambda_i^2(x)}. \tag{11}$$

One can easily verify that

$$|\lambda_{\min}(x)| \leq \|x\|_F \text{ and } |\lambda_{\max}(x)| \leq \|x\|_F. \tag{12}$$

Furthermore, we have

$$\|x^2\|_F \leq \|x\|_F^2. \tag{13}$$

In the following lemmas, we recall several important inequalities used later in the paper.

*Lemma 1*

(Lemma 2.13 in [11]) Let  $x, s \in \mathcal{V}$  and  $\langle x, s \rangle = 0$ . Then

$$-\frac{1}{4}\|x + s\|_F^2 e \preceq_{\mathcal{K}} x \circ s \preceq_{\mathcal{K}} \frac{1}{4}\|x + s\|_F^2 e.$$

*Lemma 2*

(Lemma 2.16 in [11]) Let  $x, s \in \mathcal{V}$ . Then

$$\|x \circ s\|_F \leq \frac{1}{2} \|x^2 + s^2\|_F.$$

Next lemma provides an important inequality connecting eigenvalues of  $x \circ s$  with the sum of Frobenius norms of  $x$  and  $s$ .

*Lemma 3*

(Lemma 2.3 in [25]) Let  $x, s \in \mathcal{V}$ . Then

$$\sum_{i=1}^r |\lambda_i(x \circ s)| \leq \frac{1}{2} (\|x\|_F^2 + \|s\|_F^2).$$

If  $\langle x, s \rangle = 0$ , then  $\|x + s\|_F^2 = \|x\|_F^2 + \|s\|_F^2$ . Thus, the following corollary follows immediately from Lemma 3.

*Corollary 2*

Let  $x, s \in \mathcal{V}$  and  $\langle x, s \rangle = 0$ . Then

$$\sum_{i=1}^r |\lambda_i(x \circ s)| \leq \frac{1}{2} \|x + s\|_F^2.$$

*Lemma 4*

(Lemma 2.5 in [25]) Let  $u, v \in \mathcal{V}$  and  $\langle u, v \rangle = 0$ , and suppose  $\|u + v\|_F = 2a$  with  $a < 1$ . Then

$$\langle e, (e + u \circ v)^{-1} - e \rangle \leq \frac{2a^4}{1 - a^4}.$$

Lemmas 3 and 4 are crucial in developing the improved complexity analysis of the algorithm presented in the next Section 3.

**2.2. Continuous Tabular Adjustment Problem**

In this subsection we provide the formulation of CTA problem as an important example of the conic problem to which the IPM developed in this paper can be efficiently applied.

The following CTA formulation is given in [13] and several other papers: Given the following set of parameters:

- (i) A set of cells  $a_i, i \in \mathcal{N} = \{1, \dots, n\}$ . The vector  $a = (a_1, \dots, a_n)^T$  satisfies certain linear system  $Aa = b$  where  $A \in \mathbb{R}^{m \times n}$  is an  $m \times n$  matrix and  $b \in \mathbb{R}^m$  is  $m$ -vector. The system usually describes the fact that sum of elements in each row and column should remain unchanged, i.e. constant.
- (ii) A lower, and upper bound for each cell,  $l_{a_i} \leq a_i \leq u_{a_i}$  for  $i \in \mathcal{N}$ , which are considered known by any attacker.
- (iii) A set of indices of sensitive cells,  $\mathcal{S} = \{i_1, i_2, \dots, i_s\} \subseteq \mathcal{N}$ .
- (iv) A lower and upper protection level for each sensitive cell  $i \in \mathcal{S}$  respectively,  $lpl_i$  and  $upl_i$ , such that the released values must be outside of the interval  $(a_i - lpl_i, a_i + upl_i)$ .
- (v) A set of weights,  $w_i, i \in \mathcal{N}$  used in measuring the deviation of the released data values from the original data values.

A CTA problem is a problem of finding values  $z_i, i \in \mathcal{N}$ , such that  $z_i, i \in \mathcal{S}$  are safe values and the weighted distance between released values  $z_i$  and original values  $a_i$ , denoted as  $\|z - a\|_{l(w)}$ , is minimized, which leads to solving the following optimization problem

$$\begin{aligned} \min_z \quad & \|z - a\|_{l(w)} \\ \text{s.t.} \quad & Az = b, \\ & l_{a_i} \leq z_i \leq u_{a_i}, i \in \mathcal{N}, \\ & z_i, i \in \mathcal{S} \text{ are safe values.} \end{aligned} \tag{14}$$

As indicated in the assumption (iv) above, safe values are the values that satisfy

$$z_i \leq a_i - lpl_i \text{ or } z_i \geq a_i + upl_i, i \in \mathcal{S}. \tag{15}$$

By introducing a vector of binary variables  $y \in \{0, 1\}^s$  the constraint (15) can be written as

$$\begin{aligned} z_i \geq \quad & -M(1 - y_i) + (a_i + upl_i)y_i, \quad i \in \mathcal{S}, \\ z_i \leq \quad & My_i + (a_i - lpl_i)(1 - y_i), \quad i \in \mathcal{S}, \end{aligned} \tag{16}$$

where  $M \gg 0$  is a large positive number. Constraints (16) enforce the upper safe value if  $y_i = 1$  or the lower safe value if  $y_i = 0$ .

Replacing the last constraint in the CTA model (14) with (16) leads to a mixed integer convex optimization problem (MIOP) which is, in general, a difficult problem to solve; however, it provides solutions with high data utility [3]. The alternative approach is to fix binary variables up front, which leads to a CTA that is a continuous convex optimization problem because all binary variables are replaced with values 0 or 1. The continuous CTA is easier to solve; however, the obtained solution may have a lower data utility because the optimal solution of the continuous CTA is either feasible or infeasible solution of the corresponding MIOP depending on the values that were assigned to the binary variables. The strategies on how to avoid a wrong assignment of binary variables that may result in the MIOP being infeasible are discussed in [4, 5].

