



# A regularized interior point method for sparse optimal transport on graphs

S. Cipolla\*, J. Gondzio, F. Zanetti

The University of Edinburgh, James Clerk Maxwell Building, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, United Kingdom

## ARTICLE INFO

### Keywords:

Convex programming  
Primal–dual regularized interior point methods  
Optimal transport on graphs  
Polynomial complexity  
Inexact interior point methods

## ABSTRACT

In this work, the authors address the Optimal Transport (OT) problem on graphs using a proximal stabilized Interior Point Method (IPM). In particular, strongly leveraging on the induced primal–dual regularization, the authors propose to solve large scale OT problems on sparse graphs using a bespoke IPM algorithm able to suitably exploit primal–dual regularization in order to enforce scalability. Indeed, the authors prove that the introduction of the regularization allows to use sparsified versions of the normal Newton equations to inexpensively generate IPM search directions. A detailed theoretical analysis is carried out showing the polynomial convergence of the inner algorithm in the proposed computational framework. Moreover, the presented numerical results showcase the efficiency and robustness of the proposed approach when compared to network simplex solvers.

## 1. Introduction

The Optimal Transport (OT) problem requires to move a certain distribution of mass from one configuration into another, minimizing the total cost required for the operation. It has been studied extensively, from the early work of Kantorovich (Kantorovich, 1958), to the development of ever faster algorithms for various OT formulations, e.g. Cuturi (2013), Gottschlich and Schuhmacher (2014), Ling and Okada (2007), Merigot (2011), Peyre and Cuturi (2019), Schrieber, Schuhmacher, and Gottschlich (2017). Recently, there has been a growing interest in using Interior Point Methods (IPMs) (Gondzio, 2012a) in applications that involve optimal transport, in particular for very large scale instances of such problems, see e.g. Natale and Todeschi (2021), Wijesinghe and Chen (2022), Zanetti and Gondzio (2023).

A particularly interesting problem is the optimal transport over sparse graphs: in this case, the transport of mass is only possible along a specific subset of connections, which is noticeably smaller than the full list of edges of a fully connected bipartite graph, as it would happen in a standard discrete OT formulation. The use of OT and the Wasserstein distance (i.e. the optimal objective function of the OT problem) is becoming more and more common in many practical applications, e.g. neural networks (Gulrajani, Ahmed, Arjovsky, Dumoulin, & Courville, 2017), image processing (Haker, Zhu, Tannenbaum, & Angenent, 2004), inverse problems (Métivier, Brossier, Méridot, Oudet, & Virieux, 2016) and in the analysis of large complex networks (Ni, Lin, Luo, & Gao, 2019).

The specific formulation of the problem is the following: suppose that  $G = (V, E)$  is a connected graph with directed edges  $E \subset V \times V$

and weights  $\mathbf{c} \in \mathbb{R}_+^{|E|}$ . Define the incidence matrix  $A \in \{-1, 0, 1\}^{|V| \times |E|}$  as

$$A_{ve} := \begin{cases} -1, & \text{if } e = (v, w) \text{ for some } w \in V \\ 1, & \text{if } e = (w, v) \text{ for some } w \in V \\ 0, & \text{otherwise.} \end{cases}$$

We consider the optimal transport problem in the Beckmann form (Essid & Solomon, 2018):

$$\mathcal{W}_1(\rho_0, \rho_1) := \begin{cases} \min_{\mathbf{x} \in \mathbb{R}^{|E|}} & \sum_{e \in E} c_e x_e \\ \text{s.t.} & A\mathbf{x} = \rho_1 - \rho_0 \\ & \mathbf{x} \geq 0, \end{cases} \quad (1)$$

where  $\rho_0, \rho_1 \in \{\rho \in \mathbb{R}^{|V|} : \mathbf{1}^T \rho = 1 \text{ and } \rho \geq 0\} =: \text{Prob}(V)$ . In the following we will define  $|E| := n$  and  $|V| := m$ .

**Assumption 1.** In the remainder of this work we will assume that the graph  $G = (V, E)$  is sparse. In particular, we will assume that there exists a constant  $C_{\text{deg}} \in \mathbb{N}$  independent of  $n$  such that  $\max_{v \in V} \text{deg}(v) \leq C_{\text{deg}}$ .

Moreover, we assume that the graph is not a trivial one and that it has more edges than nodes, i.e.  $|E| \geq |V|$ .

Most practical sparse graphs satisfy this assumption and in particular the graphs used in the numerical experiments of the paper do. Notice however that some particular sparse graphs may not satisfy this assumption, e.g., if there exists a node in the graph having connections to most or all of the other nodes.

\* Corresponding author.

E-mail address: [s.cipolla@soton.ac.uk](mailto:s.cipolla@soton.ac.uk) (S. Cipolla).

OT on graphs has been recently studied in [Essid and Solomon \(2018\)](#), [Facca and Benzi \(2021\)](#) and, in this formulation, it is similar to the more general minimum cost flow problem on networks ([Ahuja, Magnanti, & Orlin, 1993](#)), which has also seen extensive use of IPMs, e.g. [Castro and Nasini \(2021\)](#), [Frangioni and Gentile \(2004\)](#), [Portugal, Resende, Veiga, and Júdice \(2000\)](#), [Resende and Veiga \(1993\)](#).

Sparse graphs have on average very few edges per node, which can lead to nearly disconnected regions and seriously limit the possible paths where mass can be moved. As a result, finding a solution to the optimal transport problem on a sparse graph requires more sophisticated algorithms and may be more computationally challenging compared to solving the same problem on a denser graph. In particular, first order methods like the network simplex may struggle and move slowly towards optimality, due to the limited number of edges available, while an interior point method manages to identify quickly the subset of basic variables (i.e. the subset of edges with non-zero flow) and converges faster.

In this work, the authors address the efficient solution of the optimal transport problem (1) considering the Proximal-Stabilized Interior Point framework (PS-IPM), recently introduced and analysed in [Cipolla and Gondzio \(2023\)](#).

As originally observed in [Portugal et al. \(2000\)](#), when IPMs are used to solve the minimum cost flow problem on networks, the normal form of the related Newton systems is structured as a Laplacian matrix of the graph (defined as the difference of the diagonal matrix of the vertex degrees minus the adjacency matrix) and the iterates of IPM determine the associate weights of this matrix, see also Eq. (38). In [Daitch and Spielman \(2008\)](#), this observation was exploited to solve such Laplacian linear systems (which are, in turn, particular instances of symmetric M-matrices) through the fast specialized solution of  $O(\ln m)$  linear systems involving symmetric diagonally dominant matrices ([Spielman & Teng, 2014](#)). We refer the interested reader to [Vishnoi \(2012\)](#) for a survey on fast Laplacian solvers and to [Frangioni and Serra Capizzano \(2001\)](#) for information concerning the distribution of Laplacian’s singular values.

### 1.1. Contribution and organization

This work focuses on the efficient solution of large scale OT problems on sparse graphs (see [Assumption 1](#)) using a bespoke IPM algorithm able to suitably exploit primal–dual regularization in order to enforce scalability. The organization of the work and its main contributions can be summarized as follows:

- In [Section 2](#), the authors briefly recall the proximal stabilized framework responsible for the primal–dual regularization of the IPMs here considered.
- In [Section 3](#), the authors provide a detailed convergence analysis of the inexact infeasible primal–dual regularized IPM, when a proximal stabilization procedure is used. Moreover, they prove its polynomial complexity.
- In [Section 4](#), the authors prove that the normal form of the related Newton system is naturally structured as a shifted Laplacian matrix characterized by a strict diagonal dominance. Such feature consistently simplifies the factorization of the normal equations and allows the use of standard libraries for the solution of the corresponding linear systems. On the other hand, such factorizations could incur a significant fill-in even when the original graph is sparse, hence limiting the applicability of the proposed approach for the solution of large scale problems.
- In [Section 5](#), to overcome potential scalability issues related to the fill-in mentioned above, the authors propose to generate IPM search directions using *sparsified* versions of the IPM normal equations. In particular, when classical primal–dual regularized IPMs are considered, the normal matrix takes the form  $A(\Theta^{-1} + \rho I)^{-1} A^T + \delta I$ , where  $\rho, \delta$  are regularization parameters and  $\Theta$  is a diagonal matrix related to the IPM barrier parameter  $\mu$ ; the

authors propose to use a *perturbed* normal matrix, where the entries of  $(\Theta^{-1} + \rho I)^{-1}$  that are sufficiently small (when compared to  $\mu$ ) are set to zero (completely ignoring the corresponding columns of matrix  $A$ ). This strategy reduces the time required to assemble and solve the normal equations systems, providing a fundamental advantage to the algorithm.

The resulting sparsified linear systems are solved either using a Cholesky factorization (if that displays only negligible fill-in) or using the conjugate gradient method and employing a simple and inexpensive incomplete Cholesky preconditioner. In both these cases either the *complete* or the *incomplete* Cholesky factorization remains very sparse, and this translates into an outstanding efficiency of the proposed method. Moreover, the authors are able to interpret the Newton directions generated using sparsified Newton matrices as *inexact* Newton directions. Relying on the convergence theory developed in [Section 3](#), the authors are able to prove that, under suitable choice of the sparsification parameters, the above described approach gives rise to a polynomially convergent algorithm.

- In [Section 6](#), the authors present experimental results which demonstrate the efficiency and robustness of the proposed approach. Extensive numerical experiments, involving very large and sparse graphs coming from public domain random generators as well as from real world applications, show that, for sufficiently large problems, the approach presented in this work consistently outperforms, in terms of computational time, the Lemon network simplex implementation ([Kovacs, 2015](#)), one of the state-of-the-art solvers available for network problems.

### 1.2. Notation

In the paper, vectors are indicated with bold letters.  $\|\cdot\|$  indicates the Euclidean norm.  $I$  represents the identity matrix and  $\mathbf{e}$  the vector of all ones. Given a vertex  $v$  of a graph  $G$ , we denote as  $deg(v)$  its degree, i.e. the number of edges that are incident to  $v$ . Concerning the variables inside the algorithm, we use a subscript  $k$  to indicate the external proximal iteration and a superscript  $j$  to indicate the internal IPM iteration. Given a sequence  $\{\mu^j\}_{j \in \mathbb{N}}$  and a continuous function  $f$ , the big-O notation  $O(\cdot)$  is used as follows:

$$\{\mu^j\}_{j \in \mathbb{N}} \in O(f(\mu^j)) \text{ iff } \exists C > 0 \text{ s.t. } \mu^j \leq C f(\mu^j) \text{ for all } j \in \mathbb{N}.$$

## 2. Computational framework

### 2.1. Proximal-stabilized interior point method

Let us consider the following primal–dual formulation of a Linear Program (LP):

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{c}^T \mathbf{x} & \qquad \max_{\mathbf{s} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m} \mathbf{b}^T \mathbf{y} \\ \text{s.t. } \mathbf{A} \mathbf{x} = \mathbf{b} & \qquad \text{s.t. } \mathbf{c} - \mathbf{A}^T \mathbf{y} - \mathbf{s} = 0 \\ \mathbf{x} \geq 0 & \qquad \mathbf{s} \geq 0 \end{aligned} \tag{2}$$

where  $A \in \mathbb{R}^{m \times n}$  with  $m \leq n$  is not required to have full rank. Notice that problem (1), under [Assumption 1](#), is indeed formulated in this way.

We solve this problem using PS-IPM ([Cipolla & Gondzio, 2023](#)), which is a *proximal-stabilized* version of classic *Interior Point Method*. Broadly speaking, PS-IPM resorts to the Proximal Point Method (PPM) ([Rockafellar, 1970](#)) to produce primal–dual regularized forms of problem (2). Indeed, given an approximation  $(\mathbf{x}_k, \mathbf{y}_k)$  of the solution of such problem, PS-IPM uses interior point methods to produce the next PPM step  $(\mathbf{x}_{k+1}, \mathbf{y}_{k+1})$ , which, in turn, represents a *better* approximation of the solution of problem (2).

In this regard, the problem that needs to be solved at every PPM step takes the form

$$\begin{aligned} \min_{\substack{\mathbf{x} \in \mathbb{R}^n \\ \mathbf{y} \in \mathbb{R}^m}} \mathbf{c}^T \mathbf{x} + \frac{\rho}{2} \|\mathbf{x} - \mathbf{x}_k\|^2 + \frac{\delta}{2} \|\mathbf{y}\|^2 & \quad \max_{\substack{\mathbf{x} \in \mathbb{R}^n \\ \mathbf{y} \in \mathbb{R}^m}} \mathbf{y}^T \mathbf{b} - \frac{\rho}{2} \|\mathbf{x}\|^2 - \frac{\delta}{2} \|\mathbf{y} - \mathbf{y}_k\|^2 \\ \text{s.t. } \mathbf{A}\mathbf{x} + \delta(\mathbf{y} - \mathbf{y}_k) = \mathbf{b} & \quad \text{s.t. } \rho(\mathbf{x} - \mathbf{x}_k) - \mathbf{A}^T \mathbf{y} - \mathbf{s} + \mathbf{c} = \mathbf{0} \\ \mathbf{x} \geq 0, & \quad \mathbf{s} \geq 0 \end{aligned} \tag{PPM(k)}$$

**Definition 1.** Solution of problem (PPM(k)) Using standard duality theory, we say that  $(\mathbf{x}_k^*, \mathbf{y}_k^*, \mathbf{s}_k^*)$  is a solution of problem (PPM(k)) if the following identities hold

$$\begin{aligned} \mathbf{A}\mathbf{x}_k^* + \delta(\mathbf{y}_k^* - \mathbf{y}_k) - \mathbf{b} &= 0 \\ \rho(\mathbf{x}_k^* - \mathbf{x}_k) - \mathbf{A}^T \mathbf{y}_k^* - \mathbf{s}_k^* + \mathbf{c} &= 0 \\ (\mathbf{x}_k^*)^T \mathbf{s}_k^* = 0 \text{ and } (\mathbf{x}_k^*, \mathbf{s}_k^*) &\geq 0 \end{aligned} \tag{3}$$

More in particular, the PS-IPM here considered uses two nested cycles to solve problem (2). The outer loop uses an inexact proximal point method (Luque, 1984), as shown in Algorithm 1: the current approximate solution  $(\mathbf{x}_k, \mathbf{y}_k)$  is used to regularize the LP problem, which is then solved using an IPM to find the next approximate solution  $(\mathbf{x}_{k+1}, \mathbf{y}_{k+1}) \approx (\mathbf{x}_k^*, \mathbf{y}_k^*)$ . And indeed, at the inner loop level, an inexact infeasible interior point method is used to solve the PPM sub-problems, see Algorithm 2.

Notice that both methods are *inexact*: the outer cycle is inexact because the sub-problems are solved approximately by an IPM; the IPM is inexact because the Newton systems are also solved inexactly (see Section 2.2 for more details). Notice also that the IPM is referred to as *infeasible* because the intermediate iterates are not required to be inside the feasible region. We also call the inner loop *regularized*, because it is a primal–dual regularized version of the original LP (2).

Regularization in interior point methods was originally introduced in Saunders and Tomlin (1996) and extensively used in Altman and Gondzio (1999), as a tool to stabilize and improve the linear algebra routines needed for their efficient implementation. In this work and in Cipolla and Gondzio (2023), the regularization is introduced as a result of the application of the PPM at the outer cycle level. To summarize, in the following we use three acronyms: PPM refers to the outer cycle; IPM refers to the inner cycle; PS-IPM refers to the overall procedure, combining PPM and IPM.

