STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–20.
Published online in International Academic Press (www.IAPress.org)

New search direction based on a class of parametric kernel functions with a Full Newton step Infeasible $\mathcal{O}(nL)$ Interior point Methods for Linear Optimization

S. Bouali*

Department of Data, Content and Knowledge Engineering,
Laboratory of Computer Sciences, Data Sciences and Artificial Intelligence, Morocco

Abstract In this paper, which is motivated by the work of Roos [7] (SIAM J. Optim. 16(4):1110-1136, 2006), we examine a new search direction derived from a family of parametric kernel functions for IIPM algorithms. The main iteration of the algorithm is composed of one feasibility step followed by several centrality steps. The neighborhood of Newton process is more wider using a sharper quadratic convergence results. The algorithm has polynomial complexity and matches the best known iteration bound based on centrality steps. Furthermore, the numerical experiments demonstrate the efficiency of this class of functions, providing increased flexibility in selecting the search direction for solving problems.

Keywords Linear programming, Infeasible interior-point method, Full-Newton step, Polynomial complexity, Kernel function

DOI: 10.19139/soic-2310-5070-3076

1. Introduction

In this paper, we consider the linear optimization (LO) problem in the standard form :

$$(\mathrm{P}) \quad \min \{c^T x \ : \ Ax = b, \ x \geq 0\},$$

with its dual problem

$$(\mathrm{D}) \quad \max \{b^T y \ : \ A^T y + s = c, \ s \geq 0\},$$

where $c, x, s \in \mathbb{R}^n$, $b, y \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$ is of full row rank.

Due to the efficiency from a computational point of view, the use of Interior Point Method based on the kernel function becomes more attractive.

Since the appearance of the Roos's paper [7], much works regarding this special topic of Infeasible interior point algorithms for linear optimization, has been done. These works has essentially focused on the search direction by trying to modify the KKT system of the primal and dual problem, especially the third equation namely the centering equation of the system below (2) and consequently analyzing the algorithm's convergence[1],[6]. Some works have proposed an algorithms for Infeasible Interior point method IIPM with a new search directions based on some specific kernel functions [10],[4],[3]. Our algorithm is an extension of the original work of Roos [7] and Liu [10] in which no line search is needed, it uses a full Newton step instead of a damped step.

*Correspondence to: Samir Bouali (Email: samir.bouali34@gmail.com). Laboratory of Computer Sciences, Data Sciences and Artificial Intelligence, School of Information Sciences. Avenue Allal El Fassi, Al Irfane City, Mailbox: 6204, Rabat Institutes Morocco.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

In this paper, we propose a new search direction introduced by a class of parametric kernel functions whose parameter p belongs to the interval $[0, 1]$, giving a more slightly wider neighborhood of quadratic convergence for feasibility steps which guarantees that the proximity-measure will be smallest than a threshold τ in a finite number of iterations (do not exceed 4 in our case). The direction used in our work is more natural and better intuitively. The analysis of the convergence and complexity of the algorithm follows. For the survey of IIPM we refer to the introduction of Roos [7].

The paper is organized as follows. In Section 2 we present some useful properties in the analysis of feasible IPM which will be exploited in the analysis of our IIPM. In Section 3 we present our full-Newton step IIPM. Each main step of the method consists of a feasibility step and several centering steps. For the centering steps we exploit a sharper quadratic convergence result which is done in a slightly wider neighborhood for the feasibility steps. Section 4 is devoted to the analysis of our feasibility step. In Section 5 we obtain the complexity result of our IIPM algorithm. In section 6, we present the numerical results of our algorithm. Finally we give some concluding remarks in Section 7.

2. Feasible Newton step for IPMs

In our analysis we recall some useful properties of central path and feasible full Newton step. For more details, we refer to [8],[9]. To solve (P) and (D), one needs to find a solution of the following system of equations.

$$\begin{aligned} Ax &= b, \quad x \geq 0, \\ A^T y + s &= c, \quad s \geq 0, \\ xs &= 0, \end{aligned} \tag{1}$$

In these so-called optimality conditions the first two constraints represent primal and dual feasibility, whereas the last equation is the so-called complementary condition. The nonnegativity constraints in the feasibility conditions make the problem already nontrivial: only iterative methods can find solutions of linear systems involving inequality constraints. The complementary condition is nonlinear, which makes it extra hard to solve this system.

2.1. Central Path

IPMs replace the complementarity condition by the so-called centering condition $xs = \mu e$, where μ may be any positive number. This yields the system

$$\begin{aligned} Ax &= b, \quad x \geq 0, \\ A^T y + s &= c, \quad s \geq 0, \\ xs &= \mu e. \end{aligned} \tag{2}$$

Surprisingly enough, if this system has a solution for some $\mu > 0$, then a solution exists for every $\mu > 0$, and this solution is unique. This happens if and only if problems (P) and (D) satisfy the interior-point condition (IPC); i.e., if (P) has a feasible solution $x > 0$ and (D) has a solution (y, s) with $s > 0$. If the IPC is satisfied, then the solution of (2) is denoted by $(x(\mu), y(\mu), s(\mu))$ and is called the μ-center of (P) and (D). The set of all ì-centers forms a path, which is called the central path. As μ goes to zero, $(x(\mu), y(\mu), s(\mu))$ converge to optimal solutions of problems (P) and (D). Of course, the system (2) is still hard to solve, but by applying Newtons method one can easily find approximate solutions.

2.2. *Properties of the Newton step*

We proceed by describing Newtons method for solving (3), with μ fixed. Given any primal feasible $x > 0$, dual feasible y and $s > 0$, we want to find displacements Δx, Δy and Δs such that

$$\begin{aligned} A(x + \Delta x) &= b, \\ A^T(y + \Delta y) + (s + \Delta s) &= c, \\ (x + \Delta x)(s + \Delta s) &= \mu e. \end{aligned}$$

According to Newton's method for solving nonlinear equations, we obtain the linear system in the search directions Δx, Δy and Δs:

$$\begin{aligned} A\Delta x &= b - Ax, \\ A^T\Delta y + \Delta s &= c - A^T y - s, \\ x\Delta s + s\Delta x &= \mu e - xs. \end{aligned} \tag{3}$$

Since A has full row rank, and since the vectors x and s are positive, one may easily verify that the coefficient matrix in the linear system (3) is nonsingular. Hence, this system uniquely defines the search directions Δx, Δy and Δs. These search directions are used in all existing primal-dual (feasible and infeasible) IPMs.

If x is primal feasible and (y, s) is dual feasible pair, then $b - Ax = 0$ and $c - A^T y - s = 0$, whence the above system reduces to

$$\begin{aligned} A\Delta x &= 0, \\ A^T\Delta y + \Delta s &= 0, \\ x\Delta s + s\Delta x &= \mu e - xs, \end{aligned} \tag{4}$$

which gives the usual search directions for feasible primal-dual IPMs. Then The new iterates are given by

$$\begin{aligned} x^+ &= x + \Delta x, \\ y^+ &= y + \Delta y, \\ s^+ &= s + \Delta s. \end{aligned}$$

An important observation is that Δx lies in the null space of A, whereas Δs belongs to the row space of A. This implies that Δx and Δs are orthogonal, i.e., $\Delta x^T \Delta s = 0$. As a consequence, we have the important property that, after a full-Newton step, the duality gap assumes the same value as at the μ-centers, namely $n\mu$.

Lemma 1. *(See [8], Lemma $II.47$) After a primal-dual Newton step, one has $(x^+)^T s^+ = n\mu$.*

We measure proximity of iterates (x, y, s) to the μ-center $(x(\mu), y(\mu), s(\mu))$ by the quantity $\delta(x, s; \mu)$, which is defined as follows:

$$\delta(x, s; \mu) := \delta(v) := \frac{1}{2}\|v - v^{-1}\|, \quad \text{where } v := \sqrt{\frac{xs}{\mu}}. \tag{5}$$

In the analysis of the algorithm, the effect of the proximity $\delta(x, s; \mu)$ on a full-Newton step targeting the μ-center of (P) and (D), will be essential.

We recall the following interesting Theorem which implies that the Newton process is locally quadratically convergent. This property has been crucial in the analysis in many papers as [8], [10], [6].

Theorem 2. *(See [8], Theorem II.52) If $\delta(x, s; \mu) < 1$, then*

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

The quadratic convergence can be also obtained by using a tighter upper bound for $\delta(x^+, s^+; \mu)$, which provides a slightly wider neighborhood for the feasibility step of our IIPM. We can now deduce the following trivial corollary which we state without proof.

Corollary 3. *If* $\delta(x, s; \mu) \leq 1/\sqrt[4]{2}$, *then* $\delta(x^+, s^+; \mu) \leq \delta^2$.