In what follows, we consider a continuous CTA where binary variables in MIOP are fixed with certain values of 0 or 1, and vector  $z$  is replaced by the vector of *cell deviations*  $x = z - a$ . Then, the CTA (14) reduces to the following convex optimization problem:

$$\begin{aligned} \min_x \quad & \|x\|_{l(w)} \\ \text{s.t.} \quad & Ax = 0, \\ & l \leq x \leq u, \end{aligned} \tag{17}$$

where upper and lower bounds for  $x_i, i \in \mathcal{N}$  are defined as follows:

$$l_i = \begin{cases} upl_i & \text{if } i \in \mathcal{S} \text{ and } y_i = 1 \\ l_{a_i} - a_i & \text{if } (i \in \mathcal{N} \setminus \mathcal{S}) \text{ or } (i \in \mathcal{S} \text{ and } y_i = 0) \end{cases} \tag{18}$$

$$u_i = \begin{cases} -lpl_i & \text{if } i \in \mathcal{S} \text{ and } y_i = 0 \\ u_{a_i} - a_i & \text{if } (i \in \mathcal{N} \setminus \mathcal{S}) \text{ or } (i \in \mathcal{S} \text{ and } y_i = 1). \end{cases} \tag{19}$$

The two most commonly used norms in problem (17) are the  $\ell_1$  and  $\ell_2$  norms. For the  $\ell_2$ -norm the problem, (17) reduces to the following  $\ell_2$ -CTA model:

$$\begin{aligned} \min_x \quad & \sum_{i=1}^n w_i x_i^2 \\ \text{s.t.} \quad & Ax = 0, \\ & l \leq x \leq u. \end{aligned} \tag{20}$$

The above problem is a standard QO problem that can be efficiently solved using IPM or other methods.

For the  $\ell_1$ -norm the problem, (17) reduces to the following  $\ell_1$ -CTA model:

$$\begin{aligned} \min_x \quad & \sum_{i=1}^n w_i |x_i| \\ \text{s.t.} \quad & Ax = 0, \\ & l \leq x \leq u. \end{aligned} \tag{21}$$

The above  $\ell_1$ -CTA model (21) is a convex optimization problem; however, the objective function is not differentiable at  $x = 0$ . Since most of the algorithms, including IPMs, require differentiability of the objective function, problem (21) needs to be reformulated.

The standard reformulation is the transformation of a model (21) to the following LO model:

$$\begin{aligned} \min_{x^-, x^+} \quad & \sum_{i=1}^n w_i (x_i^+ + x_i^-) \\ \text{s.t.} \quad & A(x_i^+ - x_i^-) = 0, \\ & l^+ \leq x^+ \leq u^+, \\ & l^- \leq x^- \leq u^-, \end{aligned} \tag{22}$$

where

$$x^+ = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0, \end{cases} \quad x^- = \begin{cases} 0 & \text{if } x > 0 \\ -x & \text{if } x \leq 0, \end{cases} \tag{23}$$

The drawback of the above LO reformulation is that number of variables and inequality constraints doubles. In [13] an alternative SOC reformulation of  $\ell_1$ -CTA is proposed where the dimension of the problem does not increases as much. It is based on the fact that absolute value has an obvious SOC representation since the epigraph of the absolute value function is exactly SOC, that is,

$$t_i = |x_i| \quad \longrightarrow \quad \mathcal{K}_i = \left\{ (x_i, t_i) \in \mathbb{R}^2 : t_i \geq \sqrt{x_i^2} \right\}.$$

A SOC formulation of the  $\ell_1$ -CTA (21) is given below

$$\begin{aligned} \min_x \quad & \sum_{i=1}^n w_i t_i \\ \text{s.t.} \quad & Ax = 0, \\ & (x_i, t_i) \in \mathcal{K}_i; \quad i = 1, \dots, n, \\ & l \leq x \leq u. \end{aligned} \tag{24}$$

The above conic formulation of continuous CTA problem is an important example of the conic problem to which the IPM developed and analyzed in the rest this paper can be efficiently applied.

### 3. A Brief Outline of the Full NT-Step Feasible IPM

In this section, a brief outline of the feasible algorithm presented in [25] is given.

Let  $(\mathcal{V}, \circ)$  be an  $n$ -dimensional EJA with rank  $r$  equipped with the standard inner product  $\langle x, s \rangle = \text{tr}(x \circ s)$  and  $\mathcal{K}$  be the corresponding symmetric cone. Moreover, we always assume that there exists an identity element  $e \in \mathcal{V}$  such that  $e \circ x = x$  for all  $x \in \mathcal{V}$ . Additional facts regarding EJAs and SCs are listed in the previous Section 2 and references cited in that section.

We consider the LO problem over symmetric cones, or shortly, the SO problem given in the standard form

$$\begin{aligned} (SOP) \quad \min \quad & \langle c, x \rangle \\ \text{s.t.} \quad & Ax = b, \quad x \in \mathcal{K}, \end{aligned}$$



and its dual problem

$$(SOD) \quad \begin{aligned} \max \quad & b^T y \\ \text{s.t.} \quad & A^T y + s = c, \quad s \in \mathcal{K}, \end{aligned}$$

where  $A$  is a linear operator from  $\mathcal{V}$  to  $\mathbf{R}^m$ ,  $c$  and the rows of  $A$  lie in  $\mathcal{V}$ ,  $b \in \mathbf{R}^m$ , and  $A^T$  is the adjoint of  $A$ . Let  $a_i \in \mathcal{V}$  be the  $i$ th row of  $A$ , then  $Ax = b$  means that  $\langle a_i, x \rangle = b_i$ , for each  $i = 1, \dots, m$ , while  $A^T y + s = c$  means that  $\sum_{i=1}^m y_i a_i + s = c$ . Without loss of generality, we assume that the rows of  $A$  are linearly independent.

Additionally, without loss of generality we can assume that both  $(SOP)$  and  $(SOD)$  satisfy the interior-point condition (IPC) [22], i.e., there exists  $(x^0, y^0, s^0)$  such that

$$Ax^0 = b, x^0 \in \text{int}\mathcal{K}, A^T y^0 + s^0 = c, s^0 \in \text{int}\mathcal{K}.$$

The perturbed Karush-Kuhn-Tucker (KKT) conditions for  $(SOP)$  and  $(SOD)$  are given by

$$\begin{aligned} Ax &= b, x \in \mathcal{K}, \\ A^T y + s &= c, s \in \mathcal{K}, \\ x \circ s &= \mu e. \end{aligned} \tag{25}$$

The parameterized system (25) has a unique solution  $(x(\mu), y(\mu), s(\mu))$  for each  $\mu > 0$ . The set of  $\mu$ -centers forms a homotopy path with  $\mu$  running through all positive real numbers, which is called the central path. If  $\mu \rightarrow 0$ , then the limit of the central path exists and since the limit points satisfy the complementarity condition, i.e.,  $x \circ s = 0$ , it naturally yields an optimal solution for  $(SOP)$  and  $(SOD)$  (see, e.g., [8, 22]).