**Input:**  $tol > 0, \sigma_r \in (0, 1), \tau_1 > 0$ .  
**Initialization:** Iteration counter  $k = 0$ ; initial point  $(\mathbf{x}_0, \mathbf{y}_0)$   
**1 while Stopping Criterion (43) False do**  
**2** Use Algorithm 2 with starting point  $(\mathbf{x}_k^0, \mathbf{y}_k^0) = (\mathbf{x}_k, \mathbf{y}_k)$  to find  $(\mathbf{x}_{k+1}, \mathbf{y}_{k+1})$  s.t.  

$$\|\mathbf{r}_k(\mathbf{x}_{k+1}, \mathbf{y}_{k+1})\| < \frac{(\sigma_r)^k}{\tau_1} \min\{1, \|(\mathbf{x}_{k+1}, \mathbf{y}_{k+1}) - (\mathbf{x}_k, \mathbf{y}_k)\|\} \tag{4}$$
  
**3** Update the iteration counter:  $k := k + 1$ .  
**4 end**

**Algorithm 1:** PPM, outer loop of PS-IPM

Concerning the stopping criteria, we finally highlight that Algorithm 1 is stopped based on the criterion (43). Algorithm 2 instead, is stopped according to the accuracy that is required for the solution of current sub-problem and based on the following *natural residual*, see Luque (1984), of problem (PPM(k)):

**Definition 2 (Natural Residual).**

$$\mathbf{r}_k(\mathbf{x}, \mathbf{y}) := \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} - \Pi_D \left( \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} - \begin{bmatrix} \rho(\mathbf{x} - \mathbf{x}_k) + \mathbf{c} - \mathbf{A}^T \mathbf{y} \\ \mathbf{A}\mathbf{x} - \mathbf{b} + \delta(\mathbf{y} - \mathbf{y}_k) \end{bmatrix} \right),$$

where

$$D := \mathbb{R}_{\geq 0}^n \times \mathbb{R}^m$$

and where  $\Pi_D$  is the corresponding projection operator. Moreover, it is easy to verify that  $(\mathbf{x}_k^*, \mathbf{y}_k^*, \mathbf{s}_k^*)$  is a solution of problem (PPM(k)) if and only if  $\mathbf{r}_k(\mathbf{x}_k^*, \mathbf{y}_k^*) = \mathbf{0}$ , Cipolla and Gondzio (2023, Sec. 2.3).

**2.2. Interior point method**

We now focus on the inner cycle and give a brief description of the IPM used to solve problem (PPM(k)). To this aim, we introduce the following Lagrangian function which uses a logarithmic barrier to take into account the inequality constraints

$$\begin{aligned} L_k(\mathbf{x}, \mathbf{y}) &= \frac{1}{2} [\mathbf{x}^T, \mathbf{y}^T] \begin{bmatrix} \rho I & 0 \\ 0 & \delta I \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} + [\mathbf{c}^T - \rho \mathbf{x}_k^T, 0] \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \\ &\quad - \mathbf{y}^T (\mathbf{A}\mathbf{x} + \delta(\mathbf{y} - \mathbf{y}_k) - \mathbf{b}) - \mu \sum_{i=1}^n \ln(x_i). \end{aligned} \tag{5}$$

The KKT conditions that arise from the gradients of the Lagrangian (5) are

$$\begin{aligned} \nabla_{\mathbf{x}} L_k(\mathbf{x}, \mathbf{y}) &= \rho \mathbf{x} - \mathbf{A}^T \mathbf{y} + \mathbf{c} - \rho \mathbf{x}_k - \begin{bmatrix} \frac{\mu}{x_1} \\ \vdots \\ \frac{\mu}{x_n} \end{bmatrix} = 0; \\ -\nabla_{\mathbf{y}} L_k(\mathbf{x}, \mathbf{y}) &= (\mathbf{A}\mathbf{x} + \delta(\mathbf{y} - \mathbf{y}_k) - \mathbf{b}) = 0. \end{aligned}$$

Setting  $s_i = \frac{\mu}{x_i}$  for  $i \in \{1, \dots, n\}$ , we consider the following function

$$F_k^{\mu, \sigma}(\mathbf{x}, \mathbf{y}, \mathbf{s}) := \begin{bmatrix} \rho(\mathbf{x} - \mathbf{x}_k) - \mathbf{A}^T \mathbf{y} - \mathbf{s} + \mathbf{c} \\ \mathbf{A}\mathbf{x} + \delta(\mathbf{y} - \mathbf{y}_k) - \mathbf{b} \\ S\mathbf{X}\mathbf{e} - \sigma \boldsymbol{\mu} \end{bmatrix}, \tag{6}$$

where  $\sigma \in (0, 1)$  is the barrier reduction parameter,  $S = \text{diag}(\mathbf{s})$  and  $X = \text{diag}(\mathbf{x})$ . A primal–dual interior point method applied to problem (PPM(k)) relies on the use of Newton iterations to solve a nonlinear problem of the form

$$F_k^{\mu, \sigma}(\mathbf{x}, \mathbf{y}, \mathbf{s}) = \mathbf{0}, \quad \mathbf{x}, \mathbf{s} > \mathbf{0}.$$

A Newton step for (6) from the current iterate  $(\mathbf{x}, \mathbf{y}, \mathbf{s})$  is obtained by solving the system

$$\begin{bmatrix} \rho I & -\mathbf{A}^T & -I \\ \mathbf{A} & \delta I & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x} \\ \Delta \mathbf{y} \\ \Delta \mathbf{s} \end{bmatrix} = -F_k^{\mu, \sigma}(\mathbf{x}, \mathbf{y}, \mathbf{s}) =: \begin{bmatrix} \xi_d \\ \xi_p \\ \xi_{\mu, \sigma} \end{bmatrix}, \tag{7}$$

i.e., the following relations hold:

$$\rho \Delta \mathbf{x} - \mathbf{A}^T \Delta \mathbf{y} - \Delta \mathbf{s} = \xi_d \tag{8}$$

$$\mathbf{A} \Delta \mathbf{x} + \delta \Delta \mathbf{y} = \xi_p \tag{9}$$

$$S \Delta \mathbf{x} + X \Delta \mathbf{s} = \xi_{\mu, \sigma}, \tag{10}$$

where  $(\Delta \mathbf{x}, \Delta \mathbf{y}, \Delta \mathbf{s})$  is the Newton direction to be taken at each iteration (with an appropriate stepsize).

The solution of (7) is delivered by the following computational procedure

$$(\mathbf{A}(\Theta^{-1} + \rho I)^{-1} \mathbf{A}^T + \delta I) \Delta \mathbf{y} = \xi_p - \mathbf{A}(\Theta^{-1} + \rho I)^{-1} (X^{-1} \xi_{\mu, \sigma} + \xi_d) \tag{11}$$

$$(\Theta^{-1} + \rho I) \Delta \mathbf{x} = \mathbf{A}^T \Delta \mathbf{y} + \xi_d + X^{-1} \xi_{\mu, \sigma} \tag{12}$$

$$X \Delta \mathbf{s} = (\xi_{\mu, \sigma} - S \Delta \mathbf{x}). \tag{13}$$

where  $\Theta := X S^{-1}$ . Before continuing let us give basic definitions used in the remainder of this work.

**Definition 3.** Normal Matrix:

$$S_{\rho, \delta} := \mathbf{A}(\Theta^{-1} + \rho I)^{-1} \mathbf{A}^T + \delta I. \tag{14}$$

Neighbourhood of the infeasible central path:

$$\begin{aligned} \mathcal{N}_k(\bar{\gamma}, \underline{\gamma}_p, \underline{\gamma}_d) &:= \{(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \mathbb{R}_{>0}^n \times \mathbb{R}^m \times \mathbb{R}_{>0}^n : \\ &\quad \bar{\gamma} \mathbf{x}^T \mathbf{s} / n \geq x_i s_i \geq \underline{\gamma} \mathbf{x}^T \mathbf{s} / n \text{ for } i = 1, \dots, n; \\ &\quad \mathbf{x}^T \mathbf{s} \geq \underline{\gamma}_p \|\mathbf{A}\mathbf{x} + \delta(\mathbf{y} - \mathbf{y}_k) - \mathbf{b}\|; \\ &\quad \mathbf{x}^T \mathbf{s} \geq \underline{\gamma}_d \|\rho(\mathbf{x} - \mathbf{x}_k) - \mathbf{A}^T \mathbf{y} - \mathbf{s} + \mathbf{c}\| \}, \end{aligned} \tag{15}$$

where  $\bar{\gamma} > 1 > \underline{\gamma} > 0$  and  $(\gamma_p, \gamma_d) > 0$ .

The neighbourhood here considered is standard in the analysis of infeasible IPMs (Wright, 1997): it requires the iterates to be close enough to the central path (according to parameters  $\bar{\gamma}$  and  $\underline{\gamma}$ ), and the primal–dual constraint violations to be reduced at the same rate as the complementarity product  $\mathbf{x}^T \mathbf{s}$ . Within this neighbourhood,  $\mathbf{x}^T \mathbf{s} \rightarrow 0$  guarantees convergence to a primal–dual optimal solution.

Moreover, we consider an *inexact* solution of the linear system (11):

**Assumption 2.**

$$S_{\rho, \delta} \Delta \mathbf{y} = \bar{\xi}_p + \zeta \text{ where } \|\zeta\| \leq C_{\text{inexact}} \mathbf{x}^T \mathbf{s}, \tag{16}$$

where  $C_{\text{inexact}} \in (0, 1)$  and we defined

$$\bar{\xi}_p := \xi_p - A(\Theta^{-1} + \rho I)^{-1}(X^{-1}\xi_{\mu, \sigma} + \xi_d).$$

It is important to note that the above Assumption 2 is a non-standard requirement in inexact Newton methods (Dembo, Eisenstat, & Steihaug, 1982; Kelly, 1995). Its particular form is motivated by the use of IPM and the needs of the complexity analysis in Section 3. It is chosen in agreement with the definition of the infeasible neighbourhood (15) of the central path of the sub-problem considered. Using (12) and (16) in (9), we have

$$\|A\Delta \mathbf{x} + \delta \Delta \mathbf{y} - \xi_p\| = \|S_{\rho, \delta} \Delta \mathbf{y} - \bar{\xi}_p\| = \|\zeta\|,$$

whereas Eqs. (8) and (10) are satisfied exactly. Therefore the inexact Newton directions computed according to (16) satisfy:

$$\begin{bmatrix} \rho I & -A^T & -I \\ A & \delta I & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x} \\ \Delta \mathbf{y} \\ \Delta \mathbf{s} \end{bmatrix} = \begin{bmatrix} \xi_d \\ \xi_p \\ \xi_{\mu, \sigma} \end{bmatrix} + \begin{bmatrix} 0 \\ \zeta \\ 0 \end{bmatrix}. \tag{17}$$

Define

$$\begin{bmatrix} \mathbf{x}_k^j(\alpha) \\ \mathbf{y}_k^j(\alpha) \\ \mathbf{s}_k^j(\alpha) \end{bmatrix} := \begin{bmatrix} \mathbf{x}_k^j \\ \mathbf{y}_k^j \\ \mathbf{s}_k^j \end{bmatrix} + \begin{bmatrix} \alpha \Delta \mathbf{x}_k^j \\ \alpha \Delta \mathbf{y}_k^j \\ \alpha \Delta \mathbf{s}_k^j \end{bmatrix}, \tag{18}$$

i.e.  $\mathbf{x}_k^j(\alpha)$  is the point reached from  $\mathbf{x}_k^j$  after a step of length  $\alpha$  along the Newton direction. Notice that, after selecting the correct stepsize  $\alpha'_k$ , we define  $\mathbf{x}_k^{j+1} := \mathbf{x}_k^j(\alpha'_k)$ .

We report in Algorithm 2 a prototype IPM scheme for the solution of problem (PPM(k)). The fundamental steps involved in the algorithm are: computing the Newton direction by solving (17) with a level of inexactness that satisfies (16), see Line 2; finding the largest stepsize that guarantees to remain inside the neighbourhood and to sufficiently reduce the complementarity products, see Line 7; preparing the quantities to be used in the next iteration, see Lines 8–9.

We study the convergence of Algorithm 2 in Section 3. Concerning the notation, recall that the subscript  $k$  is related to the iteration count of the outer Algorithm 1 (PPM) whereas the superscript  $j$  is related to the iteration of the inner Algorithm 2 (IPM). To avoid over-complicating the notation, notice that in the following we use  $\xi_{p,k}^j$  and  $\xi_{d,k}^j$  instead of  $(\xi_p)_k^j$  and  $(\xi_d)_k^j$ .

**3. Convergence and complexity**

In this section, we show that the particular inexact IPM in Algorithm 2, used as inner solver in Algorithm 1, is convergent. Moreover, at the end of the present section, we show that such IPM converges to an  $\epsilon$ -accurate solution in a polynomial number of iterations. Our implant of the proof is inspired by the works (Armand & Benoist, 2013; Bellavia, 1998; Cornelis & Vanroose, 2021; Friedlander & Orban, 2012; Kojima, Megiddo, & Mizuno, 1993; Korzak, 2000) but consistently differs from the hypothesis and techniques used there.

Before continuing, let us observe that the PPM iteration counter  $k$  is fixed through this section and, for the sake of readability, is used only

**Input:**  $0 < \sigma, \bar{\sigma} < 1$ , barrier reduction parameters s.t.  $\sigma < \bar{\sigma}$ ;

$C_{\text{inexact}} \in (0, 1)$  inexactness parameter s.t.  $\gamma_p C_{\text{inexact}} < \sigma$ ;

**Initialization:**

Iteration counter  $j = 0$ ;

Primal–dual point  $(\mathbf{x}_k^0, \mathbf{y}_k^0, \mathbf{s}_k^0) \in \mathcal{N}_k(\bar{\gamma}, \underline{\gamma}, \gamma_p, \gamma_d)$

Compute  $\mu_k^0 := \mathbf{x}_k^0 T \mathbf{s}_k^0 / n$ ,  $\xi_{d,k}^0$ , and  $\xi_{p,k}^0$ .

```

1 while Stopping Criterion (4) False do
2   Solve the KKT system (17) using  $[\xi_{d,k}^j, \xi_{p,k}^j, \xi_{\mu_k, \sigma}^j]^T$  with
    $\|\zeta_k^j\| \leq C_{\text{inexact}}(\mathbf{x}_k^j)^T \mathbf{s}_k$  to find  $[\Delta \mathbf{x}_k^j, \Delta \mathbf{y}_k^j, \Delta \mathbf{s}_k^j]^T$ ;
3   Compute
    $\alpha_p^{*,j} = \sup\{\alpha \in \mathbb{R} : \mathbf{x}_k^j(\alpha) \geq 0\}$ 
    $\alpha_d^{*,j} = \sup\{\alpha \in \mathbb{R} : \mathbf{s}_k^j(\alpha) \geq 0\}$ 
   and define  $\alpha^{*,j} := \min\{\alpha_p^{*,j}, \alpha_d^{*,j}\}$ ;
4   if  $(\mathbf{x}_k(\alpha^{*,j}), \mathbf{y}_k(\alpha^{*,j}), \mathbf{s}_k(\alpha^{*,j}))$  is a solution of (PPM(k)) then
5     | Stop
6   end
7   Find  $\alpha'_k$  as the maximum  $\alpha \in [0, 1]$  s.t. for all  $\alpha \in [0, \alpha'_k]$ 
    $(\mathbf{x}_k^j(\alpha), \mathbf{y}_k^j(\alpha), \mathbf{s}_k^j(\alpha)) \in \mathcal{N}_k(\bar{\gamma}, \underline{\gamma}, \gamma_p, \gamma_d)$  and
    $\mathbf{x}_k^j(\alpha)^T \mathbf{s}_k^j(\alpha) \leq (1 - (1 - \bar{\sigma})\alpha) \mathbf{x}_k^j T \mathbf{s}_k^j$ ;
8   Set  $\begin{bmatrix} \mathbf{x}_k^{j+1} \\ \mathbf{y}_k^{j+1} \\ \mathbf{s}_k^{j+1} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_k^j \\ \mathbf{y}_k^j \\ \mathbf{s}_k^j \end{bmatrix} + \begin{bmatrix} \alpha'_k \Delta \mathbf{x}_k^j \\ \alpha'_k \Delta \mathbf{y}_k^j \\ \alpha'_k \Delta \mathbf{s}_k^j \end{bmatrix}$ ;
9   Compute the infeasibilities  $\xi_{d,k}^{j+1}$ ,  $\xi_{p,k}^{j+1}$  and barrier parameter
    $\mu_k^{j+1} := (\mathbf{x}_k^{j+1})^T \mathbf{s}_k^{j+1} / n$ ;
10  Update the iteration counter:  $j := j + 1$ .
11 end

```

**Algorithm 2:** IPM: inner loop of PS-IPM

when writing the fixed PPM iteration  $(\mathbf{x}_k, \mathbf{y}_k, \mathbf{s}_k)$  and not in the context of the IPM iterations  $(\mathbf{x}_k^j, \mathbf{y}_k^j, \mathbf{s}_k^j)$ .