3. Infeasible full-Newton step for IIPM

In the case of an infeasible method, we call the triple (x, y, s) an ε-optimal solution of (P) and (D) if the 2-norms of the residual vectors $b - Ax$ and $c - A^T y - s$ do not exceed ε, and if the duality gap satisfies $x^T s \leq \varepsilon$. In this section, we present an infeasible-start algorithm that generates an ε-optimal solution of (P) and (D), if it exists, or establishes that no such solution exists.

3.1. Perturbed Problems

At the beginning, we choose arbitrarily $x^0 > 0$ and (y^0, s^0) with $s^0 > 0$ such that $x^0 s^0 = \mu^0 e$ for some positive number μ^0. We denote the initial values of the primal and dual residuals r_b^0 and r_c^0 respectively as

$$
\begin{aligned}
r_b^0 &= b - Ax^0, \\
r_c^0 &= c - A^T y^0 - s^0.
\end{aligned}
$$

For any ν such that $0 < \nu \leq 1$, we consider the following perturbed problem (P_ν), defined by

$$
(\mathrm{P}_\nu) \quad \min\{(c - \nu r_c^0)^T x \ : \ Ax = b - \nu r_b^0 \ , \ \ x \geq 0\},
$$

and its dual problem (D_ν), which is given by

$$
(\mathrm{D}_\nu) \quad \max\{(b - \nu r_b^0)^T y \ : \ A^T y + s = c - \nu r_c^0 \ , \ \ s \geq 0\},
$$

We note that if $\nu = 1$ then $x = x^0$ yields a strictly feasible solution of (P_ν), and $(y, s) = (y^0, s^0)$ a strictly feasible dual pair solution of (D_ν). We deduce that if $\nu = 1$ then (P_ν) and (D_ν) satisfy the IPC.

Lemma 4. *(See [7], Lemma 1.1]) The original problems* (P) *and* (D) *are feasible if and only if, for each* ν *satisfying* $0 < \nu \leq 1$, *the perturbed problems* (P_ν) *and* (D_ν) *satisfy the IPC.*

In the sequel, we assume that (P) and (D) are feasible.

3.2. Central Path of the Perturbed Problems

Let (P) and (D) be feasible and $0 < \nu \leq 1$. Then, Lemma (4) implies that the perturbed problems (P_ν) and (D_ν) satisfy the IPC; hence, their central paths exist. This means that the system

$$
\begin{aligned}
Ax &= b - \nu r_b^0, \quad x \geq 0, && (6)\\
A^T y + s &= c - \nu r_c^0, \quad s \geq 0, && (7)\\
xs &= \mu e,
\end{aligned}
$$

has a unique solution for every $\mu > 0$. This unique solution is denoted by $(x(\mu, \nu), y(\mu, \nu), s(\mu, \nu))$ and is the μ-center of the perturbed problems (P_ν) and (D_ν). In the sequel, the parameters μ and ν always satisfy the relation $\mu = \nu \mu^0$.
Note that since $x^0 s^0 = \mu^0 e$, x^0 is the μ^0-center of the perturbed problem (P_1) and (y^0, s^0) the μ^0-center of (D_1). In other words, $(x(\mu^0, 1), y(\mu^0, 1), s(\mu^0, 1)) = (x^0, y^0, s^0)$.

3.3. Description of the Algorithm

It is well known that the efficiency of algorithm is measured by the total number of inner iterations which is referred to as the iteration complexity of the algorithm. The best known iteration bound for IIPMs was first obtained by Mizuno [5]

$$\mathcal{O}\left(n \log \frac{\max\{(x^0)^T s^0, \|b - Ax^0\|, \|c - A^T y^0 - s^0\|\}}{\varepsilon}\right).$$

Up to a constant, this bound was slightly improved by Roos [7] and then by Gu et al. [6].
At the beginning, we specify our initial iterate (x^0, y^0, s^0). As usual in infeasible IPMs, we assume that the initial iterates are designed as follows :

$$x^0 = s^0 = \zeta e, \quad y^0 = 0, \quad \mu^0 = \zeta^2,$$

where e is the all-one vector of length n, μ^0 is the initial dual gap and $\zeta > 0$ is such that

$$\|x^* + s^*\|_\infty \leq \zeta,$$

for some optimal solution (x^*, y^*, s^*) of (P) and (D).
At the start of the algorithm, we have initially $\delta(x, s; \mu) = 0$, since if $\nu = 1$ and $\mu = \mu^0$, then $x = x^0$ is the μ-center of the perturbed problem (P_ν) and $(y, s) = (y^0, s^0)$ is the μ-center of the perturbed problem (D_ν). In the sequel, we assume that, at the start of each iteration, just before the feasibility step, $\delta(x, s; \mu)$ is smaller than or equal to a threshold value $\tau > 0$ which is ensured for the first iteration.
Now, we describe one (main) iteration of our algorithm. Suppose that, for some $\mu \in (0, \mu^0]$, we have (x, y, s) satisfying the feasibility conditions (6) and (7) with $\nu = \mu/\mu^0$ and such that $x^T s = n\mu$ and $\delta(x, s; \mu) \leq \tau$. We reduce μ to $\mu^+ = (1 - \theta)\mu$, with $\theta \in (0, 1)$, and find a new iterate (x^+, y^+, s^+) that satisfies (6) and (7), with ν replaced by $\nu^+ = (1 - \theta)\nu = \mu^+/\mu^0$, and such that $(x^+)^T s^+ = n\mu^+$ and $\delta(x^+, s^+; \mu^+) \leq \tau$.
To be more precise, this is achieved as follows. Each main iteration consists of a feasibility step and a few centering steps. The feasibility step serves to get an iterate (x^f, y^f, s^f) that is strictly feasible for (P_{ν^+}) and (D_{ν^+}) and close to their μ^+-center $(x(\nu^+), y(\nu^+), s(\nu^+))$. In fact, the feasibility step is designed in such a way that $\delta(x^f, s^f; \mu^f) \leq 1/\sqrt[4]{2}$, i.e., (x^f, y^f, s^f) belongs to the quadratic convergence neighborhood with respect to the μ^+-center of (P_{ν^+}) and (D_{ν^+}). Then we can easily get an iterate (x^+, y^+, s^+) that is strictly feasible for (P_{ν^+}) and (D_{ν^+}) and such that $(x^+)^T s^+ = n\mu^+$ and $\delta(x^+, s^+; \mu^+) \leq \tau$, just by performing a few centering steps starting from (x^f, y^f, s^f) and targeting the μ^+-center of (P_{ν^+}) and (D_{ν^+}).
In what follows, we describe the feasibility step in more detail. Suppose that we have a strictly feasible iterate (x, y, s) for (P_ν) and (D_ν). This means that (x, y, s) satisfies (6) and (7), with $\nu = \mu/\mu^0$. We need displacements $\Delta^f x$, $\Delta^f y$, $\Delta^f s$ such that

$$\begin{aligned} x^f &= x + \Delta^f x, \\ y^f &= y + \Delta^f y, \\ s^f &= s + \Delta^f s, \end{aligned}$$

are feasible for (P_{ν^+}) and (D_{ν^+}). One may verify easily that (x^f, y^f, s^f) satisfies (6) and (7), with ν replaced by $\nu^+ = (1 - \theta)\nu$, only if the first two equations in the following system are satisfied:

$$A\Delta^f x = \theta\nu r_b^0, \tag{8}$$
$$A^T \Delta^f y + \Delta^f s = \theta\nu r_c^0, \tag{9}$$
$$s\Delta^f x + x\Delta^f s = \mu e - xs, \tag{10}$$

We conclude that, after the feasibility step, the iterate satisfies the affine equations (6) and (7), with $\nu = \nu^+$. The hard part in the analysis is to guarantee that x^f and s^f are positive and satisfy $\delta(x^f, s^f; \mu^+) \leq 1/\sqrt[4]{2}$. After the feasibility step, we perform a few centering steps in order to get iterate (x^+, y^+, s^+) which satisfies

$(x^+)^T s^+ = n\mu^+$ and $\delta(x^+, s^+; \mu^+) \leq \tau$. By using Corollary (3), the required number of centering steps can be obtained easily. Indeed, assuming $\delta = \delta(x^f, s^f; \mu^+) \leq 1/\sqrt[4]{2}$, after k centering steps we will have iterates (x^+, y^+, s^+) that are still feasible for (P_{ν^+}) and (D_{ν^+}) and satisfy

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

From this, one deduces easily that $\delta(x^+, s^+; \mu^+) \leq \tau$ holds after at most

$$2 + \left\lceil \log_2 \left(\log_2 \frac{1}{\tau} \right) \right\rceil \tag{11}$$

centering steps.