IPMs follow the central path approximately and find an approximate solution of the underlying problems  $(P)$  and  $(D)$  as  $\mu$  gradually decreases to zero. Just like the case of a linear SDO, linearizing the third equation in (25) may not lead to a unique element in  $\mathcal{V}$ . Thus, it is necessary to symmetrize that equation before linearizing it. To overcome this difficulty, the third equation of the system (25) is replaced by the following equivalent scaled equation (Lemma 28 in [22])

$$P(u)x \circ P(u^{-1})s = \mu e,$$

where  $u$  is a scaling point from the interior of the cone  $\mathcal{K}$  (i.e.,  $\text{int}\mathcal{K}$ ).

Applying Newton's method, we have

$$\begin{aligned} A\Delta x &= 0, \\ A^T \Delta y + \Delta s &= 0, \\ P(u)x \circ P(u^{-1})\Delta s + P(u^{-1})s \circ P(u)\Delta x &= \mu e - P(u)x \circ P(u^{-1})s. \end{aligned} \tag{26}$$

The appropriate choices of  $u$  that lead to obtaining the unique search directions from the above system are called commutative class of search directions (see, e.g., [22]). In this paper, we consider the so-called NT-scaling scheme, the resulting direction is called NT search direction. This scaling scheme was first proposed by Nesterov and Todd [17, 18] for self-scaled cones and then adapted by Faybusovich [8, 9] for symmetric cones.

*Lemma 5* (Lemma 3.2 in [9])

Let  $x, s \in \text{int}\mathcal{K}$ . Then there exists a unique  $w \in \text{int}\mathcal{K}$  such that

$$x = P(w)s.$$

Moreover,

$$w = P(x)^{\frac{1}{2}} \left( P(x^{\frac{1}{2}})s \right)^{-\frac{1}{2}} \left[ = P(s^{-\frac{1}{2}}) \left( P(s^{\frac{1}{2}})x \right)^{\frac{1}{2}} \right].$$

The point  $w$  is called the scaling point of  $x$  and  $s$  (in this order). Let  $u = w^{-\frac{1}{2}}$ , where  $w$  is the NT-scaling point of  $x$  and  $s$ . Introducing the variance vector

$$v := \frac{P(w)^{-\frac{1}{2}}x}{\sqrt{\mu}} \left[ = \frac{P(w)^{\frac{1}{2}}s}{\sqrt{\mu}} \right], \quad (27)$$

and the scaled search directions

$$d_x := \frac{P(w)^{-\frac{1}{2}}\Delta x}{\sqrt{\mu}} \text{ and } d_s := \frac{P(w)^{\frac{1}{2}}\Delta s}{\sqrt{\mu}}, \quad (28)$$

the system (26) is further simplified

$$\begin{aligned} \bar{A}d_x &= 0, \\ \bar{A}^T \Delta y + d_s &= 0, \\ d_x + d_s &= v^{-1} - v, \end{aligned} \quad (29)$$

where  $\bar{A} := \frac{1}{\sqrt{\mu}}AP(w)^{\frac{1}{2}}$ . This system has a unique solution  $(d_x, \Delta y, d_s)$ . The original search directions can then be obtained through (28). If  $(x, y, s) \neq (x(\mu), y(\mu), s(\mu))$ , then  $(\Delta x, \Delta y, \Delta s)$  is nonzero. The new iterate is obtained by taking full NT-steps

$$x^+ := x + \Delta x, \quad y^+ := y + \Delta y, \quad \text{and } s^+ := s + \Delta s. \quad (30)$$

From the first two equations of the system (29), one can easily verify that the scaled search directions  $d_x$  and  $d_s$  are orthogonal with respect to the trace inner product, i.e.,  $\langle d_x, d_s \rangle = 0$ . This implies that  $\Delta x$  and  $\Delta s$  also are orthogonal, i.e.,  $\langle \Delta x, \Delta s \rangle = 0$ . As a consequence, we have the important property that, after a full NT-step, the duality gap assumes the same value as at the  $\mu$ -centers, namely  $r\mu$ .

*Lemma 6* (Lemma 3.4 in [11])

After a full NT-step, the duality gap is given by

$$\langle x^+, s^+ \rangle = r\mu.$$

To measure the distance of an iterate to the corresponding  $\mu$ -center, a norm-based proximity measure  $\delta(x, s; \mu)$  is introduced

$$\delta(v) := \delta(x, s; \mu) := \frac{1}{2} \|v^{-1} - v\|_F. \quad (31)$$

One can easily verify that

$$\delta(v) = 0 \Leftrightarrow v = e \Leftrightarrow d_x = d_s = 0 \Leftrightarrow x \circ s = \mu e, \quad (32)$$

which implies that the value of  $\delta(v)$  can indeed be considered as a measure of the distance between the given iterate and the corresponding  $\mu$ -center.

It is crucial for us to investigate the effect on the proximity measure  $\delta(x, s; \mu)$  of a full NT-step to the target point  $(x(\mu), y(\mu), s(\mu))$ . For this purpose, Wang et al. [25] established a sharper quadratic convergence result than the one mentioned in [11]. Their derivation is based on the generalization of Theorem II.52 in [21] for LO. This leads to a wider quadratic convergence neighborhood of the central path for the algorithm than the one used in [11].

*Theorem 3* (Theorem 3.2 in [25])

Let  $\delta := \delta(x, s; \mu) < 1$ . Then, the full NT-step is strictly feasible and

$$\delta(x^+, s^+; \mu) \leq \frac{\delta^2}{\sqrt{2(1 - \delta^4)}}.$$

The following corollary shows the quadratic convergence of the full NT-step to the target  $\mu$ -center  $(x(\mu), y(\mu), s(\mu))$  in the wider neighborhood determined by  $1/\sqrt[3]{2}$  as opposed to  $1/\sqrt{2}$  in [11].

*Corollary 3*

Let  $\delta := \delta(x, s; \mu) \leq \frac{1}{\sqrt[4]{2}}$ . Then the full NT-step is strictly feasible and

$$\delta(x^+, s^+; \mu) \leq \delta^2.$$

The following theorem investigates the effect on the proximity measure when  $(x, y, s)$  is kept fixed and  $\mu$  is updated to  $\mu^+ = (1 - \theta)\mu$ .