We start from analysing the progress made in a single Newton iteration. Using (6), (7), and (8) we obtain

$$\begin{aligned} \rho(\mathbf{x}^j(\alpha) - \mathbf{x}_k) - A^T \mathbf{y}^j(\alpha) - \mathbf{s}^j(\alpha) + \mathbf{c} \\ = (\rho(\mathbf{x}^j - \mathbf{x}_k) - A^T \mathbf{y}^j - \mathbf{s}^j + \mathbf{c}) + \alpha(\rho \Delta \mathbf{x}^j - A^T \Delta \mathbf{y}^j - \Delta \mathbf{s}^j) \\ = (1 - \alpha)(\rho(\mathbf{x}^j - \mathbf{x}_k) - A^T \mathbf{y}^j - \mathbf{s}^j + \mathbf{c}), \end{aligned} \tag{20}$$

whereas, using (6) and (17) we have

$$\begin{aligned} A \mathbf{x}^j(\alpha) + \delta(\mathbf{y}^j(\alpha) - \mathbf{y}_k) - \mathbf{b} \\ = (A \mathbf{x}^j + \delta(\mathbf{y}^j - \mathbf{y}_k) - \mathbf{b}) + \alpha(A \Delta \mathbf{x}^j + \delta \Delta \mathbf{y}^j) \\ = (1 - \alpha)(A \mathbf{x}^j + \delta(\mathbf{y}^j - \mathbf{y}_k) - \mathbf{b}) + \alpha \zeta^j. \end{aligned} \tag{21}$$

The last block equation in (17) yields

$$(\mathbf{s}^j)^T \Delta \mathbf{x}^j + (\mathbf{x}^j)^T \Delta \mathbf{s}^j = -(\mathbf{x}^j)^T \mathbf{s}^j + \sigma n \mu^j = (\sigma - 1)(\mathbf{x}^j)^T \mathbf{s}^j \tag{22}$$

and

$$s_i \Delta x_i + x_i \Delta s_i = \sigma \frac{\mathbf{x}^T \mathbf{s}}{n} - x_i s_i. \tag{23}$$

Finally, using (22), we state the following identity

$$(\mathbf{x}^j + \alpha \Delta \mathbf{x}^j)^T (\mathbf{s}^j + \alpha \Delta \mathbf{s}^j) = (\mathbf{x}^j)^T \mathbf{s}^j (1 + \alpha(\sigma - 1)) + \alpha^2 (\Delta \mathbf{x}^j)^T \Delta \mathbf{s}^j. \tag{24}$$

With the next Theorem 1 we prove that Algorithm 2 is well-defined: at each iteration, there exist a non-empty interval of values for the stepsize  $\alpha$  such that the next iterate still lies in the neighbourhood  $\mathcal{N}_k(\bar{\gamma}, \underline{\gamma}, \gamma_p, \gamma_d)$  and such that the complementarity product  $(\mathbf{x}_k^j)^T \mathbf{s}_k^j$  is reduced by a sufficient amount, as required in (19).



**Theorem 1.** Let us suppose that  $(\mathbf{x}^j, \mathbf{y}^j, \mathbf{s}^j) \in \mathcal{N}_k(\bar{\gamma}, \underline{\gamma}, \gamma_p, \gamma_d)$  s.t.  $(\mathbf{x}^j)^T \mathbf{s}^j > 0$  is given. If the stopping conditions at Line 4 of Algorithm 2 are not satisfied, then there exists  $0 < \hat{\alpha}^j < \alpha^{*j}$  such that conditions (19) are satisfied for all  $\alpha \in [0, \hat{\alpha}^j]$ .

**Proof.** In this proof we omit also the IPM iterate counter  $j$ , i.e.  $(\mathbf{x}^j, \mathbf{y}^j, \mathbf{s}^j) \equiv (\mathbf{x}, \mathbf{y}, \mathbf{s})$ . Let us define the following functions, for all  $i = 1, \dots, n$

$$\begin{aligned} f_i(\alpha) &:= (x_i + \alpha \Delta x_i)(s_i + \alpha \Delta s_i) - \underline{\gamma}(\mathbf{x} + \alpha \Delta \mathbf{x})^T (\mathbf{s} + \alpha \Delta \mathbf{s})/n, \\ \bar{f}_i(\alpha) &:= \bar{\gamma}(\mathbf{x} + \alpha \Delta \mathbf{x})^T (\mathbf{s} + \alpha \Delta \mathbf{s})/n - (x_i + \alpha \Delta x_i)(s_i + \alpha \Delta s_i), \\ h(\alpha) &:= (1 - (1 - \bar{\sigma})\alpha)\mathbf{x}^T \mathbf{s} - (\mathbf{x} + \alpha \Delta \mathbf{x})^T (\mathbf{s} + \alpha \Delta \mathbf{s}), \\ g_d(\alpha) &:= (\mathbf{x} + \alpha \Delta \mathbf{x})^T (\mathbf{s} + \alpha \Delta \mathbf{s}) \\ &\quad - \gamma_d \|\rho(\mathbf{x} + \alpha \Delta \mathbf{x} - \mathbf{x}_k) - A^T(\mathbf{y} + \alpha \Delta \mathbf{y}) - (\mathbf{s} + \alpha \Delta \mathbf{s}) + \mathbf{c}\|, \\ g_p(\alpha) &:= (\mathbf{x} + \alpha \Delta \mathbf{x})^T (\mathbf{s} + \alpha \Delta \mathbf{s}) \\ &\quad - \gamma_p \|A(\mathbf{x} + \alpha \Delta \mathbf{x}) + \delta(\mathbf{y} + \alpha \Delta \mathbf{y} - \mathbf{y}_k) - \mathbf{b}\|. \end{aligned}$$

Using (20) in the expressions of  $g_d(\alpha)$  we have

$$g_d(\alpha) = (\mathbf{x} + \alpha \Delta \mathbf{x})^T (\mathbf{s} + \alpha \Delta \mathbf{s}) - \gamma_d(1 - \alpha)\|\rho(\mathbf{x} - \mathbf{x}_k) - A^T \mathbf{y} - \mathbf{s} + \mathbf{c}\|,$$

whereas using (21) in the expressions of  $g_p(\alpha)$  we have

$$g_p(\alpha) \geq (\mathbf{x} + \alpha \Delta \mathbf{x})^T (\mathbf{s} + \alpha \Delta \mathbf{s}) - \gamma_p((1 - \alpha)\|A\mathbf{x} + \delta(\mathbf{y} - \mathbf{y}_k) - \mathbf{b}\| + \alpha\|\xi\|).$$

We start proving that there exists  $\hat{\alpha}^j > 0$  such that

$$f_i(\alpha) \geq 0, \bar{f}_i(\alpha) \geq 0, h(\alpha) \geq 0, g_p(\alpha) \geq 0, g_d(\alpha) \geq 0$$

for all  $i = 1, \dots, n$  and for all  $\alpha \in [0, \hat{\alpha}^j]$ . In the following we will use extensively the identity (24). We have

$$\begin{aligned} f_i(\alpha) &= \underbrace{(1 - \alpha)(x_i s_i - \underline{\gamma} \frac{\mathbf{x}^T \mathbf{s}}{n})}_{\geq 0} + \alpha^2 (\Delta x_i \Delta s_i - \underline{\gamma} \frac{(\Delta \mathbf{x})^T \Delta \mathbf{s}}{n}) + \alpha \sigma (1 - \underline{\gamma}) \frac{\mathbf{x}^T \mathbf{s}}{n} \\ &\geq \alpha^2 (\Delta x_i \Delta s_i - \underline{\gamma} \frac{(\Delta \mathbf{x})^T \Delta \mathbf{s}}{n}) + \alpha \sigma (1 - \underline{\gamma}) \frac{\mathbf{x}^T \mathbf{s}}{n}. \end{aligned} \tag{25}$$

Since  $\mathbf{x}^T \mathbf{s} > 0$ , using a simple continuity argument, we can infer the existence of a small enough  $\underline{f}_i > 0$  s.t.  $f_i(\alpha) \geq 0$  for all  $\alpha \in [0, \underline{f}_i]$ . Reasoning analogously, we have

$$\begin{aligned} \bar{f}_i(\alpha) &= \underbrace{(1 - \alpha)(\bar{\gamma} \frac{\mathbf{x}^T \mathbf{s}}{n} - x_i s_i)}_{\geq 0} + \alpha^2 (\bar{\gamma} \frac{(\Delta \mathbf{x})^T \Delta \mathbf{s}}{n} - \Delta x_i \Delta s_i) + \alpha \sigma (\bar{\gamma} - 1) \frac{\mathbf{x}^T \mathbf{s}}{n} \\ &\geq \alpha^2 (\bar{\gamma} \frac{(\Delta \mathbf{x})^T \Delta \mathbf{s}}{n} - \Delta x_i \Delta s_i) + \alpha \sigma (\bar{\gamma} - 1) \frac{\mathbf{x}^T \mathbf{s}}{n}, \end{aligned} \tag{26}$$

and hence there exists a small enough  $\bar{f}_i > 0$  s.t.  $\bar{f}_i(\alpha) \geq 0$  for all  $\alpha \in [0, \bar{f}_i]$ .

Concerning  $h(\alpha)$ , we have

$$h(\alpha) = \mathbf{x}^T \mathbf{s}(\bar{\sigma} - \sigma) - \alpha^2 (\Delta \mathbf{x})^T \Delta \mathbf{s}, \tag{27}$$

and, since  $\mathbf{x}^T \mathbf{s}(\bar{\sigma} - \sigma) > 0$ , there exists  $\hat{h} > 0$  small enough s.t.  $h(\alpha) \geq 0$  for all  $\alpha \in [0, \hat{h}]$ .

Concerning  $g_d(\alpha)$ , we have

$$\begin{aligned} g_d(\alpha) &= \underbrace{(1 - \alpha)(\mathbf{x}^T \mathbf{s} - \gamma_d \|\rho(\mathbf{x} - \mathbf{x}_k) - A^T \mathbf{y} - \mathbf{s} + \mathbf{c}\|)}_{\geq 0} + \alpha \sigma \mathbf{x}^T \mathbf{s} + \alpha^2 (\Delta \mathbf{x})^T \Delta \mathbf{s} \\ &\geq \alpha \sigma \mathbf{x}^T \mathbf{s} + \alpha^2 (\Delta \mathbf{x})^T \Delta \mathbf{s}, \end{aligned} \tag{28}$$

and hence there exists  $\hat{g}_d > 0$  small enough s.t.  $g_d(\alpha) \geq 0$  for all  $\alpha \in [0, \hat{g}_d]$ .

Finally, concerning  $g_p(\alpha)$ , we have

$$\begin{aligned} g_p(\alpha) &\geq \underbrace{(1 - \alpha)(\mathbf{x}^T \mathbf{s} - \gamma_p \|A\mathbf{x} + \delta(\mathbf{y} - \mathbf{y}_k) - \mathbf{b}\|)}_{\geq 0} + \alpha \sigma \mathbf{x}^T \mathbf{s} \\ &\quad + \alpha^2 (\Delta \mathbf{x})^T \Delta \mathbf{s} - \alpha \gamma_p \|\xi\| \\ &\geq \alpha(\sigma - \gamma_p C_{\text{inexact}}) \mathbf{x}^T \mathbf{s} + \alpha^2 (\Delta \mathbf{x})^T \Delta \mathbf{s}, \end{aligned} \tag{29}$$

and hence there exists  $\hat{g}_p > 0$  small enough s.t.  $g_p(\alpha) \geq 0$  for all  $\alpha \in [0, \hat{g}_p]$ .

Let us define

$$\hat{\alpha}^j = \min\{\min_i \underline{f}_i, \min_i \bar{f}_i, \hat{h}, \hat{g}_d, \hat{g}_p, 1\} > 0.$$

To prove the thesis, it remains to show that  $\alpha^{*j} > \hat{\alpha}^j$ , i.e. that

$$(\mathbf{x}(\alpha), \mathbf{y}(\alpha), \mathbf{s}(\alpha)) \in \mathbb{R}_{>0}^n \times \mathbb{R}^m \times \mathbb{R}_{>0}^n \text{ for all } \alpha \in [0, \hat{\alpha}^j].$$

To this aim, let us suppose by contradiction that  $\alpha^{*j} \leq \hat{\alpha}^j$ . By definition of  $\alpha^{*j}$ , there exists  $\bar{\ell} \in \{1, \dots, n\}$  s.t.  $(x_{\bar{\ell}} + \alpha^{*j} \Delta x_{\bar{\ell}})(s_{\bar{\ell}} + \alpha^{*j} \Delta s_{\bar{\ell}}) = 0$ . We have hence

$$f_{\bar{\ell}}(\alpha^{*j}) = -\underline{\gamma}(\mathbf{x}(\alpha^{*j}))^T \mathbf{s}(\alpha^{*j})/n \geq 0 \Rightarrow (\mathbf{x}(\alpha^{*j}))^T \mathbf{s}(\alpha^{*j}) = 0.$$

From the above implication, using  $g_d(\alpha^{*j})$  and  $g_p(\alpha^{*j})$ , we obtain that

$$\begin{aligned} A\mathbf{x}(\alpha^{*j}) + \delta(\mathbf{y}(\alpha^{*j}) - \mathbf{y}_k) - \mathbf{b} &= 0 \\ \rho(\mathbf{x}(\alpha^{*j}) - \mathbf{x}_k) - A^T \mathbf{y}(\alpha^{*j}) - \mathbf{s}(\alpha^{*j}) + \mathbf{c} &= 0, \end{aligned}$$

i.e.  $(\mathbf{x}(\alpha^{*j}), \mathbf{y}(\alpha^{*j}), \mathbf{s}(\alpha^{*j}))$  is a solution of problem (PPM( $k$ )). We have hence obtained a contradiction since we are supposing that Algorithm 2 did not stop at Line 4.  $\square$

Before proving the convergence of Algorithm 2 we would like to emphasize that the above proof complements and expands (Kojima et al., 1993, Remark 3.1).