We give below a more formal description of the algorithm as follows

Input :
- parameter p in $[0, 1]$;
- bound parameter ζ ;
- threshold parameter $\tau > 0$;
- accuracy parameter $\varepsilon > 0$;
- barrier update parameter θ in $]0, 1[$.

begin
- $x := \zeta e$; $y := 0$;$s := \zeta e$; $\nu := 1$;
- **while** $\max\{x^T s,\ \|b - Ax\|,\ \|c - A^T y - s\|\} \geq \varepsilon$ **do**
  - **begin**
  - feasibility step $(x, y, s) := (x, y, s) + (\Delta^f x, \Delta^f y, \Delta^f s)$;
  - μ-update: $\mu := (1 - \theta)\mu$;
  - centrality steps:
  - **while** $\delta(x, s; \mu) > \tau$ **do**
  - $(x, y, s) := (x, y, s) + (\Delta x, \Delta y, \Delta s)$;
  - **end while**
- **end**
- **end while**

end

Algorithm 1: Primal-Dual Infeasible IPMs Algorithm

Now we introduce the definition of a kernel function.

Definition 5. *We call $\psi : (0, \infty) \to [0, \infty)$ a kernel function if ψ is twice differentiable and the following conditions are satisfied*

(i) $\psi'(1) = \psi(1) = 0$,
(ii) $\psi''(t) > 0$ *for all* $t > 0$,

We define

$$\overline{A} = AV^{-1}X, \quad V = diag(v), \quad X = diag(x),$$
$$d_x^f = \frac{v\Delta^f x}{x}, \quad d_s^f = \frac{v\Delta^f s}{s} \tag{12}$$

The system (8)-(10) which defines the search directions $\Delta^f x$, $\Delta^f y$ and $\Delta^f s$, can be expressed in terms of the scaled directions d_x^f and d_s^f as follows:

$$\begin{aligned} \overline{A}d_x^f &= \theta\nu r_b^0, \\ \overline{A}^T\frac{\Delta^f y}{\mu} + d_s^f &= \theta\nu v s^{-1} r_c^0, \\ d_x^f + d_s^f &= v^{-1} - v, \end{aligned} \tag{13}$$

It is clear that the right-hand side of the equation (13) is the negative gradient direction of the logarithmic barrier function :

$$\Psi(v) := \sum_{i=1}^{n} \psi(v_i), \quad v_i = \sqrt{\frac{x_i s_i}{\mu}}$$

whose kernel function is

$$\psi(t) = \frac{t^2-1}{2} - \log(t).$$

In this paper we make a slight modification of the standard Newton direction. The new system is then defined as follows

$$\begin{aligned} \overline{A}d_x^f &= \theta\nu r_b^0, \\ \overline{A}^T\frac{\Delta^f y}{\mu} + d_s^f &= \theta\nu v s^{-1} r_c^0, \\ d_x^f + d_s^f &= -\nabla\Psi(v). \end{aligned} \tag{14}$$

where our kernel function of Ψ is given by

$$\psi_p(t) = \begin{cases} \frac{t^{1+p}-1}{1+p} + \frac{t^{1-q}-1}{q-1} & p,q \in]0,1[\quad p+q=1 \\ \frac{t^2-1}{2} - t + 1 & p = 1. \end{cases} \tag{15}$$

According to the definition (5), $\psi(t)$ is obviously a kernel function.

Since $\psi_p^{'}(t) = t^p - t^{-q}$, equation (14) can be rewritten as

$$d_x^f + d_s^f = v^{-q} - v^p. \tag{16}$$

In the sequel, the feasibility step will be based on the last equation. we can now define the following proximity measure induced by our kernel function

$$\sigma(v) := \frac{1}{2}\left\|\nabla\Psi(v)\right\| = \frac{1}{2}\left\|v^{-q} - v^p\right\|. \tag{17}$$

In fact we observe that $\sigma(v) = 0$ if and only if $v = e$, thus $\sigma(v) = 0$ is also an appropriate proximity measure. Later we prove that this proximity is smaller than the one induced by the classical logarithmic barrier function (5).

3.4. *Some technical lemmas*

The following lemmas will be useful for our analysis.

Lemma 6. *For any $t > 0$, one has*

$$\left|t^{-\frac{p}{2}} - t^{\frac{p}{2}}\right| \leq \frac{1}{2}\left|t^{-1} - t\right|, \quad p \in [0,1].$$

Proof
For t in $[0,1]$, and by defining the following functions :

$$\rho(t,p) := \frac{t^{-1}}{2} + t^{\frac{p}{2}} \quad \text{and} \quad \varrho(t,p) := \frac{t}{2} + t^{-\frac{p}{2}},$$

we get :

$$\frac{1}{2}\left|t^{-1}-t\right| - \left|t^{-\frac{p}{2}} - t^{\frac{p}{2}}\right| = \frac{1}{2}\left(t^{-1}-t\right) - \left(t^{-\frac{p}{2}} - t^{\frac{p}{2}}\right) = \rho(t,p) - \varrho(t,p).$$

The right term in the last equation will be positive if the following statement is true :

$$min\{\rho(t,p) : p \in [0,1]\} \geq max\{\varrho(t,p) : p \in [0,1]\}. \tag{18}$$

Firstly, the derivative of $\varrho(t,p)$ with respect to the p variable is positive :

$$\frac{d\varrho}{dp}(t,p) = -\frac{1}{2}\log(t)t^{-\frac{p}{2}} \geq 0.$$

Which means that the function ϱ is increasing in p. So its maximum is reached for $p = 1$, i.e :

$$max\{\varrho(t,p) : p \in [0,1]\} = \varrho(t,1) = \frac{t}{2} + t^{-\frac{1}{2}}.$$

Secondly, the derivative of the function $\rho(t,p)$ with repsect to the p variable is negative :

$$\frac{d\rho}{dp}(t,p) = \frac{1}{2}\log(t)t^{\frac{p}{2}} \leq 0,$$

which means that the function ρ is decreasing in p. So its minimum is reached for $p = 1$, i.e :

$$min\{\rho(t,p) : p \in [0,1]\} = \rho(t,1) = \frac{t^{-1}}{2} + t^{\frac{1}{2}}.$$

We can now deduce that the statement (18) is true. Indeed :

$$\rho(t,1) - \varrho(t,1) = \frac{t^{-1}}{2} + t^{\frac{1}{2}} - \left(\frac{t}{2} + t^{-\frac{1}{2}}\right) = \frac{\sqrt{t}(1-t^2)}{2t\sqrt{t}} \geq 0.$$

For the case where $t > 1$ and by redefining the following functions as :

$$\rho(t,p) := \frac{t}{2} + t^{-\frac{p}{2}} \quad \text{and} \quad \varrho(t,p) := \frac{t^{-1}}{2} + t^{\frac{p}{2}},$$

we can get :

$$\frac{1}{2}\left|t^{-1}-t\right| - \left|t^{-\frac{p}{2}} - t^{\frac{p}{2}}\right| = \frac{1}{2}\left(t-t^{-1}\right) - \left(t^{\frac{p}{2}} - t^{-\frac{p}{2}}\right) = \rho(t,p) - \varrho(t,p).$$

By the same arguments above, we can easily verify that the statement (18) remain true for $t > 0$. The derivatives of the last functions with respect to p can be easily obtained as follows :

$$\frac{d\varrho}{dp}(t,p) = \frac{t}{2}\log(t)t^{\frac{p}{2}} \geq 0 \quad \text{and} \quad \frac{d\rho}{dp}(t,p) = -\frac{1}{2}\log(t)t^{-\frac{p}{2}} \leq 0.$$

It means that the function $\varrho(t,p)$ is increasing whence the $\rho(t,p)$ function is decreasing with respect to p. Then :

$$max\{\varrho(t,p) : p \in [0,1]\} = \varrho(t,1) = \frac{t^{-1}}{2} + t^{\frac{1}{2}},$$

and

$$min\{\rho(t,p) : p \in [0,1]\} = \rho(t,1) = \frac{t}{2} + t^{-\frac{1}{2}}.$$

Once again we easily ensure that the statement (18) is verified. Indeed :

$$\rho(t,1) - \varrho(t,1) = \frac{t}{2} + t^{-\frac{1}{2}} - \left(\frac{t^{-1}}{2} + t^{\frac{1}{2}}\right) = \frac{\sqrt{t}(t^2-1)}{2t\sqrt{t}} \geq 0,$$

and the lemma follows. □

We can now derive the following result.