*Theorem 4* (Theorem 3.3 in [25])

Let  $\delta := \delta(x, s; \mu) < 1$  and  $\mu^+ = (1 - \theta)\mu$  with  $0 < \theta < 1$ . Then

$$\delta(x, s; \mu^+)^2 = (1 - \theta)\delta^2 + \frac{r\theta^2}{4(1 - \theta)}.$$

As a consequence of Theorem 3 and Theorem 4, the following corollary readily follows.

*Corollary 4*

Let  $\delta := \delta(x, s; \mu) \leq \frac{1}{\sqrt[4]{2}}$  and  $\theta = \frac{1}{\sqrt{2r}}$  with  $r \geq 2$ . Then

$$\delta(x^+, s^+; \mu^+) \leq \frac{1}{\sqrt[4]{2}}.$$

The following theorem provides an upper bound for the total number of the iterations produced by the full-NT step feasible IPM.

*Theorem 5* (Theorem 3.4 in [25])

Let  $\tau = \frac{1}{\sqrt[4]{2}}$  and  $\theta = \frac{1}{\sqrt{2r}}$  with  $r \geq 2$ . Then the feasible algorithm requires

$$O\left(\sqrt{r} \log \frac{\langle x^0, s^0 \rangle}{\varepsilon}\right)$$

iterations to obtain an iterate  $(x, y, s)$  satisfying  $\langle x, s \rangle \leq \varepsilon$  which is an  $\varepsilon$ -approximate optimal solution of (SOP) and (SOD).

Thus, the feasible algorithm is well defined, globally convergent, and achieves quadratic convergence of full NT-steps in the wider neighborhood while still maintaining the best known iteration bound known for these types of methods, namely

$$O\left(\sqrt{r} \log \frac{r\mu_0}{\varepsilon}\right).$$

#### 4. Full NT-Step Infeasible IPM

It is well known fact that finding strictly feasible starting point may be difficult. Thus, an infeasible IPM that does not require feasible starting point may be a good alternative. First, a brief outline of the infeasible algorithm is presented. Next, we concentrate on the convergence and complexity analysis of the algorithm. The method is similar to IPM presented in [11], however, with wider neighborhood and larger steps which impacts the convergence and complexity analysis. Allowing larger steps at each iteration while still maintaining the best known iteration bound for these types of methods and having a quadratic local convergence of the proximity measure at each iteration are another advantages of the method presented in this paper.

##### 4.1. An Outline of the Full NT-Step Infeasible IPM

In what follows, we assume that the SO problem has an optimal solution  $(x^*, s^*)$  with vanishing duality gap, i.e.,  $\langle x^*, s^* \rangle = 0$ . Furthermore, we choose arbitrarily  $(x^0, s^0) \in \text{int}\mathcal{K}$  and  $\mu^0 > 0$  such that

$$x^0 = \zeta e, \quad y^0 = 0, \quad s^0 = \zeta e, \quad \text{and} \quad \mu^0 = \zeta^2, \quad (33)$$

as the starting point of the algorithm, where  $\zeta$  is a (positive) number such that

$$x^* + s^* \preceq_{\mathcal{K}} \zeta e. \tag{34}$$

The initial values of the primal and dual residual vectors are denoted as

$$r_p^0 = b - Ax^0 \text{ and } r_d^0 = c - A^T y^0 - s^0, \tag{35}$$

respectively. In general, we have  $r_p^0 \neq 0$  and  $r_d^0 \neq 0$ , i.e., the initial iterate is not feasible for SO. However, a sequence of perturbed problems is constructed below in a such a way that the initial iterate is strictly feasible for the first perturbed problem in the sequence.

For any  $\nu$  with  $0 < \nu \leq 1$ , the perturbed problems of SO given in the standard form

$$\begin{aligned} (SOP_\nu) \quad \min \quad & \langle c - \nu r_d^0, x \rangle \\ \text{s.t.} \quad & b - Ax = \nu r_p^0, \quad x \in \mathcal{K}, \end{aligned}$$

and its dual problem

$$\begin{aligned} (SOD_\nu) \quad \max \quad & (b - \nu r_p^0)^T y \\ \text{s.t.} \quad & c - A^T y - s = \nu r_d^0, \quad s \in \mathcal{K}. \end{aligned}$$

It is obvious that  $x = x^0$  is a strictly feasible solution of  $(SOP_\nu)$ , and  $(y, s) = (y^0, s^0)$  is a strictly feasible solution of  $(SOD_\nu)$  when  $\nu = 1$ , that is,  $(SOP_\nu)$  and  $(SOD_\nu)$  satisfy the IPC for  $\nu = 1$  which then straightforwardly leads to the following lemma.

*Lemma 7* (Lemma 4.1 in [11])

Let  $(SOP)$  and  $(SOD)$  be feasible and  $0 < \nu \leq 1$ . Then, the perturbed problems  $(SOP_\nu)$  and  $(SOD_\nu)$  satisfy the IPC.

Let  $(SOP)$  and  $(SOD)$  be feasible and  $0 < \nu \leq 1$ . Then Lemma 7 implies that the perturbed problems  $(SOP_\nu)$  and  $(SOD_\nu)$  satisfy the IPC, and therefore the following system

$$\begin{aligned} b - Ax &= \nu r_p^0, \quad x \in \mathcal{K}, \\ c - A^T y - s &= \nu r_d^0, \quad s \in \mathcal{K}, \\ x \circ s &= \mu e, \end{aligned} \tag{36}$$

has a unique solution  $(x(\mu, \nu), y(\mu, \nu), s(\mu, \nu))$ , for every  $\mu > 0$  that is called a  $\mu$ -center of the perturbed problems  $(SOP_\nu)$  and  $(SOD_\nu)$ . Hence, the central paths of  $(SOP_\nu)$  and  $(SOD_\nu)$  exist.

The main idea of the infeasible algorithm is to simultaneously improve feasibility by reducing  $\nu$  and optimality by reducing  $\mu$  while keeping the iterates in the certain neighborhood of the central paths of  $(SOP_\nu)$  and  $(SOD_\nu)$ .

Thus, it make sense to link the parameters  $\mu$  and  $\nu$  according to the following formula  $\mu = \nu \mu^0 = \nu \zeta^2$ . It is also worth noting that, according to (33),  $x^0 \circ s^0 = \mu^0 e$ ; hence,  $x^0$  is the  $\mu^0$ -center of the perturbed problem  $(SOP_\nu)$ , and  $(y^0, s^0)$  is the  $\mu^0$ -center of the perturbed problem  $(SOD_\nu)$  for  $\nu = 1$ . In other words,

$$(x(\mu^0, 1), y(\mu^0, 1), s(\mu^0, 1)) = (x^0, y^0, s^0)$$

and the algorithm can easily be started since we have initial starting point that is by construction exactly on the central path of  $(SOP_\nu)$  and  $(SOD_\nu)$  for  $\nu = 1$ .