The next two results establish that Algorithm 2 converges to a solution of problem (PPM( $k$ )). This is done by establishing that the right-hand sides of the Newton systems are uniformly bounded and by showing that the complementarity product  $(\mathbf{x}^j)^T \mathbf{s}^j$  cannot be bounded away from zero.

**Corollary 1.** The right-hand sides of the Newton systems are uniformly bounded.

**Proof.** As a consequence of Theorem 1, we can suppose the existence of a sequence of iterates  $\{(\mathbf{x}^j, \mathbf{y}^j, \mathbf{s}^j)\}_{j \in \mathbb{N}}$  produced by Algorithm 2 s.t.

$$(\mathbf{x}^j, \mathbf{y}^j, \mathbf{s}^j) \in \mathcal{N}_k(\bar{\gamma}, \underline{\gamma}, \gamma_p, \gamma_d).$$

Since by construction  $(\mathbf{x}^j)^T \mathbf{s}^j \leq (\mathbf{x}^0)^T \mathbf{s}^0$ , we have from (15)

$$\begin{aligned} \|A\mathbf{x}^j + \delta(\mathbf{y}^j - \mathbf{y}_k) - \mathbf{b}\| &\leq (\mathbf{x}^0)^T \mathbf{s}^0 / \gamma_p, \\ \|\rho(\mathbf{x}^j - \mathbf{x}_k) - A^T \mathbf{y}^j - \mathbf{s}^j + \mathbf{c}\| &\leq (\mathbf{x}^0)^T \mathbf{s}^0 / \gamma_d. \end{aligned}$$

Moreover, we have

$$\begin{aligned} \|S^j X^j \mathbf{e} - \sigma \mu^j \mathbf{e}\| &\leq \|S^j X^j \mathbf{e}\| + \sigma \mu^j \|\mathbf{e}\| \\ &\leq \frac{\bar{\gamma}}{\sqrt{n}} (\mathbf{x}^j)^T \mathbf{s}^j + \frac{\sigma}{\sqrt{n}} (\mathbf{x}^j)^T \mathbf{s}^j \leq \frac{\bar{\gamma} + \sigma}{\sqrt{n}} (\mathbf{x}^0)^T \mathbf{s}^0. \quad \square \end{aligned}$$

**Theorem 2.** Algorithm 2 produces a sequence of iterates in  $\mathcal{N}_k(\bar{\gamma}, \underline{\gamma}, \gamma_p, \gamma_d)$  s.t.  $\lim(\mathbf{x}^j)^T \mathbf{s}^j = 0$ , i.e.  $(\mathbf{x}^j, \mathbf{y}^j, \mathbf{s}^j)$  converges to a solution of problem (PPM( $k$ )).

**Proof.** Let us argue by contradiction supposing that there exists  $\varepsilon^* > 0$  s.t.  $(\mathbf{x}^j)^T \mathbf{s}^j > \varepsilon^*$  for all  $j \in \mathbb{N}$ .

**Claim 1** There exists a constant  $C_1$  dependent only on  $n$  s.t.

$$\|[\Delta \mathbf{x}^j, \Delta \mathbf{y}^j, \Delta \mathbf{s}^j]^T\| \leq C_1 \text{ for all } j \in \mathbb{N}.$$

The proof of this fact follows observing that the Newton matrices in (17) satisfy all the hypothesis of Armand and Benoist (2013, Th. 1),

i.e. they have a uniformly bounded inverse, and that the right-hand sides are uniformly bounded, see [Corollary 1](#). As a consequence, there exists another constant  $C_2$  s.t.

$$\begin{aligned} |\Delta x_i^j \Delta s_i^j - \underline{\gamma} \frac{(\Delta \mathbf{x}^j)^T \Delta \mathbf{s}^j}{n}| &\leq C_2, \\ |\bar{\gamma} \frac{(\Delta \mathbf{x}^j)^T \Delta \mathbf{s}^j}{n} - \Delta x_i^j \Delta s_i^j| &\leq C_2, \\ |(\Delta \mathbf{x}^j)^T \Delta \mathbf{s}^j| &\leq C_2. \end{aligned} \tag{30}$$

**Claim 2** There exists  $\alpha^* > 0$  s.t.  $\alpha^j \geq \alpha^*$  for all  $j \in \mathbb{N}$ .

Using (30) in Eqs. (25), (26), (27), (28), (29) we have

$$\begin{aligned} f_i(\alpha) &\geq \alpha^2 (\Delta x_i^j \Delta s_i^j - \underline{\gamma} \frac{(\Delta \mathbf{x}^j)^T \Delta \mathbf{s}^j}{n}) + \alpha \sigma (1 - \underline{\gamma}) \frac{(\mathbf{x}^j)^T \mathbf{s}^j}{n} \\ &\geq -C_2 \alpha^2 + \alpha \sigma (1 - \underline{\gamma}) \varepsilon^* / n, \\ \bar{f}_i(\alpha) &\geq (\bar{\gamma} \frac{(\Delta \mathbf{x}^j)^T \Delta \mathbf{s}^j}{n} - \Delta x_i^j \Delta s_i^j) + \alpha \sigma (\bar{\gamma} - 1) \frac{(\mathbf{x}^j)^T \mathbf{s}^j}{n} \\ &\geq -C_2 \alpha^2 + \alpha \sigma (\bar{\gamma} - 1) \varepsilon^* / n, \\ h(\alpha) &= (\mathbf{x}^j)^T \mathbf{s}^j (\bar{\sigma} - \sigma) \alpha - \alpha^2 (\Delta \mathbf{x}^j)^T \Delta \mathbf{s}^j \\ &\geq \varepsilon^* (\bar{\sigma} - \sigma) \alpha - C_2 \alpha^2, \\ g_d(\alpha) &\geq \alpha \sigma (\mathbf{x}^j)^T \mathbf{s}^j + \alpha^2 (\Delta \mathbf{x}^j)^T \Delta \mathbf{s}^j \geq \varepsilon^* \sigma \alpha - C_2 \alpha^2, \\ g_p(\alpha) &\geq \alpha (\sigma - \gamma_p) (\mathbf{x}^j)^T \mathbf{s}^j + \alpha^2 (\Delta \mathbf{x}^j)^T \Delta \mathbf{s}^j \geq \varepsilon^* \alpha (\sigma - \gamma_p C_{\text{inexact}}) - C_2 \alpha^2. \end{aligned}$$

Hence  $\alpha^j \geq \alpha^*$ , where

$$\alpha^* := \min \left\{ 1, \frac{\sigma(1-\underline{\gamma})\varepsilon^*/n}{C_2}, \frac{\sigma(\bar{\gamma}-1)\varepsilon^*/n}{C_2}, \frac{(\bar{\sigma}-\sigma)\varepsilon^*}{C_2}, \frac{\sigma\varepsilon^*}{C_2}, \frac{(\sigma-\gamma_p C_{\text{inexact}})\varepsilon^*}{C_2} \right\}. \tag{31}$$

The convergence claim follows observing that the inequality

$$\varepsilon^* \leq (\mathbf{x}^j)^T \mathbf{s}^j \leq (1 - (1 - \bar{\sigma})\alpha^j) (\mathbf{x}^0)^T \mathbf{s}^0$$

leads to a contradiction for  $j \rightarrow \infty$ .  $\square$

### 3.1. Polynomial complexity

In this section we show that the number of iterations needed to reduce  $\mu^j$  below a certain tolerance  $\varepsilon$  grows polynomially with the size of the problem. Moreover, it is important to note that, from the definition of the central path  $\mathcal{N}_k(\bar{\gamma}, \underline{\gamma}, \gamma_p, \gamma_d)$ , the same number of iterations is sufficient to reduce also the primal and dual infeasibility below the tolerance  $\varepsilon$ .

As it is customary in IPM literature, in this section we make the following assumption:

**Assumption 3.** The norm of  $A$ , i.e.,  $\|A\|$ , is independent of  $n$ .

In the following we will omit the index  $j$  when this does not lead to ambiguities. It is important to note that the linear system in (17) can be written in an alternative form as follows

$$\underbrace{\begin{bmatrix} \rho I & A^T & -I \\ A & -\delta I & 0 \\ S & 0 & X \end{bmatrix}}_{=: J} \begin{bmatrix} \Delta \mathbf{x} \\ -\Delta \mathbf{y} \\ \Delta \mathbf{s} \end{bmatrix} = \begin{bmatrix} \xi_d \\ \xi_p + \zeta \\ \xi_{\mu, \sigma} \end{bmatrix}.$$

Using (Armand & Benoist, 2013, Remark 1), we have that

$$J^{-1} = \begin{bmatrix} H^{-1} & \frac{1}{\delta} H^{-1} A^T & H^{-1} X^{-1} \\ \frac{1}{\delta} A H^{-1} & \frac{1}{\delta^2} A H^{-1} A^T - \frac{1}{\delta} I & \frac{1}{\delta} A H^{-1} X^{-1} \\ -\Theta^{-1} H^{-1} & -\frac{1}{\delta} \Theta^{-1} H^{-1} A^T & (I - \Theta^{-1} H^{-1}) X^{-1} \end{bmatrix}, \tag{32}$$

where  $H := \rho I + \Theta^{-1} + \frac{1}{\delta} A^T A$ .

To prove polynomial complexity, we start by bounding the terms that appear in the expression (32). The next two technical results are useful in this sense.

**Lemma 1.** We have that

$$\|H^{-1}\| \in O(\|A\| \|A^T\|).$$

**Proof.** Using the Sherman–Morrison–Woodbury formula, we get

$$H^{-1} = (\rho I + \Theta^{-1})^{-1} - \frac{1}{\delta} (\rho I + \Theta^{-1})^{-1} A^T \left( I + \frac{1}{\delta} A (\rho I + \Theta^{-1})^{-1} A^T \right)^{-1} A (\rho I + \Theta^{-1})^{-1}. \tag{33}$$

We observe

$$(\rho I + \Theta^{-1})_{ii}^{-1} = \frac{\Theta_{ii}}{\rho \Theta_{ii} + 1} = \frac{\rho \Theta_{ii}}{\rho \Theta_{ii} + 1} \frac{1}{\rho} < \frac{1}{\rho} \tag{34}$$

and hence

$$\begin{aligned} \|H^{-1}\| &\leq \|(\rho I + \Theta^{-1})^{-1}\| \cdot \\ &\quad \cdot \left( 1 + \frac{1}{\delta} \|A^T\| \|A\| \left\| \left( I + \frac{1}{\delta} A (\rho I + \Theta^{-1})^{-1} A^T \right)^{-1} \right\| \|(\rho I + \Theta^{-1})^{-1}\| \right) \\ &\leq \frac{1}{\rho} \left( 1 + \frac{1}{\delta \rho} \|A^T\| \|A\| \right), \end{aligned}$$

where we used that  $\| \left( I + \frac{1}{\delta} A (\rho I + \Theta^{-1})^{-1} A^T \right)^{-1} \| \leq 1$ .  $\square$

**Corollary 2.** We have that

$$\|\Theta^{-1} H^{-1}\| \in O(\|A\| \|A^T\|).$$

**Proof.** Using (34), we observe that

$$(\Theta^{-1} (\rho I + \Theta^{-1})^{-1})_{ii} = \frac{1}{\rho \Theta_{ii} + 1}$$

and hence, using (33), we have

$$\begin{aligned} \|\Theta^{-1} H^{-1}\| &\leq \|\Theta^{-1} (\rho I + \Theta^{-1})^{-1}\| \cdot \\ &\quad \left( 1 + \frac{1}{\delta} \|A^T\| \|A\| \left\| \left( I + \frac{1}{\delta} A (\rho I + \Theta^{-1})^{-1} A^T \right)^{-1} \right\| \|(\rho I + \Theta^{-1})^{-1}\| \right) \\ &\leq 1 + \frac{1}{\delta \rho} \|A^T\| \|A\|, \end{aligned}$$

where we used (34).  $\square$

Therefore, we know that

$$\|H^{-1}\| \leq C_3 \text{ and } \|\Theta^{-1} H^{-1}\| \leq C_4,$$

for some positive constants  $C_3$  and  $C_4$  independent of  $n$ , see [Assumption 3](#). Let us define  $C_5 := \max\{C_3, C_4\}$ .

Now that we have bounded the terms in (32), we look for an upper bound on the norms of the Newton directions that depend polynomially on the size of the problem  $n$ . This is a crucial step to find a polynomial lower bound on the minimum stepsize (31) which leads to the polynomial complexity result mentioned at the beginning of this section.

**Theorem 3.** There exists a positive constant  $C_6$  independent of  $n$  s.t.

$$\|[\Delta \mathbf{x}^j, \Delta \mathbf{y}^j, \Delta \mathbf{s}^j]^T\| \leq C_6 n \sqrt{\mu^j} \text{ for all } j \in \mathbb{N}.$$

**Proof.** Using (32), we have

$$\begin{aligned} \Delta \mathbf{x}^j &= (H^j)^{-1} \xi_d^j + \frac{1}{\delta} (H^j)^{-1} A^T (\xi_p^j + \zeta^j) + (H^j)^{-1} (X^j)^{-1} \xi_{\sigma, \mu^j}^j, \\ -\Delta \mathbf{y}^j &= \frac{1}{\delta} A (H^j)^{-1} \xi_d^j + \left( \frac{1}{\delta^2} A (H^j)^{-1} A^T - \frac{1}{\delta} I \right) (\xi_p^j + \zeta^j) \\ &\quad + \frac{1}{\delta} A (H^j)^{-1} (X^j)^{-1} \xi_{\sigma, \mu^j}^j, \end{aligned}$$

and hence

$$\begin{aligned} \|\Delta \mathbf{x}^j\| &\leq \|(H^j)^{-1}\| \|\xi_d^j\| + \frac{1}{\delta} \|(H^j)^{-1}\| \|A^T\| \|\xi_p^j + \zeta^j\| \\ &\quad + \|(H^j)^{-1/2}\| \|(H^j)^{-1/2}(X^j)^{-1/2}(S^j)^{1/2}\| \\ &\quad \cdot \|(X^j)^{1/2}(S^j)^{1/2}\mathbf{e} - \sigma\mu^j(X^j)^{-1/2}(S^j)^{-1/2}\mathbf{e}\| \\ &\leq \frac{C_5 n \mu^j}{\gamma_d} + \frac{C_5 \|A^T\| (1 + \gamma_p C_{\text{inexact}}) n \mu^j}{\delta \gamma_p} + C_5 \left( \sqrt{\bar{\gamma} n} + \sigma \frac{\sqrt{n}}{\sqrt{\bar{\gamma}}} \right) \sqrt{\mu^j} \quad (35) \\ &\leq \left( \frac{C_5 \sqrt{\mu^0}}{\gamma_d} + \frac{C_5 \|A^T\| (1 + \gamma_p C_{\text{inexact}}) \sqrt{\mu^0}}{\delta \gamma_p} + C_5 \left( \sqrt{\bar{\gamma}} + \sigma \frac{1}{\sqrt{\bar{\gamma}}} \right) \right) n \sqrt{\mu^j} \\ &\leq C_{\Delta x} n \sqrt{\mu^j}, \end{aligned}$$

where we used

$$\begin{aligned} &\|(X^j)^{1/2}(S^j)^{1/2}\mathbf{e} - \sigma\mu^j(X^j)^{-1/2}(S^j)^{-1/2}\mathbf{e}\| \\ &\leq \|(X^j)^{1/2}(S^j)^{1/2}\mathbf{e}\| + \|\sigma\mu^j(X^j)^{-1/2}(S^j)^{-1/2}\mathbf{e}\|, \end{aligned}$$

$\mu^j = \sqrt{\mu^j} \sqrt{\mu^j} \leq \sqrt{\mu^0} \sqrt{\mu^j}$  and  $\sqrt{n} \leq n$ , and where  $C_{\Delta x}$  is a positive constant independent of  $n$  thanks to Assumption 3.