Corollary 7. *For any vector $v > 0$, one has*

$$\left\|v^{-\frac{p}{2}} - v^{\frac{p}{2}}\right\| \leq \frac{1}{2}\left\|v^{-1} - v\right\|, \quad p \in [0,1].$$

Proof
Due to the lemma (6) and by developing the left term in the last inequality, one can obviously get :

$$\left\|v^{-\frac{p}{2}} - v^{\frac{p}{2}}\right\|^2 = \sum_{i=1}^{n}\left(v_i^{-\frac{p}{2}} - v_i^{\frac{p}{2}}\right)^2 \leq \sum_{i=1}^{n}\frac{1}{4}\left(v_i^{-1} - v_i\right)^2 = \frac{1}{4}\left\|v^{-1} - v\right\|^2.$$

And the Lemma follows. □

Furthermore, according to (5), we obtain:

$$\delta\left(v^{\frac{p}{2}}\right) \leq \frac{1}{2}\delta(v). \tag{19}$$

Lemma 8. *For any $t > 0$, one has : $|t^{-q} - t^p| \leq \left|t^{-1} - t\right| \quad \forall\, p, q \in [0,1]$.*

Proof
For t in $[0,1]$, we have by removing the absolute value and grouping terms in the following equation :

$$\begin{aligned}\left|t^{-1} - t\right| - \left|t^{-q} - t^p\right| &= \left(t^{-1} - t\right) - \left(t^{-q} - t^p\right) = \left(t^{-1} - t^{-q}\right) + \left(t^p - t\right)\\ &= t^{-1}\left(1 - t^{1-q}\right) + t^p\left(1 - t^{1-p}\right) \geq 0.\end{aligned}$$

For $t > 1$, we obtain by the same arguments :

$$\begin{aligned}\left|t^{-1} - t\right| - \left|t^{-q} - t^p\right| &= \left(t - t^{-1}\right) - \left(t^p - t^{-q}\right) = \left(t - t^p\right) + \left(t^{-q} - t^{-1}\right)\\ &= t^p\left(t^{1-p} - 1\right) + t^{-1}\left(t^{1-q} - 1\right) \geq 0.\end{aligned}$$

Which complete the proof. □

We can now deduce the following relation between our proximity measure already mentioned in (17) and the proximity measure defined by (5).

Corollary 9. *For any vector $v > 0$, we have : $\sigma(v) \leq \delta(v)$.*

Proof
The result is obviously derived from the last lemma. Indeed :

$$\begin{aligned}\sigma^2(v) = \frac{1}{4}\left\|v^{-q} - v^p\right\|^2 &= \frac{1}{4}\sum_{i=1}^{n}\left(v_i^{-q} - v_i^p\right)^2\\ &\leq \frac{1}{4}\sum_{i=1}^{n}\left(v_i^{-1} - v_i\right)^2 = \delta^2(v).\end{aligned}$$

And the Corollary follows. □

Thus the result is obtained

Lemma 10. *For $t > 0$ and p in $[0,1]$, one has the following two statements :*

$$\frac{1}{t} \leq \frac{1}{t^2} + \frac{1}{t^p} - \frac{1}{t^{p+1}} \tag{20}$$

$$\frac{1}{t} \geq \frac{1}{t^2 + t^p - t^{p+1}} \tag{21}$$

Proof
Since we have for any $t > 0$:

$$\frac{1}{t^2} + \frac{1}{t^p} - \frac{1}{t^{p+1}} - \frac{1}{t} = \frac{(t-1)\left(1 - t^{p-1}\right)}{t^{p+1}} \geq 0,$$

the first inequality (20) is then verified.
To prove the second inequality (21), we observe that the right hand side of the following equation :

$$\frac{1}{t} - \frac{1}{t^2 + t^p - t^{p+1}} = \frac{(t-1)\left(1 - t^{p-1}\right)}{t^2 + t^p - t^{p+1}}$$

is always positive for any $t > 0$. Thus the lemma follows. □

We can now state the following result which will play an important role regarding the feasibility analysis discussed in the next section.

Corollary 11. *For $t > 0$ and p in $[0,1]$, one has :*

$$\frac{1}{t^2 + t^p - t^{p+1}} \leq \frac{1}{t^2} + \frac{1}{t^p} - \frac{1}{t^{p+1}}$$

Proof
The result is consequently derived by combining the two inequalities in the lemma (10) where the function $t \longmapsto 1/t$ played the intermediate role between its upper bound and the lower one. □

Lemma 12. *According to the result of Lemma (1), for any $p \in [0,1]$ one has*

$$\left\| v^{\frac{p}{2}} \right\| \leq \sqrt{n}.$$

Proof
By applying Hölder's inequality, we obtain :

$$\begin{aligned}
\left\| v^{\frac{p}{2}} \right\|^2 &= \sum_{i=1}^{n} \left(v_i^{\frac{p}{2}}\right)^2 = \sum_{i=1}^{n} v_i^p \\
&= \sum_{i=1}^{n} \left(v_i^2\right)^{\frac{p}{2}} . \ (1)^{\frac{2-p}{2}} \\
&\leq \left(\sum_{i=1}^{n} v_i^2\right)^{\frac{p}{2}} . \left(\sum_{i=1}^{n} 1\right)^{\frac{2-p}{2}} \\
&= n^{\frac{p}{2}} . \, n^{\frac{2-p}{2}} = n.
\end{aligned}$$

Thus the result is obtained. □

Lemma 13. *For any vector $v >$ and $p \in [0,1]$, one has*

$$\text{(i)} \quad \left\| \sqrt{1-\theta} v^{-\frac{p}{2}} - \frac{v^{\frac{p}{2}}}{\sqrt{1-\theta}} \right\|^2 \leq (1-\theta)\,\delta^2(v) + \frac{\theta^2 n}{1-\theta}$$

and

$$\text{(ii)} \quad \left\| \sqrt{1-\theta} v^{-1} - \frac{v}{\sqrt{1-\theta}} \right\|^2 \leq 4\,(1-\theta)\,\delta^2(v) + \frac{\theta^2 n}{1-\theta}.$$

Proof
By applying the lemma (12) and (19), we obtain

$$\begin{aligned}
\left\| \sqrt{1-\theta} v^{-\frac{p}{2}} - \frac{v^{\frac{p}{2}}}{\sqrt{1-\theta}} \right\|^2 &= \left\| \sqrt{1-\theta} \left(v^{-\frac{p}{2}} - \frac{v^{\frac{p}{2}}}{1-\theta} \right) \right\|^2 \\
&= (1-\theta) \left\| v^{-\frac{p}{2}} - v^{\frac{p}{2}} + v^{\frac{p}{2}} - \frac{v^{\frac{p}{2}}}{1-\theta} \right\|^2 \\
&= (1-\theta) \left\| v^{-\frac{p}{2}} - v^{\frac{p}{2}} \right\|^2 + \frac{\theta^2}{1-\theta} \left\| v^{\frac{p}{2}} \right\|^2 - 2\theta \left(v^{-\frac{p}{2}} - v^{\frac{p}{2}} \right)^T v^{\frac{p}{2}} \\
&\leq 4\,(1-\theta)\,\delta^2\left(v^{\frac{p}{2}}\right) + \frac{\theta^2 n}{1-\theta} \\
&\leq (1-\theta)\,\delta^2(v) + \frac{\theta^2 n}{1-\theta}.
\end{aligned}$$

The second statement can be checked by following the same previous steps. Thus the lemma follows. □

Lemma 14. *(See [1], Lemma A.1) For $i = 1, ..., m$, let $f_i : \mathbb{R}_+ \to \mathbb{R}$ denote a convex function. Then, for any nonzero $z \in \mathbb{R}^n_+$, the following inequality*

$$\sum_{i=1}^{n} f_i(z_i) \leq \frac{1}{e^T z} \sum_{j=1}^{n} z_j \left(f_j(e^T z) + \sum_{i \neq j} f_i(0) \right)$$

holds.

4. Analysis of the feasibility step

Let x, y and s denote the iterates at the start of an iteration, and assume that $x^T s = n\mu$ and $\delta(v) \leq \tau$ which is true at the first iteration since $\delta(v^0) = 0$, according to the choice of (x^0, s^0) stated in Sect.(3.3).

4.1. Feasibility step

As we established in Sect.(3.3), the feasibility step generates new iterate (x^f, y^f, s^f) that satisfies the feasibility conditions for (P_{ν^+}) and (D_{ν^+}), except possibly the nonnegativity constraints. A crucial element in the analysis is to show that, after the feasibility step, we get $\delta(x^f, s^f; \mu^+) \leq 1/\sqrt[4]{2}$, i.e., the new iterates (x^f, y^f, s^f) are positive and within the neighborhood where the Newton process targeting the μ^+-center of (P_{ν^+}) and (D_{ν^+}) is quadratically convergent.