The outline of one iteration of the infeasible algorithm is as follows. Suppose that for some  $\nu \in (0, 1]$  we have an iterate  $(x, y, s)$  satisfying the feasibility condition, i.e., the first two equations of the system (36) for  $\mu = \nu \mu^0$ , and such that  $\text{tr}(x \circ s) = r\mu$  and  $\delta(x, s; \mu) \leq \tau$ . Thus, we start with the iterate in the  $\tau$ -neighborhood of the central path of  $(SOP_\nu)$  and  $(SOD_\nu)$  that targets the  $\mu$ -center on that central path. The goal is to obtain a new iterate  $(x^+, y^+, s^+)$  in the  $\tau$ -neighborhood of the central path of the new pair of problems  $(SOP_\nu^+)$  and  $(SOD_\nu^+)$  where

both  $\nu$  and  $\mu$  are reduced by a barrier parameter  $\theta \in (0, 1)$ , i.e.,  $\nu^+ = (1 - \theta)\nu$  and  $\mu^+ = (1 - \theta)\mu = \nu^+\mu^0$ . Hence,  $(x^+, y^+, s^+)$  should satisfy the first two equations of the system (36), with  $\nu$  replaced by  $\nu^+$  and  $\mu$  by  $\mu^+$ , and such that  $\text{tr}(x^+, s^+) = r\mu^+$  and  $\delta(x^+, s^+; \mu^+) \leq \tau$ .

The calculation of the new iterate is achieved in two phases, a feasibility phase where one feasibility step is taken and a centering phase where a few centering steps are performed. The feasibility step serves to get an iterate  $(x^f, y^f, s^f)$  that is strictly feasible for  $(SOP_{\nu^+})$  and  $(SOD_{\nu^+})$ , and belongs to the quadratic convergence neighborhood with respect to the  $\mu^+$ -center of  $(SOP_{\nu^+})$  and  $(SOD_{\nu^+})$ . However,  $(x^f, y^f, s^f)$  may not be in the  $\tau$ -neighborhood of the  $\mu^+$ -center; therefore, several centering steps may need to be performed to get inside the  $\tau$ -neighborhood.

Note that after each iteration the residuals and the duality gap are reduced by the factor  $(1 - \theta)$ . The algorithm stops when we obtain an iterate for which the norm of the residuals and the duality gap do not exceed the accuracy parameter  $\varepsilon$ . This iterate is called an  $\varepsilon$ -approximate optimal solution for  $(SOP)$  and  $(SOD)$ .

The feasibility step is obtained by taking full steps

$$x^f := x + \Delta^f x, \quad y^f := y + \Delta^f y, \quad \text{and} \quad s^f := s + \Delta^f s, \tag{37}$$

with NT-search directions  $(\Delta^f x, \Delta^f y, \Delta^f s)$  that are calculated from the following Newton system

$$\begin{aligned} A\Delta^f x &= \theta\nu r_p^0, \\ A^T \Delta^f y + \Delta^f s &= \theta\nu r_d^0, \\ P(w)^{-\frac{1}{2}}x \circ P(w)^{\frac{1}{2}}\Delta^f s + P(w)^{\frac{1}{2}}s \circ P(w)^{-\frac{1}{2}}\Delta^f x &= (1 - \theta)\mu e - P(w)^{-\frac{1}{2}}x \circ P(w)^{\frac{1}{2}}s. \end{aligned} \tag{38}$$

One may easily verify that  $(x^f, y^f, s^f)$  satisfies the first two equations of the system (36), with  $\nu$  replaced by  $\nu^+$  and  $\mu$  by  $\mu^+$ . The third equation indicates that the  $\mu^+$ -center of  $(SOP_{\nu^+})$  and  $(SOD_{\nu^+})$  is targeted. Targeting  $\mu^+$ -center rather than  $\mu$ -center contributes to the efficiency of the algorithm. The system (38) defines the feasibility step uniquely since the coefficient matrix of the resulting system is exactly the same as in the feasible case.

Similarly to the feasible case, given the variance vector  $v$  defined by (27) and scaled search directions

$$d_x^f := \frac{P(w)^{-\frac{1}{2}}\Delta^f x}{\sqrt{\mu}} \quad \text{and} \quad d_s^f := \frac{P(w)^{\frac{1}{2}}\Delta^f s}{\sqrt{\mu}}, \tag{39}$$

the system (38) is reduced to the following form

$$\begin{aligned} \bar{A}d_x^f &= \frac{1}{\mu}\theta\nu r_p^0, \\ \bar{A}^T \Delta^f y + d_s^f &= \frac{1}{\sqrt{\mu}}\theta\nu P(w)^{\frac{1}{2}}r_d^0, \\ d_x^f + d_s^f &= (1 - \theta)v^{-1} - v. \end{aligned} \tag{40}$$

Hence,

$$x^f = x + \Delta^f x = \sqrt{\mu}P(w)^{\frac{1}{2}}(v + d_x^f) \quad \text{and} \quad s^f = s + \Delta^f s = \sqrt{\mu}P(w)^{-\frac{1}{2}}(v + d_s^f). \tag{41}$$

Since it will be shown that  $(x^f, y^f, s^f)$  is strictly feasible and moreover in the quadratic convergence neighborhood of the  $\mu^+$ -center of  $(SOP_{\nu^+})$  and  $(SOD_{\nu^+})$ , it is possible to take few centering steps to get the new iteration in the desired  $\tau$ -neighborhood of the  $\mu^+$ -center. The centering steps are obtained by taking full steps with NT-search directions calculated from the Newton system that is the same as in the feasible case, (26), or in the scaled form, (29).

The above outline is summarized in the Fig. 1, that describes a generic full NT-step infeasible IPM.

#### 4.2. Analysis of the Full-NT Step Infeasible IPM

The analysis of the infeasible algorithm is more complicated and more involved than in the feasible case. The main reason for this is that the scaled search directions  $d_x^f$  and  $d_s^f$  are not (necessarily) orthogonal with respect to the

---

**Full NT-Step Infeasible IPM**


---

**Input:**

A threshold parameter  $0 < \tau < 1$ ;  
 an accuracy parameter  $\varepsilon > 0$ ;  
 a fixed barrier update parameter  $0 < \theta < 1$ ;  
 an initialization parameter  $\zeta > 0$ .