Analogously

$$\begin{aligned} \|\Delta \mathbf{y}^j\| &\leq \frac{1}{\delta} \|A\| \|(H^j)^{-1}\| \|\xi_d^j\| + \left( \frac{1}{\delta^2} \|(H^j)^{-1}\| \|A\| \|A^T\| + \frac{1}{\delta} \right) \|\xi_p^j + \zeta^j\| \\ &\quad + \frac{1}{\delta} \|A\| \|(H^j)^{-1/2}\| \|(H^j)^{-1/2}(X^j)^{-1/2}(S^j)^{1/2}\| \cdot \\ &\quad \cdot \|(X^j)^{1/2}(S^j)^{1/2}\mathbf{e} - \sigma\mu^j(X^j)^{-1/2}(S^j)^{-1/2}\mathbf{e}\| \\ &\leq \frac{\|A\| C_5 n \mu^j}{\delta \gamma_d} + \frac{(C_5 \|A^T\| \|A\| + \delta)(1 + \gamma_p C_{\text{inexact}}) n \mu^j}{\delta^2 \gamma_p} + \\ &\quad + \frac{C_5 \|A\|}{\delta} \left( \sqrt{\bar{\gamma} n} + \sigma \frac{\sqrt{n}}{\sqrt{\bar{\gamma}}} \right) \sqrt{\mu^j} \\ &\leq \left( \frac{\|A\| C_5 \sqrt{\mu^0}}{\delta \gamma_d} + \frac{(C_5 \|A^T\| \|A\| + \delta)(1 + \gamma_p C_{\text{inexact}}) \sqrt{\mu^0}}{\delta^2 \gamma_p} + \right. \\ &\quad \left. + \frac{C_5 \|A\|}{\delta} \left( \sqrt{\bar{\gamma}} + \sigma \frac{1}{\sqrt{\bar{\gamma}}} \right) \right) n \sqrt{\mu^j} \\ &\leq C_{\Delta y} n \sqrt{\mu^j}, \quad (36) \end{aligned}$$

where  $C_{\Delta y}$  is a positive constant independent of  $n$ , see, once more, Assumption 3. Finally, using the fact that  $\Delta \mathbf{s}^j = \rho \Delta \mathbf{x}^j - A^T \Delta \mathbf{y}^j - \xi_d^j$  and using (35), (36), and the definition of  $\mathcal{N}_k(\bar{\gamma}, \underline{\gamma}, \gamma_p, \gamma_d)$ , we have

$$\begin{aligned} \|\Delta \mathbf{s}^j\| &= \rho \|\Delta \mathbf{x}^j\| + \|A^T\| \|\Delta \mathbf{y}^j\| + \|\xi_d^j\| \\ &\leq \rho C_{\Delta x} n \sqrt{\mu^j} + \|A^T\| C_{\Delta y} n \sqrt{\mu^j} + \frac{n \mu^j}{\gamma_d} \\ &\leq \left( \rho C_{\Delta x} + \|A^T\| C_{\Delta y} + \frac{\sqrt{\mu^0}}{\gamma_d} \right) n \sqrt{\mu^j} \\ &\leq C_{\Delta s} n \sqrt{\mu^j}, \end{aligned}$$

where  $C_{\Delta s}$  is a positive constant independent of  $n$ . The thesis follows setting  $C_6 = \max(C_{\Delta x}, C_{\Delta y}, C_{\Delta s})$ .  $\square$

The next straightforward technical result specializes the polynomial bound of Theorem 3 to the terms that appear in Eqs. (25)–(29).

**Corollary 3.** *There exists a positive constant  $C_7$  independent of  $n$  such that, for all  $i$*

$$\begin{aligned} \left| \Delta x_i \Delta s_i - \underline{\gamma} \frac{(\Delta \mathbf{x})^T \Delta \mathbf{s}}{n} \right| &\leq C_7 n^2 \mu, \\ \left| \bar{\gamma} \frac{(\Delta \mathbf{x})^T \Delta \mathbf{s}}{n} - \Delta x_i \Delta s_i \right| &\leq C_7 n^2 \mu, \\ |(\Delta \mathbf{x})^T \Delta \mathbf{s}| &\leq C_7 n^2 \mu. \quad (37) \end{aligned}$$

We can now apply the previous results and obtain a bound similar to (31), but that depends polynomially on the size of the problem  $n$ . This is the last fundamental step before the polynomial complexity result can be stated.

**Theorem 4.** *There exists a constant  $\bar{\alpha}$  s.t.  $\alpha^j \geq \bar{\alpha}$  for all  $j \in \mathbb{N}$  and*

$$\bar{\alpha} \geq C_8 n^{-2},$$

where  $C_8$  is a positive constant independent of  $n$ .

**Proof.** Using (37) in Eqs. (25), (26), (27), (28), (29) we have:

$$\begin{aligned} f_i(\alpha) &\geq \alpha^2 \left( \Delta x_i \Delta s_i - \underline{\gamma} \frac{(\Delta \mathbf{x})^T \Delta \mathbf{s}}{n} \right) + \alpha \sigma (1 - \underline{\gamma}) \frac{\mathbf{x}^T \mathbf{s}}{n} \\ &\geq -C_7 n^2 \mu \alpha^2 + \alpha \sigma (1 - \underline{\gamma}) \mu, \\ \bar{f}_i(\alpha) &\geq \alpha^2 \left( \bar{\gamma} \frac{(\Delta \mathbf{x})^T \Delta \mathbf{s}}{n} - \Delta x_i \Delta s_i \right) + \alpha \sigma (\bar{\gamma} - 1) \frac{\mathbf{x}^T \mathbf{s}}{n} \\ &\geq -C_7 n^2 \mu \alpha^2 + \alpha \sigma (\bar{\gamma} - 1) \mu, \\ h(\alpha) &= \mathbf{x}^T \mathbf{s} (\bar{\sigma} - \sigma) \alpha - \alpha^2 (\Delta \mathbf{x})^T \Delta \mathbf{s} \geq n \mu (\bar{\sigma} - \sigma) \alpha - C_7 n^2 \mu \alpha^2, \\ g_d(\alpha) &\geq \alpha \sigma \mathbf{x}^T \mathbf{s} + \alpha^2 (\Delta \mathbf{x})^T \Delta \mathbf{s} \geq n \mu \sigma \alpha - C_7 n^2 \mu \alpha^2, \\ g_p(\alpha) &\geq \alpha (\sigma - \gamma_p C_{\text{inexact}}) \mathbf{x}^T \mathbf{s} + \alpha^2 (\Delta \mathbf{x})^T \Delta \mathbf{s} \geq n \mu \alpha (\sigma - \gamma_p C_{\text{inexact}}) - C_7 n^2 \mu \alpha^2. \end{aligned}$$

Hence, defining

$$\bar{\alpha} := \min \left\{ 1, \frac{\sigma(1 - \underline{\gamma})}{C_7 n^2}, \frac{\sigma(\bar{\gamma} - 1)}{C_7 n^2}, \frac{(\bar{\sigma} - \sigma)}{C_7 n}, \frac{\sigma}{C_7 n}, \frac{(\sigma - \gamma_p C_{\text{inexact}})}{C_7 n} \right\},$$

the thesis follows observing that, by definition,  $\alpha^j \geq \bar{\alpha}$ .  $\square$

Finally, we are ready to show that the number of iterations required to reduce  $\mu$  below a certain tolerance  $\varepsilon$  is proportional to  $n^2$ .

**Theorem 5.** *Algorithm 2 has polynomial complexity, i.e. given  $\varepsilon > 0$  there exists  $K \in O(n^2 \ln(\frac{1}{\varepsilon}))$  s.t.  $\mu^j \leq \varepsilon$  for all  $j \geq K$ .*

**Proof.** Thesis follows observing that

$$(\mathbf{x}^j)^T \mathbf{s}^j \leq (1 - (1 - \bar{\sigma}) \bar{\alpha})^j (\mathbf{x}^0)^T \mathbf{s}^0 \leq \left( 1 - (1 - \bar{\sigma}) \frac{C_8}{n^2} \right)^j (\mathbf{x}^0)^T \mathbf{s}^0. \quad \square$$

Before concluding this section it is important to note that, with a simple modification of the arguments here contained, it is possible to prove polynomial convergence also when  $\|A\|$  grows linearly with the dimension  $n$ . Nevertheless, in this case the term  $O(n^2 \ln(\frac{1}{\varepsilon}))$  appearing in the statement of Theorem 5 has to be replaced with  $O(h(n) \ln(\frac{1}{\varepsilon}))$  for a suitable polynomial  $h(n)$  of degree strictly greater than two.

#### 4. Properties of the regularized normal equations system

In this section we show some properties of matrix  $S_{\rho, \delta}$  that are useful for the analysis performed in the next section.

For the original graph  $G(V, E)$ , we define the adjacency matrix  $\mathcal{A} \in \mathbb{R}^{|V| \times |V|}$  such that  $\mathcal{A}_{ij} = 1$  if there exists an edge between nodes  $i$  and  $j$  and  $\mathcal{A}_{ij} = 0$  otherwise. Notice that for an undirected graph  $\mathcal{A}$  is symmetric; we assume that there are no self-loops, so that the diagonal of  $\mathcal{A}$  is made of zeros. Let us define the degree matrix of a graph  $\mathcal{D} \in \mathbb{R}^{|V| \times |V|}$  such that  $\mathcal{D}$  is diagonal and  $\mathcal{D}_{jj}$  is the degree of node  $j$ . Notice that  $\mathcal{D}_{jj} = (\mathcal{A} \mathbf{e})_j$ .

Let us define also the Laplacian matrix of a graph as  $\mathcal{L} \in \mathbb{R}^{|V| \times |V|}$  such that  $\mathcal{L} = \mathcal{D} - \mathcal{A}$ . An important relation between the Laplacian  $\mathcal{L}$  and the node-arc incidence matrix  $A$  is that  $\mathcal{L} = AA^T$ .

Given a diagonal matrix  $\Theta$  and a parameter  $\rho$ , we define the re-weighted graph  $G_\Theta$  as the graph with the same connectivity of  $G$ , in which the weight of every edge  $j$  is scaled by a factor  $\sqrt{\frac{\Theta_{jj}}{1 + \rho \Theta_{jj}}}$ . The adjacency matrix of the new graph  $\mathcal{A}_\Theta$  has the same sparsity pattern of  $\mathcal{A}$ , but takes into account the new weight of the edges. The same happens for the degree matrix  $\mathcal{D}_\Theta$ . The incidence matrix therefore becomes  $A_\Theta = A(\Theta^{-1} + \rho I)^{-1/2}$ .

The new Laplacian matrix thus reads  $\mathcal{L}_\Theta = \mathcal{D}_\Theta - \mathcal{A}_\Theta$  and can be written as

$$\mathcal{L}_\Theta = A_\Theta A_\Theta^T = A(\Theta^{-1} + \rho I)^{-1} A^T.$$

We have just shown that the normal equations matrix can be interpreted as the Laplacian matrix of the original graph, where the edges are weighted according to the diagonal entries of matrix  $\Theta$ . This result is important because solving linear systems that involve Laplacian matrices is much easier than solving general linear systems. We summarize this result in the following Lemma.

**Lemma 2.** *The matrix  $A(\Theta^{-1} + \rho I)^{-1} A^T$  is the Laplacian of a weighted undirected graph and hence, for every  $\Theta \in \mathbb{R}_+^{n \times n}$*

$$A(\Theta^{-1} + \rho I)^{-1} A^T = D_\Theta - \mathcal{A}_\Theta, \tag{38}$$

where  $D_\Theta$  and  $\mathcal{A}_\Theta$  are the degree and adjacency matrices of the weighted graph.

The next result shows that the normal matrix is strictly diagonally dominant, due to the presence of dual regularization. This property is significant because it assures that the incomplete Cholesky factorization of the normal equations matrix  $S_{\rho,\delta}$  can always be computed without the algorithm breaking down (in exact arithmetic), see e.g. Manteuffel (1980).

**Lemma 3.** *If  $\delta > 0$  the matrix  $S_{\rho,\delta}$  is strictly diagonally dominant.*

**Proof.** From Lemma 2, we have that  $A(\Theta^{-1} + \rho I)^{-1} A^T = D_\Theta - \mathcal{A}_\Theta$  and hence

$$\sum_{j \neq i} |(S_{\rho,\delta})_{ij}| = \sum_{j \neq i} |-(\mathcal{A}_\Theta)_{ij}| = \sum_{j \neq i} (\mathcal{A}_\Theta)_{ij} < (D_\Theta)_{ii} + \delta = (S_{\rho,\delta})_{ii}. \quad \square$$

The next two technical results are related to the distribution of eigenvalues of the normal matrix. They are used in the next section to show that the inexactness introduced when sparsifying the normal matrix remains bounded. Moreover, when Lemma 4 is used in conjunction with the sparsity Assumption 1, proves that the polynomial complexity stated in Theorem 5 holds. Indeed, Lemma 4 implies that Assumption 3 is satisfied for the class of problems considered in this work.

**Lemma 4.**  $\|A\|^2 := \lambda_{\max}(AA^T) \leq 2 \max_{v \in V} \text{deg}(v) \leq 2C_{\text{deg}}$ .

**Proof.** The proof of the first inequality is a straightforward application of the Gershgorin’s Circles Theorems, whereas the second one follows from Assumption 1.  $\square$

**Lemma 5.** *The eigenvalues  $\lambda$  of matrix  $S_{\rho,\delta}$  satisfy*

$$\delta \leq \lambda < \delta + \frac{2}{\rho} \max_{v \in V} \text{deg}(v).$$

**Proof.** Using the Rayleigh quotient, for some vector  $\mathbf{u}$  and  $\mathbf{v} = A^T \mathbf{u}$ , the eigenvalues can be written as

$$\lambda = \frac{\mathbf{v}^T (\Theta^{-1} + \rho I)^{-1} \mathbf{v}}{\mathbf{v}^T \mathbf{v}} \frac{\mathbf{u}^T A A^T \mathbf{u}}{\mathbf{u}^T \mathbf{u}} + \delta$$

The lower bound  $\lambda \geq \delta$  is trivial; the upper bound follows from Lemma 4 and (34).  $\square$

### 5. Sparsification of the reduced matrix

We now propose a technique to reduce the number of nonzeros in the normal equations  $S_{\rho,\delta}$ , based on the weights of the edges in the re-weighted graph (according to Lemma 2). We then show that this sparsification strategy is sound and produces a polynomially convergent interior point algorithm.