Note that (16) can be rewritten as

$$\begin{aligned} s\Delta x + x\Delta s &= \mu^{\frac{1+q}{2}}(xs)^{\frac{1-q}{2}} - \mu^{\frac{1-p}{2}}(xs)^{\frac{1+p}{2}} \\ &= \mu^{\frac{2-p}{2}}(xs)^{\frac{p}{2}} - \mu^{\frac{1-p}{2}}(xs)^{\frac{1+p}{2}}, \qquad \text{since } \ q = 1-p \\ &= \mu^{1-\frac{p}{2}}\mu^{\frac{p}{2}}v^p - \mu^{\frac{1-p}{2}}\mu^{\frac{1+p}{2}}v^{1+p} \\ &= \mu\left(v^p - v^{p+1}\right). \end{aligned} \tag{22}$$

Using $xs = \mu v^2$ and $\Delta^f x \Delta^f s = \mu d_x^f d_s^f$, we obtain

$$\begin{aligned} x^f s^f &= xs + (s\Delta^f x + x\Delta^f s) + \Delta^f x \Delta^f s \\ &= \mu v^2 + \mu v^p (e - v) + \mu d_x^f d_s^f \\ &= \mu\left(v^2 + v^p - v^{p+1} + d_x^f d_s^f\right). \end{aligned} \tag{23}$$

The feasibility condition can now be stated in the following Lemma.

Lemma 15. *The iterates (x^f, y^f, s^f) are strictly feasible if and only if*

$$v^2 + v^p - v^{p+1} + d_x^f d_s^f > 0 \ \ \forall p \in [0,1].$$

Proof
If x^f and s^f are positive then (23) makes clear that $v^2 + v^p - v^{p+1} + d_x^f d_s^f > 0$, proving the only if part of the lemma. For the proof of the converse implication, we introduce a steplength $\alpha \in [0,1]$ and we define $x^\alpha = x + \alpha\Delta^f x$ and $s^\alpha = s + \alpha\Delta^f s$.
We then have $x^0 = x$, $x^1 = x^f$ and similar relations for s. Hence we have $x^0 s^0 = xs > 0$. Using (22), $xs = \mu v^2$ and $\Delta^f x\Delta^f s = \mu d_x^f d_s^f$, we may write:

$$\begin{aligned} x^\alpha s^\alpha &= (x + \alpha\Delta^f x)(s + \alpha\Delta^f s) \\ &= xs + \alpha(s\Delta^f x + x\Delta^f s) + \alpha^2 \Delta^f x\Delta^f s \\ &= \mu v^2 + \alpha\mu\left(v^p - v^{p+1}\right) + \alpha^2\mu d_x^f d_s^f \\ &> \mu(1-\alpha)\left(v^2 + \alpha\left(v^2 + v^p - v^{p+1}\right)\right). \end{aligned}$$

The right hand-side of the last inequality is nonnegative. Indeed, if the ith coordinate v_i belongs to the interval $[0,1]$ the term

$$v_i^2 + v_i^p - v_i^{p+1} = v_i^2 + v_i^p(1 - v_i)$$

is obviously non negative.
For the case $v_i \geq 1$, the quantity

$$v_i^2 + v_i^p - v_i^{p+1} = v_i^{p+1}\left(v_i^{1-p} - 1\right) + v_i^p$$

stills also non negative.
It follows that $x^\alpha s^\alpha > 0$ for $\alpha \in [0,1]$. Hence, none of the entries of x^α and s^α vanishes for $0 \leq \alpha \leq 1$. Since x^0 and s^0 are positive, and x^α and s^α depend linearly on α, this implies that $x^\alpha > 0$ and $s^\alpha > 0$ for any α in the interval $[0,1]$. Hence, x^1 and s^1 must be positive, proving the 'if' part of the statement in the lemma. □

We define

$$\tilde{v} := v^2 + v^p - v^{p+1}. \tag{24}$$

Because we need a lower bound to the vector $\tilde{v}$, we recall the following Lemma stated in Roos's book [2].

Lemma 16. *(See Lemma II.60 in [2]) Let $\rho(\delta) := \delta + \sqrt{1+\delta^2}$, which is simply denoted later by $\rho := \rho(\delta)$. Then*

$$\frac{1}{\rho(\delta)} \le v_i \le \rho(\delta), \quad 1 \le i \le n.$$

By using the previous Lemma, we can easily derive the following bounds of $\tilde{v}_i$

$$\frac{1}{\rho^2} + \frac{1}{\rho^p} - \rho^{p+1} \le \tilde{v}_i \le \rho^2 + \rho^p - \frac{1}{\rho^{p+1}} \tag{25}$$

We denote $\tilde{\rho} := \rho^{-2} + \rho^{-p} - \rho^{p+1}$, which means that $\tilde{\rho} \le \tilde{v}_i$.
Instead of $\tilde{\rho}$, we can define and take the following value

$$\overline{\rho}_p(\rho) := 2\rho^{-2} - \rho^{p+1} \ , \tag{26}$$

as a lower bound to $\tilde{v}_i$ since obviously $\overline{\rho}_p(\rho) \le \tilde{\rho}$ because by definition $\rho \ge 1$.

Furthermore the function $t \longmapsto 2t^{-2} - t^{p+1}$ where $t \ge 1$ is decreasing and it vanishes in $t_p^* = 2^{\frac{1}{p+3}}$ which obviously depends on p.
We are so interested in the minimum value of t_p^* with respect to the variable p. In fact, the function $p \longmapsto 2^{\frac{1}{p+3}}$ is decreasing since its derivative

$$\frac{-2^{\frac{1}{p+3}}}{(p+3)^2} \log 2$$

is negative, involving

$$\min\{t_p^* \ / \ p \in [0,1]\} = t_1^* = \sqrt[4]{2} \cong 1.19 \ , \tag{27}$$

which yields without any more proof the following result

Lemma 17. *For any p in $[0,1]$, the lower bound $\overline{\rho}_p(\rho)$ of $\tilde{v}_i$ is positive if*

$$\rho \in [1, \sqrt[4]{2}[. \tag{28}$$

In the sequel we denote

$$\omega_i := \omega_i(v) := \frac{1}{2}\sqrt{|d_{xi}^f|^2 + |d_{si}^f|^2},$$

and $\omega := \omega(v) := \|(\omega_1, ..., \omega_n)\|$. This implies

$$(d_x^f)^T d_s^f \le \|d_x^f\|\|d_s^f\| \le \frac{1}{2}(\|d_x^f\|^2 + \|d_s^f\|^2) = 2\omega^2,$$
$$|d_{xi}^f d_{si}^f| = |d_{xi}^f||d_{si}^f| \le \frac{1}{2}(|d_{xi}^f|^2 + |d_{si}^f|^2) = 2\omega_i^2 \le 2\omega^2, \quad 1 \le i \le n.$$

We proceed by deriving an upper bound for $\delta(x^f, s^f; \mu^+)$. According to definition (5), one has

$$\delta(x^f, s^f; \mu^+) = \frac{1}{2}\left\|v^f - (v^f)^{-1}\right\|, \quad \text{where } \ v^f := \sqrt{\frac{x^f s^f}{\mu^+}}. \tag{29}$$

In the sequel, we denote $\delta(x^f, s^f; \mu^+)$ by $\delta(v^f)$ and we have the following result.

Lemma 18. *Assuming $v^2 + v^p - v^{p+1} + d_x^f d_s^f > 0$, one has*

$$4\delta^2\left(v^f\right) \le 5\,(1-\theta)\,\delta^2(v) + \frac{2\theta^2 n}{1-\theta} + \frac{2\omega^2}{1-\theta} + \frac{(1-\theta)2\omega^2}{\overline{\rho}_p\left(\overline{\rho}_p - 2\omega^2\right)} \ ,$$

where $\overline{\rho}_p := \overline{\rho}_p(\rho)$.