**begin**

$x^0 := \zeta e$ ;  $y^0 := 0$ ;  $s^0 := \zeta e$ ;  $\mu^0 := \zeta^2$ ;  $\nu^0 := 1$ ;  
**while**  $\max(\langle x, s \rangle, \|b - Ax\|_F, \|c - A^T y - s\|_F) \geq \varepsilon$  **do**  
**begin**  
 feasibility step:  
   calculate  $(\Delta^f x, \Delta^f y, \Delta^f s)$  using (40) and (39);  
   update  $(x, y, s) := (x, y, s) + (\Delta^f x, \Delta^f y, \Delta^f s)$ ;  
 centering steps:  
**while**  $\delta(x, s; \mu) \geq \tau$  **do**  
**begin**  
   calculate  $(\Delta x, \Delta y, \Delta s)$  using (29) and (28);  
   update  $(x, y, s) := (x, y, s) + (\Delta x, \Delta y, \Delta s)$ .  
**end**  
 update of  $\mu$  and  $\nu$ :  
    $\mu := (1 - \theta)\mu$ ;  $\nu := (1 - \theta)\nu$ .

**end****end**

Figure 1. Algorithm I

trace inner product. We omit most parts of the analysis that are unchanged from the one presented in [11] and emphasize the parts where there are differences. It is shown that the feasibility steps can be taken in the wider quadratic convergence neighborhood of the central path developed in the feasible case.

**Feasibility Step.** The lemma below provides the sufficient condition for the strict feasibility of the feasibility step  $(x^f, y^f, s^f)$ .

*Lemma 8* (Lemma 4.2 in [11])

The feasibility step  $(x^f, y^f, s^f)$  is strictly feasible if  $(1 - \theta)e + d_x^f \circ d_s^f \in \text{int}\mathcal{K}$ .

Thus, the feasibility of the  $(x^f, y^f, s^f)$  highly depends on the eigenvalues of the vector  $d_x^f \circ d_s^f$ . More specifically,  $(x^f, y^f, s^f)$  is strictly feasible if  $\|\lambda(d_x^f \circ d_s^f)\|_\infty < 1 - \theta$ .

In order to measure the distance from the  $(x^f, y^f, s^f)$  to the  $\mu^+$ -center we need an upper bound on the proximity measure  $\delta(x^f, s^f; \mu^+)$  which is for simplicity denoted also as  $\delta(v^f)$ , where  $v^f$  is a variance vector defined by (27).

The following lemma provides an upper bound for  $\delta(v^f)$ .

*Lemma 9* (Lemma 4.4 in [11])

Let  $\|\lambda(d_x^f \circ d_s^f)\|_\infty < 1 - \theta$ . Then

$$4\delta(v^f)^2 \leq \frac{\left\| \frac{d_x^f \circ d_s^f}{1 - \theta} \right\|_F^2}{1 - \left\| \frac{\lambda(d_x^f \circ d_s^f)}{1 - \theta} \right\|_\infty}.$$

The following lemma gives an important relationship between the infinite norm of the vector of eigenvalues of  $d_x \circ d_s$  and the Frobenius norms of  $d_x$  and  $d_s$ .

*Lemma 10*

One has

$$\|\lambda(d_x^f \circ d_s^f)\|_\infty \leq \|d_x^f \circ d_s^f\|_F \leq \frac{1}{2}(\|d_x^f\|_F^2 + \|d_s^f\|_F^2).$$

*Proof*

We have the following derivation:

$$\begin{aligned} \|\lambda(d_x^f \circ d_s^f)\|_\infty &\leq \|d_x^f \circ d_s^f\|_F \leq \frac{1}{2}\|(d_x^f)^2 + (d_s^f)^2\|_F \\ &\leq \frac{1}{2}\|(d_x^f)^2\|_F + \frac{1}{2}\|(d_s^f)^2\|_F \leq \frac{1}{2}(\|d_x^f\|_F^2 + \|d_s^f\|_F^2). \end{aligned}$$

The first inequality follows from the definitions of the infinite norm for vectors and the Frobenius norm for the element of  $\mathcal{V}$ . The second inequality follows from Lemma 2.16 in [11], the third inequality follows from the triangle inequality for the Frobenius norm, and the last inequality follows from Lemma 2.12 in [11]. This completes the proof of the lemma.  $\square$

Substitution of the two inequalities in Lemma 10 into the inequality in Lemma 9 yields the following upper bound

$$4\delta(v^f)^2 \leq \frac{\left\| \frac{d_x^f \circ d_s^f}{1-\theta} \right\|_F^2}{1 - \left\| \frac{\lambda(d_x^f \circ d_s^f)}{1-\theta} \right\|_\infty} \leq \frac{\frac{1}{4} \left( \frac{\|d_x^f\|_F^2 + \|d_s^f\|_F^2}{1-\theta} \right)^2}{1 - \frac{1}{2} \frac{\|d_x^f\|_F^2 + \|d_s^f\|_F^2}{1-\theta}}. \tag{42}$$

Thus, a task of finding an upper bound of  $\delta(v^f)$  reduces to finding an upper bound of  $\|d_x^f\|_F^2 + \|d_s^f\|_F^2$ .

After careful and somewhat involved analysis, details of which are omitted and can be found in [11], the following upper bound is derived:

$$\|d_x^f\|_F^2 + \|d_s^f\|_F^2 \leq 2(4(1-\theta)^2\delta^2 + \theta^2r) + 12\theta^2r^2\rho(\delta)^2, \tag{43}$$

where  $\delta := \delta(v)$  and  $\rho(\delta) := \delta + \sqrt{1 + \delta^2}$ . Thus, the upper bound essentially depends on the barrier parameter  $\theta$  and the proximity measure  $\delta$  of the old iterate  $(x, y, s)$ , which is a desired result since we want to connect new proximity measure with the old one.

In what follows, we want to choose  $\theta$ ,  $0 < \theta < 1$ , as large as possible, and such that  $(x^f, y^f, s^f)$  lies in the quadratic convergence neighborhood with respect to the  $\mu^+$ -center of the perturbed problems  $(SOP_{\nu^+})$  and  $(SOD_{\nu^+})$ . As it was shown in the feasible case, this neighborhood can be extended to  $\delta(v^f) \leq 1/\sqrt[4]{2}$  as opposed to  $\delta(v^f) \leq 1/\sqrt{2}$  in [11].