In this section we omit the IPM iteration counter  $j$  and we consider all the IPM-related quantities as a function of  $\mu \rightarrow 0$ . As IPMs progress towards optimality, we expect the following partition of the diagonal matrix  $\Theta$  contributed by the barrier term:

$$\mathcal{B} := \{i = 1, \dots, n \text{ s.t. } x_i \rightarrow x_i^* > 0, s_i \rightarrow s_i^* = 0\}$$

$$\mathcal{N} := \{i = 1, \dots, n \text{ s.t. } x_i \rightarrow x_i^* = 0, s_i \rightarrow s_i^* > 0\},$$

where the optimal solution  $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{s}^*)$  was defined in (3). Notice that  $(\mathcal{B}, \mathcal{N})$  is the partition corresponding to the optimal solution.

**Assumption 4.** We suppose that the following asymptotic estimates hold

$$s_i \in O(\mu) \text{ and } x_i \in O(1) \text{ for } i \in \mathcal{B} \tag{39}$$

$$x_i \in O(\mu) \text{ and } s_i \in O(1) \text{ for } i \in \mathcal{N}$$

and, since  $x_i^{-1} s_i \approx \mu x_i^{-2}$  when an IPM iterate is sufficiently close to the central path, using (39), we suppose

$$\Theta_{ii}^{-1} = x_i^{-1} s_i = O(\mu) \text{ for } i \in \mathcal{B} \text{ and } \Theta_{ii}^{-1} = x_i^{-1} s_i = O(\mu^{-1}) \text{ for } i \in \mathcal{N}.$$

This assumption makes sense given the neighbourhood that is considered, see e.g. Gondzio (2012b).

Due to Assumption 4, we consider the following asymptotic estimates of  $(\Theta^{-1} + \rho I)^{-1}$

$$(\Theta^{-1} + \rho I)_{ii}^{-1} = \begin{cases} O(\frac{1}{\rho + \mu}) & \text{if } i \in \mathcal{B} \\ O(\frac{\mu}{1 + \rho\mu}) & \text{if } i \in \mathcal{N}. \end{cases}$$

The diagonal entries of  $\Theta$  give a specific weight to each column of matrix  $A$  (or equivalently, give a weight to each edge of the original sparse graph, as shown in Lemma 2). The columns for which the corresponding  $\Theta_{ii}$  is  $O(\mu)$  have a very small impact on the normal matrix, but still contribute to its sparsity pattern. In order to save time and memory when forming (complete or incomplete) Cholesky factorizations, we propose the following sparsification strategy: we introduce a suitable threshold  $C_i \in \mathbb{R}_+$  and define

$$(\hat{\Theta}_{C_i, \mu, \rho}^\dagger)_{ii} := \begin{cases} (\Theta^{-1} + \rho I)_{ii}^{-1} & \text{if } (\Theta^{-1} + \rho I)_{ii}^{-1} \geq \frac{C_i \mu}{1 + \rho \mu} \\ 0 & \text{if } (\Theta^{-1} + \rho I)_{ii}^{-1} < \frac{C_i \mu}{1 + \rho \mu}. \end{cases} \tag{40}$$

We define the  $\mu$ -sparsified version  $S_{\rho,\delta}^{C_i, \mu}$  of  $S_{\rho,\delta}$  as

$$S_{\rho,\delta}^{C_i, \mu} := A \hat{\Theta}_{C_i, \mu, \rho}^\dagger A^T + \delta I. \tag{41}$$

Notice that this matrix completely ignores some of the columns of  $A$  (and some of the edges of the graph). The dual regularization  $\delta I$  guarantees that the resulting matrix is non-singular, irrespective of the level of sparsification chosen. In this paper, we consider using inexact Newton directions produced by solving linear systems with matrix  $S_{\rho,\delta}^{C_i, \mu}$ , rather than  $S_{\rho,\delta}$ .

**Remark 1.** It is important to note that, in general, the sparsity pattern of the matrix  $S_{\rho,\delta}^{C_i, \mu}$  depends on the choice of the parameter  $C_i$  and on the partitioning  $(\mathcal{B}, \mathcal{N})$ . Indeed, when  $\mu$  is sufficiently small, we expect that

$$\left| \left\{ i \in \{1, \dots, n\} \text{ s.t. } (\Theta^{-1} + \rho I)_{ii}^{-1} \geq \frac{C_i \mu}{1 + \rho \mu} \right\} \right| = |\mathcal{B}|.$$

Let us now show how the algorithm is affected by the use of the proposed sparsified normal matrix. Notice that the results presented below depend strongly on two facts: the optimization problem involves an incidence matrix of a graph and thus the normal matrix is a Laplacian, with very desirable properties; the considered IPM employs primal–dual regularization.

We start by showing how much the normal matrix deviates from its sparsified counterpart.

**Theorem 6.** *The sparsification strategy in (40) produces a matrix which is close to the original  $S_{\rho,\delta}$ , in the sense that there exists a constant  $C_9$  independent of  $n$  such that*

$$\|S_{\rho,\delta} - S_{\rho,\delta}^{C_i, \mu}\| \leq C_9 \frac{C_i \mu}{1 + \rho \mu}.$$



**Proof.** We have that  $S_{\rho,\delta} - S_{\rho,\delta}^{C_t\mu} = AE^\mu A^T$  where  $E^\mu$  is the diagonal matrix defined as

$$E_{ii}^\mu := (\Theta^{-1} + \rho I)^{-1} - (\widehat{\Theta}_{C_t\mu,\rho}^\mu)_{ii} = \begin{cases} 0 & \text{if } (\Theta^{-1} + \rho I)_{ii}^{-1} \geq \frac{C_t\mu}{1+\rho\mu} \\ (\Theta^{-1} + \rho I)_{ii}^{-1} & \text{if } (\Theta^{-1} + \rho I)_{ii}^{-1} < \frac{C_t\mu}{1+\rho\mu}. \end{cases}$$

Hence we have  $\lambda_{\max}(E^\mu) \leq \frac{C_t\mu}{1+\rho\mu}$ . This follows using Lemma 4 and observing that

$$\frac{\mathbf{v}^T (S_{\rho,\delta} - S_{\rho,\delta}^{C_t\mu}) \mathbf{v}}{\mathbf{v}^T \mathbf{v}} = \frac{\mathbf{v}^T (AE^\mu A^T) \mathbf{v}}{\mathbf{v}^T \mathbf{v}} \leq \lambda_{\max}(E^\mu) \lambda_{\max}(AA^T). \quad \square$$

We now show that the condition number of both matrices,  $S_{\rho,\delta}$  and  $S_{\rho,\delta}^{C_t\mu}$ , is uniformly bounded. This is an important property when using iterative Krylov solvers to find the Newton directions.

**Lemma 6.** When  $\mu$  is sufficiently small, the condition numbers of  $S_{\rho,\delta}$  and  $S_{\rho,\delta}^{C_t\mu}$  satisfy

$$k_2(S_{\rho,\delta}) \in O\left(1 + \frac{1}{\delta(\rho + \mu)}\right) \quad \text{and} \quad k_2(S_{\rho,\delta}^{C_t\mu}) \in O\left(1 + \frac{1}{\delta(\rho + \mu)}\right).$$

**Proof.** The thesis follows from Lemma 4 and observing that for  $\mathbf{v} = A^T \mathbf{w}$

$$\begin{aligned} \delta &\leq \frac{\mathbf{w}^T S_{\rho,\delta} \mathbf{w}}{\mathbf{w}^T \mathbf{w}} \leq \delta + \frac{\mathbf{v}^T (\Theta^{-1} + \rho I)^{-1} \mathbf{v}}{\mathbf{v}^T \mathbf{v}} \frac{\mathbf{w}^T A A^T \mathbf{w}}{\mathbf{w}^T \mathbf{w}} \\ &\leq \delta + O\left(\frac{1}{\rho + \mu} (2 \max_{v \in V} \text{deg}(v))\right). \end{aligned}$$

A similar argument holds for  $S_{\rho,\delta}^{C_t\mu}$ .  $\square$

We now show that the solution of the sparsified linear system is “close” to the solution of the original one, and the bound depends on  $\mu$ . This result depends on the spectral distributions results shown in the previous section.

**Theorem 7.** For all  $\mathbf{v} \in \mathbb{R}^m$  we have that

$$(S_{\rho,\delta}^{C_t\mu})^{-1} \mathbf{v} = S_{\rho,\delta}^{-1} \mathbf{v} + \boldsymbol{\psi}$$

where  $\|\boldsymbol{\psi}\| \leq C_{10} \frac{C_t\mu}{\delta^2(1+\rho\mu)} \|\mathbf{v}\|$  where  $C_{10}$  is independent of  $n$ .

**Proof.** Using Lemma 5 and Theorem 6, we have that

$$\|I - S_{\rho,\delta}^{-1} S_{\rho,\delta}^{C_t\mu}\| \leq \|S_{\rho,\delta}^{-1}\| \|S_{\rho,\delta} - S_{\rho,\delta}^{C_t\mu}\| \leq C_{10} \frac{C_t\mu}{\delta(1 + \rho\mu)},$$

and hence

$$\begin{aligned} \|(S_{\rho,\delta}^{-1} - (S_{\rho,\delta}^{C_t\mu})^{-1}) \mathbf{v}\| &\leq \|(S_{\rho,\delta}^{C_t\mu})^{-1}\| \|S_{\rho,\delta}^{-1}\| \|S_{\rho,\delta} - S_{\rho,\delta}^{C_t\mu}\| \|\mathbf{v}\| \\ &\in O\left(\frac{C_t\mu}{\delta^2(1 + \rho\mu)} \|\mathbf{v}\|\right). \quad \square \end{aligned}$$

The next technical result is useful for the proof of Corollary 4.

**Lemma 7.**  $\|\bar{\xi}_p^j\|$  is uniformly bounded w.r.t. the IPM iterates  $j$ , i.e., there exists a constant  $C_{11} > 0$  independent of  $n$  such that for all  $j \in \mathbb{N}$

$$\|\bar{\xi}_p^j\| \leq C_{11} n$$

**Proof.** For the sake of notational clarity, we do not include the index  $j$  in the proof. To bound  $\|\bar{\xi}_p\|$ , consider the following estimate

$$\|\bar{\xi}_p\| \leq \|\xi_p\| + \|A\| \left( \|(\Theta^{-1} + \rho I)^{-1} X^{-1} \xi_{\mu,\sigma}\| + \|(\Theta^{-1} + \rho I)^{-1}\| \|\xi_d\| \right).$$

We already know the following estimates

$$\|\xi_p\| \leq \frac{\mu n}{\gamma_p} \leq \frac{\mu^0 n}{\gamma_p}, \quad \|\xi_d\| \leq \frac{\mu n}{\gamma_d} \leq \frac{\mu^0 n}{\gamma_d}, \quad \|(\Theta^{-1} + \rho I)^{-1}\| \leq \frac{1}{\rho}.$$

To estimate  $\|(\Theta^{-1} + \rho I)^{-1} X^{-1} \xi_{\mu,\sigma}\|$ , we proceed as in (35):

$$\|(\Theta^{-1} + \rho I)^{-1} X^{-1} \xi_{\mu,\sigma}\| = \|(\Theta^{-1} + \rho I)^{-1} (S\mathbf{e} - \sigma\mu X^{-1}\mathbf{e})\| \leq$$

$$\begin{aligned} &\leq \|(\Theta^{-1} + \rho I)^{-1/2}\| \|(\Theta^{-1} + \rho I)^{-1/2} X^{-1/2} S^{1/2}\| \\ &\quad (\|X^{1/2} S^{1/2} \mathbf{e}\| + \sigma\mu \|X^{-1/2} S^{-1/2} \mathbf{e}\|). \end{aligned}$$

It is straightforward to prove that

$$\|(\Theta^{-1} + \rho I)^{-1/2}\| \leq \frac{1}{\rho^{1/2}}, \quad \|(\Theta^{-1} + \rho I)^{-1/2} X^{-1/2} S^{1/2}\| \leq 1.$$

The remaining terms can be bounded using the properties of the neighbourhood

$$\|X^{1/2} S^{1/2} \mathbf{e}\| \leq \sqrt{\mu \gamma n}, \quad \sigma\mu \|X^{-1/2} S^{-1/2} \mathbf{e}\| \leq \sigma \sqrt{\frac{\mu n}{\gamma}}.$$

Since  $\mu \leq \mu^0$ , we deduce that  $\|\bar{\xi}_p\| \leq C_{11} n$ , for some positive constant  $C_{11}$ .  $\square$

Finally, we show that, for a small enough constant  $C_t$  independent of  $n$ , the inexactness introduced by the sparsification strategy satisfies the condition in (16). Therefore, an algorithm that includes such a sparsification strategy retains the polynomial complexity of the inexact IPM shown in Section 3.

**Corollary 4.** If in Algorithm 2 we generate the search directions using  $(S_{\rho,\delta}^{C_t\mu})^{-1}$  with  $C_t$  sufficiently small, i.e. if we compute the search directions using (11), (12) and (13) where  $S_{\rho,\delta}^{C_t\mu}$  substitutes  $S_{\rho,\delta}$ , then Algorithm 2 is convergent.

**Proof.** Using Theorem 7, we have

$$(S_{\rho,\delta}^{C_t\mu})^{-1} \bar{\xi}_p = S_{\rho,\delta}^{-1} \bar{\xi}_p + \boldsymbol{\psi}$$

where, using Lemma 7 we have

$$\|\boldsymbol{\psi}\| \leq C_{10} \frac{C_t\mu}{\delta^2(1 + \rho\mu)} \|\bar{\xi}_p\| \leq C_{10} C_{11} n \frac{C_t\mu}{\delta^2(1 + \rho\mu)} \quad (42)$$

for some constants  $C_{10}, C_{11} > 0$  independent of  $n$ . Hence

$$S_{\rho,\delta} (S_{\rho,\delta}^{C_t\mu})^{-1} \bar{\xi}_p = \bar{\xi}_p + S_{\rho,\delta} \boldsymbol{\psi} \stackrel{=dy}{\approx} \bar{\xi}_p.$$

Using Lemma 5, Assumption 1 and Eq. (42), we have that there exists a constant  $C_{12} > 0$  independent of  $n$  s.t.

$$\|S_{\rho,\delta} \boldsymbol{\psi}\| \leq C_{12} n \frac{C_t\mu}{\delta^2(1 + \rho\mu)}$$

The thesis follows using Eq. (16) and observing that

$$C_{12} n \frac{C_t\mu}{\delta^2(1 + \rho\mu)} \leq C_{\text{inexact}} \mathbf{x}^T \mathbf{s}$$

holds if

$$C_t < \frac{\delta^2(1 + \rho\mu) C_{\text{inexact}}}{C_{12}}. \quad \square$$

## 6. Numerical results

The proposed method is compared with Lemon (Library for Efficient Modelling and Optimization on Networks) (Kovacs, 2015), an extremely efficient C++ implementation of the network simplex method (Orlin, 1997), that has been shown to significantly outperform other popular implementations, like Cplex, see e.g. Castro and Nasini (2021), Zanetti and Gondzio (2023). The network simplex method has been shown (Schrieber et al., 2017) to be very competitive against other algorithms specifically developed for discrete OT, while remaining very robust and adaptable to many types of problems. Let us highlight that the other algorithms available in Lemon (cost scaling, capacity scaling, cycle cancelling) produced worse results than the network simplex.