Proof
According to (29) and (23), one has

$$\left(v^f\right)^2 = \frac{x^f s^f}{\mu(1-\theta)} = \frac{v^2 + v^p - v^{p+1} + d_x^f d_s^f}{1-\theta} = \frac{\tilde{v} + d_x^f d_s^f}{1-\theta}$$

and

$$\begin{aligned} 4\delta^2\left(v^f\right) &= \sum_{i=1}^{n}\left((v_i^f)^2 + (v_i^f)^{-2} - 2\right) \\ &= \sum_{i=1}^{n}\left(\frac{\tilde{v}_i + d_{xi}^f d_{si}^f}{1-\theta} + \frac{1-\theta}{\tilde{v}_i + d_{xi}^f d_{si}^f} - 2\right) \\ &\leq \sum_{i=1}^{n}\left(\frac{\tilde{v}_i + 2\omega_i^2}{1-\theta} + \frac{1-\theta}{\tilde{v}_i - 2\omega_i^2} - 2\right). \end{aligned}$$

For each i we define the function

$$f_i(z_i) = \frac{\tilde{v}_i + z_i}{1-\theta} + \frac{1-\theta}{\tilde{v}_i - z_i} - 2, \quad i = 1, ..., n.$$

One can easily verify that if $\tilde{v}_i - z_i > 0$ then $f_i(z_i)$ is convex in z_i. By taking $z_i = 2\omega_i^2$, we can require

$$\tilde{v}_i - 2\omega_i^2 > 0 \ .$$

By using (25) and (26), this holds if

$$2\omega^2 \leq \overline{\rho}_p. \tag{30}$$

Furthermore, we can use Lemma (14) to get

$$\begin{aligned} 4\delta^2\left(v^f\right) &\leq \sum_{i=1}^{n} f_j(\omega_j) \leq \frac{1}{2\omega^2}\sum_{j=1}^{n} 2\omega_j^2\left(f_j(2\omega^2) + \sum_{i\neq j} f_i(0)\right) \\ &= \frac{1}{2\omega^2}\sum_{j=1}^{n}\left[2\omega_j^2\left(\left(\frac{\tilde{v}_j + 2\omega^2}{1-\theta} + \frac{1-\theta}{\tilde{v}_j - 2\omega^2} - 2\right) + \sum_{i\neq j} f_i(0)\right)\right]. \end{aligned}$$

Using Corollary (11) and Lemma (13) we obtain

$$\begin{aligned}
\sum_{i=1}^{n}\left(\frac{\tilde{v}_i}{1-\theta}+\frac{1-\theta}{\tilde{v}_i}-2\right) &= \sum_{i=1}^{n}\left(\frac{v^2+v^p-v^{p+1}}{1-\theta}+\frac{1-\theta}{v^2+v^p-v^{p+1}}-2\right)\\
&\leq \sum_{i=1}^{n}\left(\frac{v^2+v^p-v^{p+1}}{1-\theta}+\frac{1-\theta}{v^2}+\frac{1-\theta}{v^p}-\frac{1-\theta}{v^{p+1}}-2\right)\\
&= \sum_{i=1}^{n}\left(\frac{v^2}{1-\theta}+\frac{1-\theta}{v^2}-2\right)+\sum_{i=1}^{n}\left(\frac{v^p}{1-\theta}+\frac{1-\theta}{v^p}-2\right)\\
&\qquad -\sum_{i=1}^{n}\left(\frac{v^{p+1}}{1-\theta}+\frac{1-\theta}{v^{p+1}}-2\right)\\
&= \left\|\sqrt{1-\theta}v^{-1}-\frac{v}{\sqrt{1-\theta}}\right\|^2+\left\|\sqrt{1-\theta}v^{-\frac{p}{2}}-\frac{v^{\frac{p}{2}}}{\sqrt{1-\theta}}\right\|^2\\
&\qquad -\left\|\sqrt{1-\theta}v^{-\frac{p+1}{2}}-\frac{v^{\frac{p+1}{2}}}{\sqrt{1-\theta}}\right\|^2\\
&\leq \left\|\sqrt{1-\theta}v^{-1}-\frac{v}{\sqrt{1-\theta}}\right\|^2+\left\|\sqrt{1-\theta}v^{-\frac{p}{2}}-\frac{v^{\frac{p}{2}}}{\sqrt{1-\theta}}\right\|^2\\
&\leq 5\left(1-\theta\right)\delta^2(v)+\frac{2\theta^2 n}{1-\theta},
\end{aligned}$$

which implies

$$\begin{aligned}
\sum_{i\neq j} f_i(0) &= \sum_{i\neq j}\left(\frac{\tilde{v}_i}{1-\theta}+\frac{1-\theta}{\tilde{v}_i}-2\right)\\
&= \left[\sum_{i=1}^{n}\left(\frac{\tilde{v}_i}{1-\theta}+\frac{1-\theta}{\tilde{v}_i}-2\right)\right]-\left(\frac{\tilde{v}_j}{1-\theta}+\frac{1-\theta}{\tilde{v}_j}-2\right)\\
&\leq 5\left(1-\theta\right)\delta^2(v)+\frac{2\theta^2 n}{1-\theta}-\left(\frac{\tilde{v}_j}{1-\theta}+\frac{1-\theta}{\tilde{v}_j}-2\right).
\end{aligned}$$

Then

$$\begin{aligned}
4\delta^2\left(v^f\right) &\leq 5\left(1-\theta\right)\delta^2(v)+\frac{2\theta^2 n}{1-\theta}\\
&\quad +\frac{1}{2\omega^2}\sum_{i=j}^{n}2\omega_j^2\left(\frac{\tilde{v}_j+2\omega^2}{1-\theta}+\frac{1-\theta}{\tilde{v}_j-2\omega^2}-2-\left(\frac{\tilde{v}_j}{1-\theta}+\frac{1-\theta}{\tilde{v}_j}-2\right)\right)\\
&= 5\left(1-\theta\right)\delta^2(v)+\frac{2\theta^2 n}{1-\theta}\\
&\quad +\frac{2\omega^2}{1-\theta}+\frac{1}{2\omega^2}\sum_{i=j}^{n}2\omega_j^2\left(1-\theta\right)\left(\frac{1}{\tilde{v}_j-2\omega^2}-\frac{1}{\tilde{v}_j}\right)\\
&= 5\left(1-\theta\right)\delta^2(v)+\frac{2\theta^2 n}{1-\theta}+\frac{2\omega^2}{1-\theta}+\frac{1}{2\omega^2}\sum_{i=j}^{n}2\omega_j^2\frac{(1-\theta)2\omega^2}{\tilde{v}_j\left(\tilde{v}_j-2\omega^2\right)}.
\end{aligned}$$

Since the last term in the right hand side is decreasing with respect to $\tilde{v_j}$ and by using again (25) and (26), we deduce

$$4\delta^2\left(v^f\right) \leq 5\,(1-\theta)\,\delta^2(v) + \frac{2\theta^2 n}{1-\theta} + \frac{2\omega^2}{1-\theta} + \frac{(1-\theta)2\omega^2}{\overline{\rho}_p\left(\overline{\rho}_p - 2\omega^2\right)}\ ,$$

which completes the proof of the lemma. □

We choose now

$$\tau = \frac{1}{16}, \qquad \theta = \frac{\alpha}{2\sqrt{n}}, \qquad 0 \leq \alpha \leq 1. \tag{31}$$

Because we need to have $\delta(v^f) \leq 1/\sqrt[4]{2}$, it follows from Lemma (18) that it suffices if

$$5\,(1-\theta)\,\delta^2(v) + \frac{2\theta^2 n}{1-\theta} + \frac{2\omega^2}{1-\theta} + \frac{(1-\theta)2\omega^2}{\overline{\rho}_p\left(\overline{\rho}_p - 2\omega^2\right)} \leq 2\sqrt{2}. \tag{32}$$

Once again, since the function $\overline{\rho}_p \longmapsto \frac{1}{\overline{\rho}_p\left(\overline{\rho}_p - 2\omega^2\right)}$ is decreasing, we need a lower bound of $\overline{\rho}_p$, which depends on p and ρ, in order to maximize the last term in the left hand side of inequality (32). One can easily check that $\overline{\rho}_p$ is decreasing with respect to the p variable since its derivative

$$-\rho^{p+1}\log\rho$$

is negative where $\rho \geq 1$. So the minimum value of $\overline{\rho}_p$ is reached for $p = 1$

$$\min\{\overline{\rho}_p\} = \overline{\rho}_1 = 2\rho^{-2} - \rho^2.$$

On the other hand, the function $\overline{\rho}_p$ is also decreasing with respect to ρ variable. So its minimum value is reached for the following value of ρ denoted by

$$\widehat{\rho} := \frac{1}{16} + \sqrt{1 + \left(\frac{1}{16}\right)^2} \cong 1.0645\ , \tag{33}$$

since $\delta \leq \tau$. In addition, the condition (28) is satisfied because $\widehat{\rho} < \sqrt[4]{2}$, which guarantees that $\overline{\rho}_p$ remains positive. Thus by replacing the value $\widehat{\rho}$ into the function $\overline{\rho}_p$, we can easily deduce

$$\begin{aligned}\min\left\{\overline{\rho}_p \text{ such that } \ p \in [0,1] \text{ and } \rho \in [1, \sqrt[4]{2}[\right\} &= 2\,(\widehat{\rho})^{-2} - (\widehat{\rho})^2 \\ &\cong 0.6321\ .\end{aligned} \tag{34}$$

Furthermore one can check that the left-hand side of (32) is monotonically increasing with respect to ω^2. By some elementary calculations, for $n \geq 1$ and $\delta(v) \leq \tau$, we obtain

$$\omega \leq 0.3735 \qquad \Longrightarrow \qquad \delta(v^f) \leq \frac{1}{\sqrt[4]{2}}. \tag{35}$$

We note that with the previous choice of ω, the condition (30) still satisfied. Later in this paper, we will give a precise value of α.