From (42), we know that  $\delta(v^f) \leq 1/\sqrt[4]{2}$  holds if

$$\frac{\frac{1}{4} \left( \frac{\|d_x^f\|_F^2 + \|d_s^f\|_F^2}{1-\theta} \right)^2}{1 - \frac{1}{2} \frac{\|d_x^f\|_F^2 + \|d_s^f\|_F^2}{1-\theta}} \leq 2\sqrt{2}, \tag{44}$$

which leads to the following inequality

$$\frac{\|d_x^f\|_F^2 + \|d_s^f\|_F^2}{1-\theta} \leq 2\sqrt{2} \left( \sqrt{1 + \sqrt{2}} - 1 \right) \approx 1.566. \tag{45}$$

Substituting (43) into the above inequality (45) we obtain

$$2(4(1-\theta)^2\delta^2 + \theta^2r) + 12\theta^2r^2\rho(\delta)^2 \leq 2\sqrt{2} \left( \sqrt{1 + \sqrt{2}} - 1 \right) (1-\theta). \tag{46}$$

One can easily verify that the largest values of  $\theta$  and  $\tau$  for which inequality (45) holds are

$$\tau = \frac{1}{16} \text{ and } \theta = \frac{1}{4r}. \tag{47}$$

Furthermore, from Lemma 10 and (45) we obtain

$$\|\lambda(d_x^f \circ d_s^f)\|_\infty \leq \frac{1}{2}(\|d_x^f\|_F^2 + \|d_s^f\|_F^2) \leq \sqrt{2} \left(\sqrt{1 + \sqrt{2}} - 1\right) (1 - \theta) < 1 - \theta. \tag{48}$$

Lemma 8 then implies that with the above choice of parameters  $\theta$  and  $\tau$ ,  $(x^f, y^f, s^f)$  is indeed strictly feasible.

**Centering Steps.** After the feasibility step we perform centering steps in order to get an iterate  $(x^+, y^+, s^+)$  that is in the  $\tau$ -neighborhood of the  $\mu^+$ -center, i.e. satisfies  $\text{tr}(x^+ \circ s^+) = r\mu^+$  and  $\delta(x^+, s^+; \mu^+) \leq \tau$ . Using Corollary 3, the required number of centering steps can easily be obtained. Indeed, since  $(x^f, y^f, s^f)$  is in the quadratic convergence neighborhood of the  $\mu^+$ -center, i.e.  $\delta = \delta(x^f, s^f; \mu^+) \leq 1/\sqrt[4]{2}$ , after  $k$  centering steps we will have an iterate  $(x^+, y^+, s^+)$  that is still feasible for  $(SOP_{\nu^+})$  and  $(SOD_{\nu^+})$  and satisfies

$$\delta(x^+, s^+; \mu^+) \leq \left(\frac{1}{\sqrt[4]{2}}\right)^{2^k}.$$

Hence,  $\delta(x^+, y^+, s^+) \leq \tau$  is satisfied if  $k$  satisfies

$$\left(\frac{1}{\sqrt[4]{2}}\right)^{2^k} \leq \tau.$$

Thus,  $\delta(x^+, s^+; \mu^+) \leq \tau$  will be obtained after at most

$$2 + \left\lceil \log_2 \left( \log_2 \frac{1}{\tau} \right) \right\rceil \tag{49}$$

centering steps.

Substituting  $\tau = 1/16$  into the above expression leads to

$$2 + \left\lceil \log_2 \left( \log_2 \left( \frac{1}{\tau} \right) \right) \right\rceil = 2 + \lceil \log_2 (\log_2 16) \rceil = 4.$$

Hence, at most four centering steps are needed to get the iterate  $(x^+, y^+, s^+)$  that satisfies  $\delta(x^+, s^+; \mu^+) \leq \tau$ , i.e., the iterate that is in the  $\tau$ -neighborhood of the  $\mu^+$ -center again.

**Iteration Bound.** To summarize, each main iteration consists of at most five inner iterations, one feasibility step, and at most four centering steps. In each main iteration both the duality gap and the norms of the residual vectors are reduced by the factor  $(1 - \theta)$ . Hence, using  $\text{tr}(x^0 \circ s^0) = r\zeta^2$ , the total number of main iterations is bounded above by

$$\frac{1}{\theta} \log \frac{\max\{r\zeta^2, \|r_p^0\|_F, \|r_d^0\|_F\}}{\varepsilon}$$

Since  $\theta = \frac{1}{4r}$  and at most five inner iterations per the main iteration are needed, the main result can be stated in the following theorem.

*Theorem 6*

Suppose  $(SOP)$  has an optimal solution  $x^*$  and  $(SOD)$  has an optimal solution  $(y^*, s^*)$ , which satisfy  $\text{tr}(x^* \circ s^*) = 0$  and  $x^* + s^* \preceq_{\mathcal{K}} \zeta e$  for some  $\zeta > 0$ . If the values of parameters  $\tau$  and  $\theta$  are chosen as  $\tau = \frac{1}{16}$  and  $\theta = \frac{1}{4r}$ , then at most

$$20r \log \frac{\max\{r\zeta^2, \|r_p^0\|_F, \|r_d^0\|_F\}}{\varepsilon}$$

inner iterations of the algorithm in Fig. 1, are needed to find an  $\varepsilon$ -approximate optimal solution of  $(SOP)$  and  $(SOD)$ .



In conclusion, the infeasible algorithm in Fig 1 is well defined, globally convergent, and achieves quadratic convergence of full NT-feasibility steps in the wider neighborhood of the central path while still maintaining the best-known iteration bound known for these types of methods, namely

$$O\left(r \log \frac{\max\{r\zeta^2, \|r_p^0\|_F, \|r_d^0\|_F\}}{\varepsilon}\right).$$

*Remark 1*

Similarly to LO, the iteration bound in Theorem 6 is derived under the assumption that there exists an optimal solution pair  $(x^*, y^*, s^*)$  of (SOP) and (SOD) with vanishing duality gap and satisfying  $x^* + s^* \preceq_{\mathcal{K}} \zeta e$ . During the course of the algorithm, if at some main iteration, the proximity measure  $\delta$  after the feasibility step exceeds  $1/\sqrt[4]{2}$ , then it tells us that the above assumption does not hold. It may happen that the value of  $\zeta$  has been chosen too small. In this case, one might run the algorithm once more with a larger value of  $\zeta$ . If this does not help, then eventually one should realize that (SOP) and/or (SOD) do not have optimal solutions at all, or they have optimal solutions with a positive duality gap.