All the computational tests discussed in this section are performed using a Dell PowerEdge R740 running Scientific Linux 7 with 4x Intel Gold 6234 3.3G, 8C/16T, 10.4GT/s, 24.75M Cache, Turbo, HT (130 W) DDR4-2933, with 500 GB of memory. The PS-IPM implementation

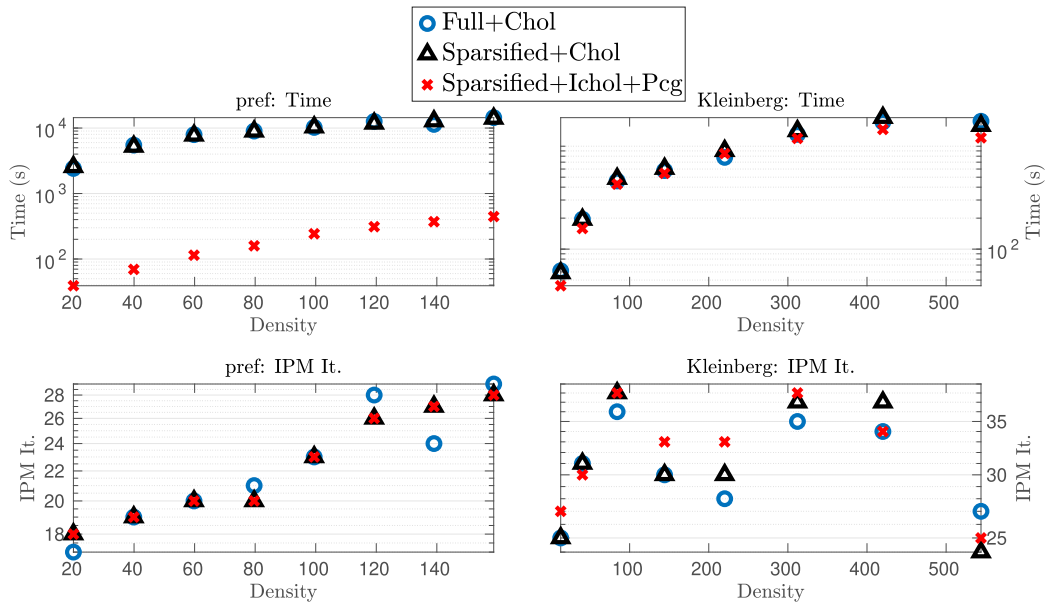


Fig. 1. Comparison of sparsified and full normal equations approach, using full Cholesky factorization or incomplete Cholesky as preconditioner for PCG, in terms of IPM iterations and computational time, for problems pref and kleinberg.

closely follows the one from Cipolla and Gondzio (2023) and is written in Matlab®. The software versions used for the numerical experiments are as follows: Matlab R2022a, Lemon 1.3.1 and GCC 4.8.5 as the C++ compiler.

We stop Algorithm 1, when

$$\|g - A^T y - s\|_\infty \leq R \cdot tol \wedge \|b - Ax\|_1 \leq R \cdot tol \wedge C_{x,s} \leq tol, \quad (43)$$

where

$$tol = 10^{-10}, \quad R := \max\{\|A\|_\infty, \|b\|_1, \|c\|_1\},$$

and

$$C_{x,s} := \max_i \{\min\{|(x_i, s_i)|, |x_i|, |s_i|\}\}.$$

Concerning the choice of the parameters in Algorithm 1, we set  $\sigma_r = 0.7$ . Moreover, to prevent wasting time on finding excessively accurate solutions in the early PPM sub-problems, we set  $\tau_1 = 10^{-4}$ , i.e. we use as inexactness criterion for the PPM method

$$\|r_k(x_{k+1}, y_{k+1})\| < 10^4 \sigma_r^k \min\{1, \|(x_{k+1}, y_{k+1}) - (x_k, y_k)\|\}.$$

Indeed, in our computational experience, we have found that driving the IPM solver to a high accuracy in the initial PPM iterations is unnecessary and usually leads to a significant deterioration of the overall performance.

Concerning Algorithm 2, we set as regularization parameters  $\rho = 10^{-4}$  and  $\delta = 10^{-6}$ . Moreover, in order to find the search direction, we employ a widely used predictor–corrector method (Mehrotra, 1992). This issue represents the main point where practical implementation deviates from the theory in order to gain computational efficiency and is a quite common procedure in the IPM community. We refer the interested reader to Gondzio (2012a, Sec. 4.4) and references therein for further explanation of such predictor–corrector technique.

Finally, concerning the test problems, in all the following experiments we generate the load vector  $\rho_1 - \rho_0$  in (1) randomly and such that the sum of its entries is zero (to guarantee feasibility of the optimization problem), with only 10% of them being nonzeros. Moreover, we fix the weight of each edge at 1.

### 6.1. Analysis of the sparsification strategy

In this section, we compare three possible solution strategies inside the PS-IPM: Cholesky factorization (using Matlab’s chol function)

applied to the full normal equations matrix (14); Cholesky factorization (always using Matlab’s chol function) applied to the sparsified matrix (41); preconditioned conjugate gradient (PCG) (using Matlab’s pcg function) applied to the sparsified matrix (41) with incomplete Cholesky preconditioner (computed using Matlab’s ichol function). More in particular, as sparsification parameter in (40) we use  $C_t = 0.4$ , ‘droptol’ =  $10^{-3}$  in ichol and ‘tol’ =  $10^{-1} \mu$  in pcg.

We test the above mentioned solution strategies on various instances generated with the CONTEST generator (Taylor & Higham, 2009); in particular, we considered the graphs pref, kleinberg, smallw and erdrey, with a fixed number of 100,000 nodes and different densities (i.e. average number of edges per node). Therefore, for these instances,  $m = |V| = 100,000$  and  $n = |E| = m \cdot \text{density}$ .

In the upper panels of Figs. 1 and 2 we report the computational time of the three approaches for various values of densities (chosen in relation to the properties of the graph), whereas in the lower panels we report the total number of IPM iterations. From the presented numerical results, it is clear that the sparsification strategy, in conjunction with the iterative solution of the linear systems, provides a clear advantage over the use of a direct factorization. As can be expected, the iterative method and the sparsification strategy become more advantageous when the size of the problem (number of edges) increases. On the other hand, it is important to note that the use of the sparsified Newton equations in conjunction with the full Cholesky factorization presents only limited advantages in terms of computational time when compared to the Cholesky factorization of the full Newton normal equation. This is the case because the resulting inexact IPM requires, generally, more iterations to converge (see lower panels of Figs. 1 and 2). Advantages of the proposed approach become clearer when the graphs are denser.

### 6.2. Results on randomly generated graphs

In this section, we compare the PS-IPM algorithm, using the sparsified normal equations matrix and the PCG, with the network simplex solver of Lemon. For PS-IPM we use the same parameters as proposed in Section 6.1. The graphs used in this section come from the generator developed in Viger and Latapy (2005) and already used for OT on graphs in Essid and Solomon (2018). This generator produces random connected graphs with a number of nodes varying from 1000 to 10,000,000 and degrees of each node in the range [1, 10], with an

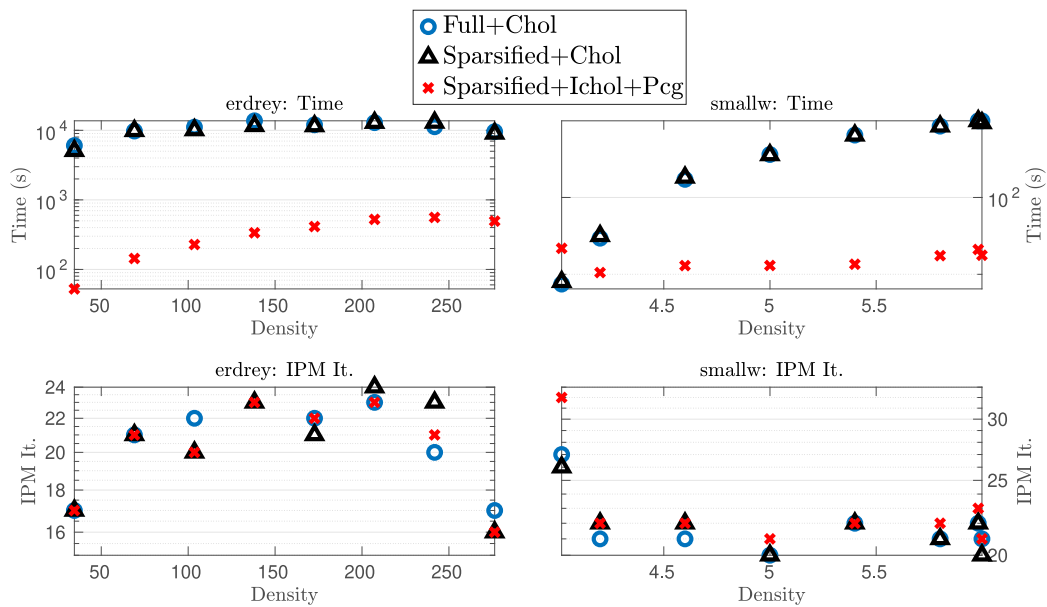


Fig. 2. Comparison of sparsified and full normal equations approach, using full Cholesky factorization or incomplete Cholesky as preconditioner for PCG, in terms of IPM iterations and computational time, for problems erdrey and smallw.

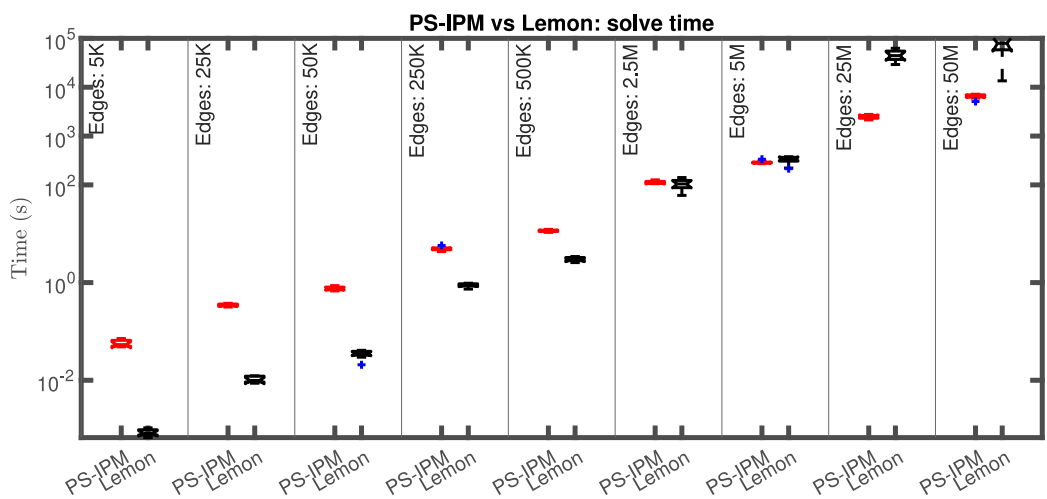


Fig. 3. Box Plots of the computational times of PS-IPM and Lemon, for randomly generated graphs. The red and black intervals show the spread of the measured computational times (red, on the left, is PS-IPM; black, on the right, is Lemon), while the (blue) crosses indicate the outliers. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

average of 5 edges per node. For each size, 10 graphs and load vectors are generated and tested. These parameters closely resemble the ones used in Essid and Solomon (2018).

Fig. 3 shows the comparison of the computational time between PS-IPM and Lemon: for each size of the problem (indicated by the total number of edges), we report the summary statistics of the execution times using Matlab's boxplot.

For small size problems, Lemon is the clear winner, by two orders of magnitude; however, as the size increases, the performance difference between the two methods reduces and for the largest instance considered, Lemon becomes one order of magnitude slower than PS-IPM.

Fig. 4 shows the average computational time against the number of edges (from 5000 to 50M) in a logarithmic scale, the corresponding regression lines and their slopes. From the computational results presented, we can estimate the practical time complexity of both methods. Recall that, in a log–log plot, polynomials of the type  $x^m$  appear as straight lines with slope  $m$ . Using linear regression, we can estimate

that the time taken by Lemon grows with exponent approximately 2.06, while the time taken by PS-IPM grows with exponent approximately 1.28, providing a considerable advantage for large sizes and confirming that, in practice, a polynomial-type convergence can be expected for the PS-IPM here introduced.

Finally, looking at the full set of results as reported in Fig. 3, let us mention the fact that the variance of the computational times over the 10 runs for a given problem size is smaller when using PS-IPM, especially for large sizes, indicating that the method is more robust and less dependent on the specific problem being solved. This is a very desirable property.

### 6.3. Results on SuiteSparse graphs

Results on randomly generated problems do not necessarily represent the ability of an optimization method to tackle problems coming from real world applications. Therefore, in this section, we show the results of applying PS-IPM and Lemon to some sparse graphs from the

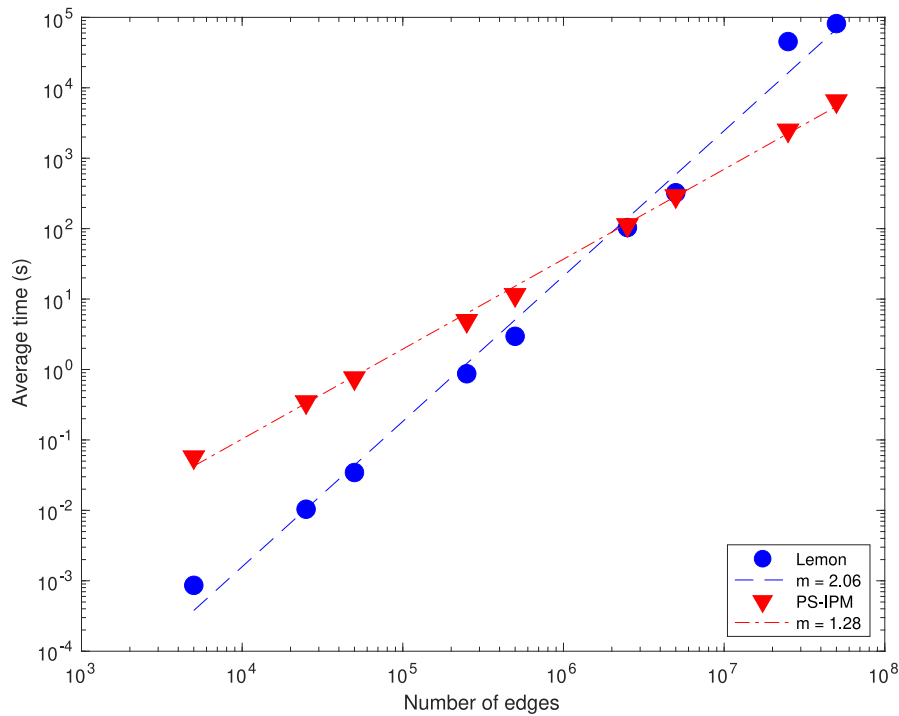


Fig. 4. Logarithmic plot of the computational time for randomly generated graphs. The time taken by Lemon (blue circles) grows as (number of edges)<sup>2.06</sup>; the time taken by PS-IPM (red triangles) grows as (number of edges)<sup>1.28</sup>. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

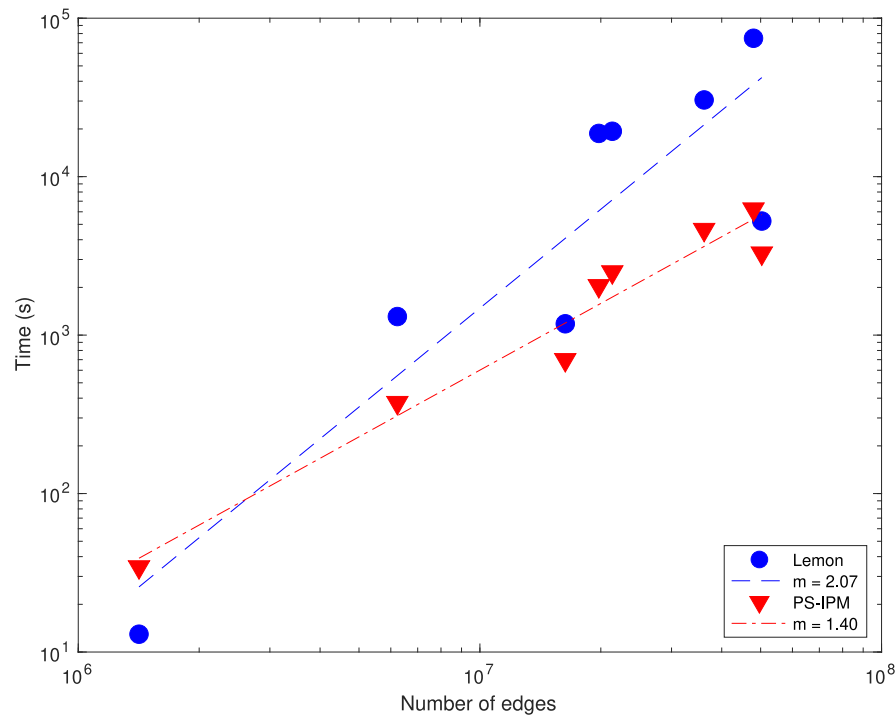


Fig. 5. Logarithmic plot of the computational time for the SuiteSparse problems. The time taken by Lemon (blue circles) grows as (number of edges)<sup>2.07</sup>; the time taken by PS-IPM (red triangles) grows as (number of edges)<sup>1.40</sup>. The instances are ordered as in Table 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

SuiteSparse matrix collection (Davis & Hu, 2011). The characteristics of the graphs considered are shown in Table 1: the number of nodes, edges and the average number of edges per node. All the graphs are undirected and connected. Due to the fact that the considered graphs are particularly sparse, in the numerical results presented in

this section, we solve the sparsified normal equations using the full Cholesky factorization.