4.2. *Upper Bound for* $\omega(v)$

Let us denote the null space of the matrix $\overline{A}$ as $\mathcal{L}$. So

$$\mathcal{L} := \{\xi \in \mathbb{R}^n \ : \ \overline{A}\xi = 0\}.$$

Then the affine space $\{\xi \in \mathbb{R}^n \ : \ \overline{A}\xi = \theta\nu r_b^0\}$ equals $d_x + \mathcal{L}$. Note that the row space of $\overline{A}$ equals the orthogonal complement $\mathcal{L}^{\perp}$ of $\mathcal{L}$, and $d_s^f \in \theta\nu v s^{-1} r_c^0 + \mathcal{L}^{\perp}$. We recall the following result from Roos [7].

Lemma 19. *(See Lemma 4.6 in [7]) Let q be the (unique) point in the intersection of the affine spaces $d_x + \mathcal{L}$ and $d_s + \mathcal{L}^{\perp}$. Then*

$$2\omega(v) \leq \sqrt{\|q\|^2 + (\|q\| + 2\sigma(v))^2}.$$

Note that (35) implies that we must have $\omega \leq 0.3735$ to guarantee $\delta(v^f) \leq 1/\sqrt[4]{2}$. Due to Lemma (19) this will certainly hold if $\|q\|$ satisfies

$$\|q\|^2 + (\|q\| + 2\delta(v))^2 \leq 4 * (0.3735)^2. \tag{36}$$

Furthermore from Ross [7], we can have

$$\sqrt{\mu}\|q\| \leq \theta\nu\zeta\sqrt{e^T\left(\frac{x}{s} + \frac{s}{x}\right)}. \tag{37}$$

In what follows, we give bounds for the rate vectors x/s and s/x.

4.3. *Bounds for* x/s *and* s/x

Note that x is feasible for P_ν and (y, s) for D_ν and moreover $\delta(x, s; \mu) \leq \tau$, i.e., these iterates are close to the μ-centers of P_ν and D_ν. Based on this information we need to estimate the sizes of the entries of the vectors x/s and s/x. We recall now a useful result from Roos [7], which is still available for our choice of $\tau = \frac{1}{16}$.

Corollary 20. *Let $\tau = \frac{1}{16}$ and $\delta(v) \leq \tau$. Then*

$$\sqrt{\frac{x}{s}} \leq \sqrt{2}\frac{x(\mu,\nu)}{\sqrt{\mu}}, \qquad \sqrt{\frac{s}{x}} \leq \sqrt{2}\frac{s(\mu,\nu)}{\sqrt{\mu}}. \tag{38}$$

Proof

Following the same notations as in the Appendix of Roos ([7], Corollary A.10 and Theorem A.9), and by some elementary calculations, one may easily verify that if $\tau = \frac{1}{16}$, then $\tau' \approx 0.0041$, $\varrho(\tau') \approx 1.0645$ and $\chi(\tau') \approx 0.9369$, which gives

$$\frac{\varrho(\tau')}{\chi(\tau')} \approx 1.1362 < \sqrt{2}\ (\approx 1.4142).$$

Thus the result follows. □

Using (38) and substituting into (37) yields

$$\sqrt{\mu}\|q\| \leq \theta\nu\zeta\sqrt{2e^T\left(\frac{x(\mu,\nu)^2}{\mu} + \frac{s(\mu,\nu)}{\mu}\right)}. \tag{39}$$

This gives

$$\mu\|q\| \leq \sqrt{2}\theta\nu\zeta\sqrt{\|x(\mu,\nu)\|^2 + \|s(\mu,\nu)\|^2}. \tag{40}$$

By substitution of $\mu = \mu^0\nu = \zeta^2\nu =$ and $\theta = \frac{\alpha}{2\sqrt{n}}$ into (40), we obtain the following upper bound for the norm of $\|q\|$

$$\|q\| \leq \frac{\alpha}{\sqrt{2n}\ \zeta}\sqrt{\|x(\mu,\nu)\|^2 + \|s(\mu,\nu)\|^2}. \tag{41}$$

We define

$$\kappa(\zeta,\nu) = \frac{\sqrt{\|x(\mu,\nu)\|^2 + \|s(\mu,\nu)\|^2}}{\sqrt{2n}\,\zeta}, \quad 0 < \nu \leq 1,\ \mu = \mu^0\nu,$$

and

$$\overline{\kappa}(\zeta) = \max_{0<\nu\leq 1} \kappa(\zeta,\nu),$$

we obtain so

$$\|q\| \leq \alpha\,\overline{\kappa}(\zeta).$$

Because we are looking for the value that we do not allow ω to exceed and in order to guarantee that $\delta^f(v) \leq 1/\sqrt[4]{2}$, (36) holds if $\|q\|$ satisfies $\|q\|^2 + \left(\|q\| + \frac{1}{8}\right)^2 \leq 4 * (0.3735)^2$, since $\delta(v) \leq 1/16$. This will be certainly satisfied if $\|q\| \leq 0.462$. We may now deduce the value of α:

$$\alpha = \frac{0.462}{\overline{\kappa}(\zeta)}\ , \tag{42}$$

which guarantee that $\delta^f(v) \leq 1/\sqrt[4]{2}$ holds. Furthermore, following Sect.4.6 in Roos [7], we can prove that $\overline{\kappa}(\zeta) = \sqrt{2n}$. We then conclude by substitution into (42) that

$$\theta = \frac{0.462}{2\sqrt{2}\,n}. \tag{43}$$

5. Iteration Bound

In the previous sections, we have found that, if at the start of an iteration the iterates satisfy $\delta(x,s;\mu) \leq \tau$, with $\tau = 1/16$, then after the feasibility step, with θ as defined in (43), the feasible iterates satisfy $\delta(x^f, s^f; \mu^+) \leq 1/\sqrt[4]{2}$.

According to (11), so each main iteration consists of at most

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

so-called inner iterations, in each of which we need to compute a new search direction to get the iterate $(x^+, s^+; \mu^+)$ that satisfies $\delta(x^+, s^+; \mu^+) \leq \tau$. In each main iteration both the duality gap and the norms of the residual vectors are reduced by the factor $1 - \theta$. Hence, using $(x^0)^T s^0 = n\zeta^2$, the total number of main iterations is bounded above by

$$\frac{1}{\theta} \log \frac{\max\{n\zeta^2, \|r_b^0\|, \|r_c^0\|\}}{\varepsilon}.$$

By using (43), the total number of inner iterations is so bounded above by

$$17\sqrt{2}n \log \frac{\max\{n\zeta^2, \|r_b^0\|, \|r_c^0\|\}}{\varepsilon}.$$

6. Numerical results

To carry out the numerical tests, we chose four kernel functions ψ_1, ψ_2, ψ_3, ψ_4, according to the definition (15), that correspond respectively to the following parameters : $1, 0.85, 0.5$ and 0.2. The Figure 1 displays the curves of the four kernel functions used in the numerical tests to provide a geometric illustration of their behavior. For each of these functions, we performed tests under the MATLAB† environment, on 57 problems belonging to the Netlib

†https://www.mathworks.com/

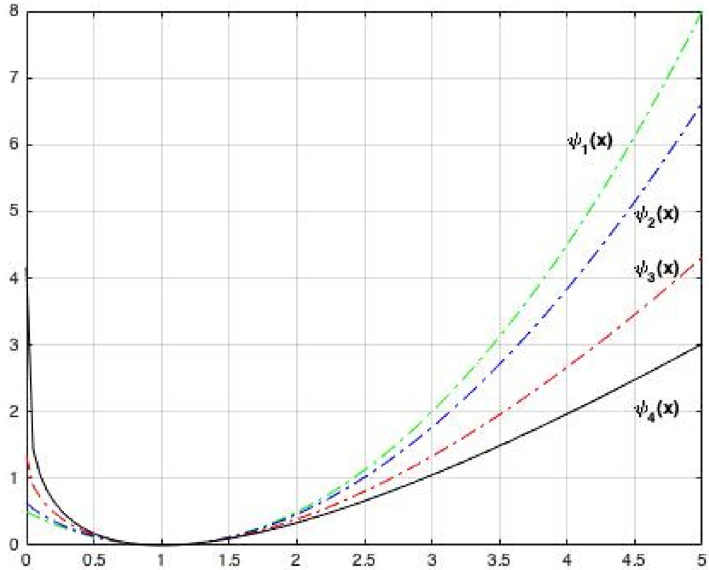

Figure 1. The curves of the four chosen kernel functions ψ_1, ψ_2, ψ_3 and ψ_4

Standard Library for linear programming[‡].