*Remark 2*

In [11] the number of centering steps per the main iteration is three. In our paper, the ‘price to pay for expanding the quadratic convergence neighborhood of the central path is a possible additional centering step which slightly increases the constant in the upper bound on the total number of inner iterations from 16 to 20; however, that does not change the order of magnitude of the required number of iterations, it still matches the best-known iteration bound for the infeasible algorithms mentioned above. It is also worth mentioning that in practice all four centering steps may not always be needed, very often only one or two suffice.

## 5. Concluding remarks

In this paper, an infeasible version of the full NT-step IPM for SO in a wider neighborhood of the central path than the one in [11] is presented and convergence analysis is given. Wider quadratic convergence neighborhood of the central path characterized by  $1/\sqrt[4]{2}$  is carried over from the feasible case and applied to the feasibility steps of the infeasible algorithm resulting in larger steps. However, despite full NT-steps in the wider neighborhood of the central path, the best complexity known for the infeasible algorithm is still maintained.

Future research is planned in two directions. The first direction is implementation and numerical testing of the method on a set of conic CTA problems as well as other conic problems. The second direction is theoretical and involves the generalization of this IPM to other optimization problems such as Linear Complementarity Problems over symmetric cones.

## Acknowledgments

Any opinions, findings, and conclusions, or recommendations expressed in this publication are those of the authors only and do not necessarily reflect the views of the Centers for Disease Control and Prevention.

## REFERENCES

- [1] Anjos M.F., Lasserre, J.B.: Handbook on Semidefinite, Conic and Polynomial Optimization: Theory, Algorithms, Software and Applications. International Series in Operational Research and Management Science. Volume 166, Springer, New York, USA (2012)
- [2] Castro, J.: Minimum-distance controlled perturbation methods for large-scale tabular data protection. European J. Oper. Res., 171, 39-52 (2006)
- [3] Castro, J., Gonzalez, J.A.: Assessing the information loss of controlled adjustment methods in two-way tables. *Privacy in Statistical Databases 2014*, LNCS, 8744, 79-88 (2014)

- [4] Castro, J., Gonzalez, J.A.: A fast CTA method without complicating binary decisions. Documents of the Joint UNECE / Eurostat Work Session on Statistical Data Confidentiality, Statistics Canada, Ottawa, 1-7 (2013)
- [5] Castro, J., Gonzalez, J.A.: A multiobjective LP approach for controlled tabular adjustment in statistical disclosure control. Working paper, Department of Statistics and Operations Research, Universitat Politecnica de Catalunya (2014)
- [6] Dandekar, R. A., Cox, L.H.: Synthetic tabular Data: an alternative to complementary cell suppression. Manuscript, Energy Information Administration, U.S. (2002)
- [7] Faraut, J., Koranyi, A.: Analysis on Symmetric Cones, Oxford University Press, New York, USA (1994)
- [8] Faybusovich L.: Euclidean Jordan algebras and interior-point algorithms. Positivity 1, 331-35 (1997)
- [9] Faybusovich L.: A Jordan-algebraic approach to potential-reduction algorithms. Math. Z. 239(1), 117-129 (2002)
- [10] Gu, G., Mansouri, H., Zangiabadi, M., Bai, Y.Q., Roos, C.: Improved full-Newton step  $O(nL)$  infeasible interior-point method for linear optimization. J. Optim. Theory Appl. 145(2), 271-288 (2010)
- [11] Gu, G., Zangiabadi, M., Roos, C.: Full Nesterov-Todd step infeasible interior-point method for symmetric optimization. European J. Oper. Res. 214(3), 473-484 (2011)
- [12] Güler, O.: Barrier functions in interior-point methods. Math. Oper. Res. 21(4), 860-885 (1996)
- [13] Lesaja, G., Castro, J., Oganian, A.: Second Order Cone formulation of Continuous CTA Model. Lecture Notes in Computer Science 9867, Springer, 41-53, (2016)
- [14] Liu, C.H., Liu, H.W., Liu, X.Z.: Polynomial convergence of second-order mehrotra type predictor-corrector algorithms over symmetric cones. J. Optim. Theory Appl. 154(3), 949-965 (2012)
- [15] Liu H.W., Yang X.M., Liu C.H.: A new wide neighborhood primal-dual infeasible interior-point method for symmetric cone programming[J]. J. Optim. Theory Appl. 158(3), 796-815 (2013)
- [16] Muramatsu, M.: On a commutative class of search directions for linear programming over symmetric cones. J. Optim. Theory Appl. 112(3), 595-625 (2002)
- [17] Nesterov Y.E., Todd M.J.: Self-scaled barriers and interior-point methods for convex programming. Math. Oper. Res. 22(1), 1-42 (1997)
- [18] Nesterov Y.E., Todd M.J.: Primal-dual interior-point methods for self-scaled cones. SIAM J. Optim. 8(2), 324-364 (1998)
- [19] Rangarajan, B.K.: Polynomial convergence of infeasible-interior-point methods over symmetric cones. SIAM J. Optim. 16(4), 1211-1229 (2006)
- [20] Roos, C.: A full-Newton step  $O(n)$  infeasible interior-point algorithm for linear optimization. SIAM J. Optim. 16(4), 1110-1136 (2006)
- [21] Roos C., Terlaky T., Vial J.-Ph.: Theory and Algorithms for Linear Optimization: An Interior-Point Approach. John Wiley & Sons, Chichester, UK (1997)
- [22] Schmieta, S.H., Alizadeh, F.: Extension of primal-dual interior-point algorithms to symmetric cones. Math. Program. 96(3), 409-438 (2003)
- [23] Vieira, M.V.C.: Jordan algebraic approach to symmetric optimization. Ph.D thesis, Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, The Netherlands (2007)
- [24] Wang, G.Q., Bai, Y.Q.: A new full Nesterov-Todd step primal-dual path-following interior-point algorithm for symmetric optimization. J. Optim. Theory Appl. 154(3), 966-985 (2012)
- [25] Wang, G.Q., Kong, L.C., Tao, J.Y., Lesaja G.: Improved complexity analysis of full Nesterov-Todd step feasible interior-point method for symmetric optimization. J. Optim. Theory Appl. 166(2), 588-604, (2015)
- [26] Wright, S.J.: *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia (1996)