Fig. 5 shows the computational times for the eight problems considered, using PS-IPM and Lemon. Apart from the problem nc2010, which represents a relatively small instance in our dataset, on all the

**Table 1**

Details of the graphs from the SuiteSparse matrix collection, ordered by increasing number of edges.

Name	Nodes	Edges	Density
nc2010	288,987	1,416,620	4.9
NACA0015	1,039,183	6,229,636	6.0
great-britain-osm	7,733,822	16,313,034	2.1
hugetric-00010	6,592,765	19,771,708	3.0
hugetric-00020	7,122,792	21,361,554	3.0
hugetrace-00010	12,057,441	36,164,358	3.0
hugetrace-00020	16,002,413	47,997,626	3.0
delaunay-n23	8,388,608	50,331,568	6.0

other problems PS-IPM consistently outperforms Lemon in terms of required computational time. In particular, for the problems `hugetric-00010`, `hugetric-00020`, `hugetrace-00010` and `hugetrace-00020`, which reach up to 16 million nodes and 48 million edges, PS-IPM is one order of magnitude faster than Lemon.

Notice that graphs of these sizes (and larger) appear in many modern practical applications, e.g. social networks, PageRank, analysis of rail/road networks, energy models, to mention a few.

Looking at the regression lines and their slopes, we notice that the time taken by Lemon grows with exponent approximately 2.07 while the time taken by PS-IPM grows with exponent approximately 1.40. These values are very close to the ones found previously for randomly generated graphs. The data of Fig. 5 however has a more erratic behaviour than the times shown in Fig. 4, because the properties of each graph considered are different and because we are not averaging over 10 different instances of each problem.

Let us highlight also that the time taken by Lemon seems to be more problem dependent, while PS-IPM looks more consistent and robust.

## 7. Conclusion

An efficient computational framework for the solution of Optimal Transport problems on graphs has been presented in this paper. Such framework relies on Proximal-Stabilized Interior Point Method and clever sparsifications of the normal Newton equations to compute the inexact search directions. The proposed technique is sound and polynomial convergence guarantee has been established for the inner inexact IPM. Extensive numerical experiments show that for large scale problems, a simple prototype Matlab implementation is able to outperform consistently a highly specialized and very efficient C++ implementation of the network simplex method.

We highlight also that Interior Point Methods are more easily parallelizable than simplex-like methods; for huge scale problems, for which high performance computing resources need to be used, the use of IPMs with proper parallelization may be the only viable strategy to solve these problems.

## References

- Ahuja, R. K., Magnanti, T. L., & Orlin, J. B. (1993). *Network flows*. Englewood Cliffs, NJ: Prentice-Hall.
- Altman, A., & Gondzio, J. (1999). Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization. *Optimization Methods & Software*, 11/12(1–4), 275–302. <http://dx.doi.org/10.1080/10556789908805754>.
- Armand, P., & Benoist, J. (2013). Uniform boundedness of the inverse of a Jacobian matrix arising in regularized interior-point methods. *Mathematical Programming*, 137(1–2, Ser. A), 587–592. <http://dx.doi.org/10.1007/s10107-011-0498-3>.
- Bellavia, S. (1998). Inexact interior-point method. *Journal of Optimization Theory and Applications*, 96(1), 109–121. <http://dx.doi.org/10.1023/A:1022663100715>.
- Castro, J., & Nasini, S. (2021). A specialized interior-point algorithm for huge minimum convex cost flows in bipartite networks. *European Journal of Operational Research*, 290(3), 857–869. <http://dx.doi.org/10.1016/j.ejor.2020.10.027>.
- Cipolla, S., & Gondzio, J. (2023). Proximal stabilized interior point methods and low-frequency-update preconditioning techniques. *Journal of Optimization Theory and Applications*, 197, 1061–1103. <http://dx.doi.org/10.1007/s10957-023-02194-4>.

- Cornelis, J., & Vanroose, W. (2021). Convergence analysis of a regularized inexact interior-point method for linear programming problems. <http://dx.doi.org/10.48550/ARXIV.2105.01333>, [math.OC].
- Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. In C. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Weinberger (Eds.), *Vol. 26, Adv. in neural information processing systems*. Curran Associates, Inc..
- Daitch, S. L., & Spielman, D. A. (2008). Faster approximate lossy generalized flow via interior point algorithms. In *Proceedings of the fortieth annual ACM symposium on theory of computing STOC '08*, (pp. 451–460). New York, NY, USA: Association for Computing Machinery, <http://dx.doi.org/10.1145/1374376.1374441>.
- Davis, T. A., & Hu, Y. (2011). The university of florida sparse matrix collection. *ACM Transactions on Mathematical Software*, 38(1), Art. 1, 25. <http://dx.doi.org/10.1145/2049662.2049663>.
- Dembo, R. S., Eisenstat, S. C., & Steihaug, T. (1982). Inexact Newton methods. *SIAM Journal on Numerical Analysis*, 19(2), 400–408. <http://dx.doi.org/10.1137/0719025>.
- Essid, M., & Solomon, J. (2018). Quadratically regularized optimal transport on graphs. *SIAM Journal on Scientific Computing*, 40(4), A1961–A1986. <http://dx.doi.org/10.1137/17M1132665>.
- Facca, E., & Benzi, M. (2021). Fast iterative solution of the optimal transport problem on graphs. *SIAM Journal on Scientific Computing*, 43(3), A2295–A2319. <http://dx.doi.org/10.1137/20M137015X>.
- Frangioni, A., & Gentile, C. (2004). New preconditioners for KKT systems of network flow problems. *SIAM Journal on Optimization*, 14(3), 894–913. <http://dx.doi.org/10.1137/S105262340240519X>.
- Frangioni, A., & Serra Capizzano, S. (2001). Spectral analysis of (sequences of) graph matrices. *SIAM Journal of Mathematical Analysis*, 23(2), 339–348. <http://dx.doi.org/10.1137/S089547989935366X>.
- Friedlander, M. P., & Orban, D. (2012). A primal-dual regularized interior-point method for convex quadratic programs. *Mathematical Programming Computation*, 4(1), 71–107. <http://dx.doi.org/10.1007/s12532-012-0035-2>.
- Gondzio, J. (2012a). Interior point methods 25 years later. *European Journal of Operational Research*, 218(3), 587–601. <http://dx.doi.org/10.1016/j.ejor.2011.09.017>.
- Gondzio, J. (2012b). Matrix-free interior point method. *Computational Optimization and Applications*, 51, 457–480. <http://dx.doi.org/10.1007/s10589-010-9361-3>.
- Gottschlich, C., & Schuhmacher, D. (2014). The shortlist method for fast computation of the Earth mover's distance and finding optimal solutions to transportation problems. *PLoS ONE*, 9(10), Article e110214. <http://dx.doi.org/10.1371/journal.pone.0110214>.
- Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., & Courville, A. (2017). Improved training of wasserstein GANs. In *Proceedings of the 31st international conference on neural information processing systems NIPS '17*, (pp. 5769–5779). Red Hook, NY, USA: Curran Associates Inc..
- Haker, S., Zhu, L., Tannenbaum, A., & Angenent, S. (2004). Optimal mass transport for registration and warping. *International Journal of Computer Vision*, 60(3), 225–240. <http://dx.doi.org/10.1023/B:VISI.0000036836.66311.97>.
- Kantorovich, L. (1958). On the translocation of masses. *Management Science*, 5, 1–4.
- Kelly, C. T. (1995). *Iterative methods for linear and nonlinear equations*. Philadelphia, USA: SIAM.
- Kojima, M., Megiddo, N., & Mizuno, S. (1993). A primal-dual infeasible interior-point algorithm for linear programming. *Mathematical Programming*, 61(3, Ser. A), 263–280. <http://dx.doi.org/10.1007/BF01582151>.
- Korzak, J. (2000). Convergence analysis of inexact infeasible-interior-point algorithms for solving linear programming problems. *SIAM Journal on Optimization*, 11(1), 133–148. <http://dx.doi.org/10.1137/S1052623497329993>.
- Kovacs, P. (2015). Minimum-cost flow algorithms: an experimental evaluation. *Optimization Methods & Software*, 30(1), 94–127.
- Ling, H., & Okada, K. (2007). An efficient Earth mover's distance algorithm for robust histogram comparison. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(5), 840–853. <http://dx.doi.org/10.1109/TPAMI.2007.1058>.
- Luque, F. J. (1984). Asymptotic convergence analysis of the proximal point algorithm. *SIAM Journal on Control and Optimization*, 22(2), 277–293. <http://dx.doi.org/10.1137/0322019>.
- Manteuffel, T. A. (1980). An incomplete factorization technique for positive definite linear systems. *Mathematics of Computation*, 34(150), 473–497. <http://dx.doi.org/10.1090/S0025-5718-1980-0559197-0>.
- Mehrotra, S. (1992). On the implementation of a primal-dual interior-point methods. *SIAM Journal on Optimization*, 2(4), 575–601. <http://dx.doi.org/10.1137/0802028>.
- Mérigot, Q. (2011). A multiscale approach to optimal transport. *Computer Graphics Forum*, 30(5), 1583–1592. <http://dx.doi.org/10.1111/j.1467-8659.2011.02032.x>.
- Métivier, L., Brossier, R., Mérigot, Q., Oudet, E., & Virieux, J. (2016). An optimal transport approach for seismic tomography: application to 3D full waveform inversion. *Inverse Problems*, 32(11), Article 115008. <http://dx.doi.org/10.1088/0266-5611/32/11/115008>.
- Natale, A., & Todeschi, G. (2021). Computation of optimal transport with finite volumes. *ESAIM: M2AN*, 55(5), 1847–1871.
- Ni, C.-C., Lin, Y.-Y., Luo, F., & Gao, J. (2019). Community detection on networks with Ricci flow. *Scientific Reports*, 9(1), 9984. <http://dx.doi.org/10.1038/s41598-019-46380-9>.



- Orlin, J. B. (1997). A polynomial time primal network simplex algorithm for minimum cost flows. *Mathematical Programming*, 78(2), 109–129. <http://dx.doi.org/10.1007/BF02614365>.
- Peyre, G., & Cuturi, M. (2019). Computational optimal transport: with applications to data science. *Foundations Trends and Trends in Machine Learning*, 11, 355–607. <http://dx.doi.org/10.1561/22000000073>.
- Portugal, L. F., Resende, M. G. C., Veiga, G., & Júdice, J. J. (2000). A truncated primal-infeasible dual-feasible network interior point method. *Networks*, 35(2), 91–108.
- Resende, M. G., & Veiga, G. (1993). An implementation of the dual affine scaling algorithm for minimum-cost flow on bipartite uncapacitated networks. *SIAM Journal on Optimization*, 3(3), 516–537. <http://dx.doi.org/10.1137/0803025>.
- Rockafellar, R. T. (1970). Monotone operators associated with saddle-functions and minimax problems. In *Nonlinear functional analysis (proc. sympos. pure math., Vol. XVIII, Part I, Chicago, Ill., 1968)* (pp. 241–250). Amer. Math. Soc., Providence, R.I.
- Saunders, M., & Tomlin, J. A. (1996). Solving regularized linear programs using barrier methods and KKT systems. *Technical Report SOL 96-4*, Systems Optimization Laboratory, Dept. of Operations Research, Stanford University, Stanford, CA 94305, USA.
- Schrieber, J., Schuhmacher, D., & Gottschlich, C. (2017). DOTmark - a benchmark for discrete optimal transport. *IEEE Access*, 5, 271–282. <http://dx.doi.org/10.1109/ACCESS.2016.2639065>.
- Spielman, D. A., & Teng, S.-H. (2014). Nearly linear time algorithms for preconditioning and solving symmetric, diagonally dominant linear systems. *SIAM Journal of Mathematical Analysis*, 35(3), 835–885. <http://dx.doi.org/10.1137/090771430>.
- Taylor, A., & Higham, D. J. (2009). CONTEST: A controllable test matrix toolbox for MATLAB. *ACM Transactions on Mathematical Software*, 35(4), <http://dx.doi.org/10.1145/1462173.1462175>.
- Viger, F., & Latapy, M. (2005). Efficient and simple generation of random simple connected graphs with prescribed degree sequence. In *International computing and combinatorics conference* (pp. 440–449). Springer.
- Vishnoi, N. K. (2012).  $Lx = b$  Laplacian solvers and their algorithmic applications. *Foundations and Trends in Theoretical Computer Science*, 8(1–2), <http://dx.doi.org/10.1561/04000000054>, front matter, 1–141 (2013).
- Wijesinghe, J., & Chen, P. (2022). Matrix-free interior point methods for point set matching problems. [arXiv:2202.09763](https://arxiv.org/abs/2202.09763), [math.OC].
- Wright, S. J. (1997). *Primal-dual interior-point methods*. Society for Industrial and Applied Mathematics, Philadelphia, PA, [http://dx.doi.org/10.1137/1.9781611971453,xx+289](http://dx.doi.org/10.1137/1.9781611971453.xx+289).
- Zanetti, F., & Gondzio, J. (2023). An interior-point-inspired algorithm for linear programs arising in discrete optimal transport. *INFORMS Journal on Computing*, 35(5), <http://dx.doi.org/10.1287/ijoc.2022.0184>.