By comparing these 4 functions with respect to each problem, we notice that overall the number of iterations increases when the value of the parameter p decreases. Despite the performance of the function ψ_1, which generally recorded the best score in terms of number of iterations, we observed that the function ψ_4, whose parameter $p = 0.2$, outperforms the other test functions for the problems: GREENBEB, BORE3D, RECIPELP, even though its parameter approaches zero. Furthermore, the number of iterations of the ψ_4 function is greater than or equal to that recorded by ψ_2 and ψ_3 for the problems AGG3, CZPROB, ISRAEL, MAROS, SCORPION, SCTAP1, SCTAP2, SHARE1B, SHARE2B, SHELL, SHIP08S, and SHIP12L, as shown in Table 6. These numerical results show that the selection of the function relative to its parameter p could be beneficial for certain problems. This choice is, of course, legitimate, since the work presented in this article theoretically proves convergence for a spectrum of parametric functions where the parameter p continuously ranges over the interval $]0, 1]$ (excluding 0).

Although the performance of the MATLAB code for this algorithm, whether in number of iterations or in resolution time, approaches that recorded by LIPSOL solver [11], for the majority of LP problems, we observed some limitations of our implementation in solving certain problems. As an example, for the BOEING1 problem, the best value reached is approximately $-3.344605e$+02 for $p = 1$, with a primal residual of $4.38e$-07 while the dual residual is $7.08e$+01. This gap between the residuals directly impacts the convergence to the final solution, which equals $-3.3521356751e$+02. In principle, our code should simultaneously monitor the calculation errors of both primal and dual residuals to ensure a cautious displacement for such problems, whose conditioning requires specific pre-processing. This also applies to the problems: BOEING2, CAPRI, and DFL001.

[‡]Linear programming test problems can be found at https://www.netlib.org/lp/data

LP Problem	ψ_1 $(p = 1)$		ψ_2 $(p = 0.85)$		ψ_3 $(p = 0.5)$		ψ_4 $(p = 0.2)$	
	iter	cpu-time	iter	cpu-time	iter	cpu-time	iter	cpu-time
25FV47	28	0.74	35	0.92	39	1.03	39	1.08
80BAU3B	90	5.26	83	5.13	86	5.21	89	5.78
adlittle	14	0.15	34	0.22	36	0.29	38	0.32
AFIRO	8	0.09	30	0.26	33	0.28	33	0.27
AGG	22	0.46	55	1.08	59	1.10	59	1.07
AGG2	20	0.41	39	0.80	50	0.96	51	1.06
AGG3	22	0.52	40	0.87	49	1.13	**29**	0.38
BANDM	20	0.30	29	0.36	32	0.33	34	0.38
BLEND	14	0.12	25	0.21	31	0.25	28	0.18
BNL1	30	0.48	41	0.73	45	0.80	47	0.89
BNL2	36	1.56	45	2.04	48	2.18	50	2.28
BOEING1	!	!	!	!	!	!	!	!
BOEING2	!	!	!	!	!	!	!	!
BORE3D	64	0.19	49	0.13	52	0.15	**37**	0.07
BRANDY	20	0.27	32	0.43	38	0.48	39	0.49
CAPRI	!	!	!	!	!	!	!	!
CZPROB	41	0.19	66	1.31	69	1.47	**48**	1.15
D2Q06C	36	3.01	47	3.87	49	4.10	50	4.14
D6CUBE	31	1.36	69	3.24	41	1.93	70	3.32
DEGEN2	19	0.38	69	1.46	88	1.91	96	2.05
DEGEN3	58	4.69	85	6.51	100	7.80	120	9.26
DFL001	!	!	!	!	!	!	!	!
E226	24	0.40	35	0.48	39	0.57	40	0.60
ETAMACRO	!	!	!	!	!	!	!	!
FFFFF800	29	0.57	35	0.80	41	0.95	46	1.09
FINNIS	!	!	!	!	!	!	!	!
FIT1D	35	0.86	32	0.80	55	1.36	53	1.11
FORPLAN	37	0.64	36	0.62	41	0.70	43	0.72
GANGES	70	1.75	67	1.69	54	1.36	50	1.24
GREENBEB	85	4.79	77	4.64	66	3.87	**59**	3.54
ISRAEL	28	0.44	38	0.70	50	0.71	**32**	0.47
LOTFI	23	0.21	32	0.29	40	0.42	42	0.46
MAROS	34	0.86	53	1.34	54	1.37	**52**	1.34
RECIPELP	73	0.66	54	0.51	47	0.43	**45**	0.40
SCORPION	19	0.16	41	0.48	39	0.42	**30**	0.31
SCTAP1	18	0.25	35	0.44	35	0.45	**27**	0.35
SCTAP2	20	0.42	34	0.69	40	0.82	**34**	0.71
SCTAP3	22	0.38	35	0.84	40	0.92	44	1.03
SHARE1B	22	0.29	40	0.40	38	0.38	36	0.34
SHARE2B	14	0.16	27	0.23	29	0.26	**22**	0.23
SHELL	78	1.47	73	1.38	60	1.06	56	0.96
SHIP04l	16	0.33	35	0.63	34	0.58	44	0.78
SHIP04S	18	0.27	30	0.48	37	0.60	42	0.66
SHIP08L	16	0.48	35	0.95	41	1.05	42	1.07
SHIP08S	18	0.36	39	0.77	43	0.79	39	0.74
SHIP12L	24	0.61	35	0.90	37	0.98	37	0.95
SHIP12S	20	0.38	36	0.65	32	0.58	39	0.74
SIERRA	60	1.69	65	1.97	63	1.75	64	1.81
STAIR	!	!	51	1.06	45	0.89	47	0.95
STANDATA	52	0.78	55	0.99	58	1.00	59	1.13
STANDGUB	44	0.62	51	0.98	52	0.94	55	0.99
STANDMPS	42	0.78	40	0.63	46	0.80	47	0.85
STOCFOR1	16	0.13	30	0.23	33	0.26	33	0.29
STOCFOR2	25	0.65	38	1.11	38	1.09	39	1.11
TRUSS	21	0.85	33	1.28	33	1.34	32	1.30
WOOD1P	28	1.19	41	1.67	42	1.69	**39**	1.55
WOODW	34	1.38	42	1.90	45	2.06	46	2.09

Table 1. Number of iterations (iter) and CPU-time in seconds

7. Concluding Remarks

In this paper we propose a new search direction based on a class of kernel function depending on the parameter $p \in]0,1]$ to generate the infeasible interior-point algorithm with full-Newton steps for linear optimization. We benefits the nice property of a sharper quadratic convergence which results in a wider neighborhood for the feasibility steps, making algorithm more stable. The iteration bound coincides with the currently best known bound for IIPMs. Future research might focuses firstly on the analysis of this class of kernel functions with the parameter $p \geq 1$ to provide more extension of the interval $]0,1]$ and also on the generalization to other classes of optimization problems, as second-order cone optimization, semi-definite optimization, and also P_*-matrix LCP.

REFERENCES

1. H. Mansouri C. Roos. Simplified $\mathcal{O}(nL)$ infeasible interior-point algorithm for linear optimization using full-newton steps. *Optim. Methods Softw*, 22:519–530, 2007.
2. T. Terlaky C. Roos and J.-Ph. Vial. Theory and algorithms for linear optimization. an interior-point approach. 2005.
3. B. Kheirfam and M. Haghighi. A full-newton step infeasible interior-point algorithm for monotone lcp based on a locally kernel function. *A Journal of Mathematical Programming and Operations Research*, 65:841–857, 2016.
4. Y. Bai L. Zhang and Y. Xu. A full-newton step infeasible interior-point method for linear optimization based on a trigonometric kernel function. *Numerical Algorithms*, 61:57–81, 2012.
5. S. Mizuno. Polynomiality of infeasible-interior-point algorithms for linear programming. *Math. Program*, 67:109–119, 1994.
6. G.Gu H.Mansouri M.Zangiabadi, Y.Q.Bai and C.Roos. Improved full-newton step $\mathcal{O}(nL)$ infeasible interior-point method for linear optimization. *Journal Optimization Theory Application*, 145:271–288, 2010.
7. C. Roos. A full-newton step $\mathcal{O}(n)$ infeasible interior-point algorithm for linear optimization. *SIAM J. Optim*, 16:1110–1136, 2006.
8. C. Roos T. Terlaky and J.-Ph. Vial. Interior point methods for linear optimization.
9. Y.Ye. Interior point algorithms. wiley-interscience series in discrete mathematics and optimization.
10. W. Sun Z. Liu and F. Tian. A full-newton step infeasible interior-point algorithm for linear programming based on a kernel function. *Applied Mathematics and Optimization*, 60:237–251, 2009.
11. Y. Zhang. Solving large-scale linear programs by interior-point methods under the matlab environment. *Optimization Methods and Software*, 10:1–31, 1998.